Reconstructing and Selecting Electron Neutrino and Anti-Neutrino Interactions on Argon in the Short-Baseline Near Detector

Edward Tyley

Supervisor: Prof. Neil Spooner

Department of Physics & Astronomy
The University of Sheffield

This dissertation is submitted for the degree of
Doctor of Philosophy

February 2023

Abstract

As the field of neutrino physics approaches the precision era, where all of the parameters describing neutrino oscillations have been measured, the understating of neutrino-nucleus interactions must continue to develop to facilitate the next generation of oscillation experiments. This advancement relies upon developments in both the theoretical modelling and experimental measurements of these interactions. The Short-Baseline Near Detector (SBND) is a 112 ton Liquid Argon Time Projection Chamber (LAr TPC) which aims to provide such interaction measurements in the Booster Neutrino Beam (BNB), in addition to constraining uncertainties for the sterile neutrino search of the Short-Baseline Neutrino (SBN) program, once it turns on in 2023. The capability of SBND to make a measurement of ν_e Charged Current (CC) interaction cross section is assessed in this thesis. Improvements to the reconstruction chain within SBND, with a focus on electron showers, are presented which enhance the ability to both select and measure these interactions. These ν_e CC interactions are selected with a 28.1% efficiency whilst rejecting over 99% of neutrino induced and 99.999% of cosmogenic backgrounds. A procedure to perform a ν_e CC inclusive cross measurement is outlined, including an assessment of the impact of both statistical and systematic uncertainties. The power of this measurement to differentiate between neutrino interaction models is subsequently presented.

Declaration

This thesis represents original work by the author except for where specific references are made. The work presented has not been submitted for any other degree or qualification at this, or any other university. Due to the collaborative nature of modern particle physics experiments, the work relies upon work performed by SBND collaborators, as described below.

The overviews of neutrino physics and LAr TPCs given in Chapters 2 and 3 represent work which the author did not contribute, references assign credit for the work and figures presented. The overview of SBND given in Chapter 4 relies on work performed by the entire SBND collaboration, in particular the SBN proposal[1]. Plots not made by the author are labelled with a reference to the original source. The flux prediction in SBND relies upon the BNB simulation developed in MiniBooNE and the flux reader module developed by M. Del Tutto[2]. The author implemented the dirt simulation for SBND using the existing GENIE rockbox feature with the help of M. Del Tutto, R. Hatcher, and M. Roda.

The Pandora pattern recognition package was largely developed in μBooNE prior to the authors involvement, including the development of the unambiguous cosmic removal and Multi Variate Analyses (MVAs) that are presented in Chapter 5[3, 4]. However, all analysis of the performance of these algorithms and the implementation of the MVAs in SBND were performed by the author. The recursive shower growing algorithm was an original algorithm developed by the author nevertheless. Similarly, the high level reconstruction of tracks discussed in Chapter 6 relies upon existing code within the LArSoft framework[5]. The shower reconstruction was developed by the author in conjunction with D. Barker, with significant contributions from D. Brailsford, building upon existing tools by M. Wallbank and the Pandora team[6, 7].

The selection presented in Chapter 7 utilises the Common Analysis Framework (CAF) developed within SBN, based on the original framework developed initially within the NuMI Off-axis ν_e Appearance experiment (NOνA), and later the Deep Underground Neutrino Experiment (DUNE)[8]. The development of this framework was performed by many people including the author, B. Howard, D. Mendez, F. Psihas, and G. Putnam. The simulated samples used for analysing the performance of the selection and cross section extraction were produced by the SBND production team, consisting of M. Carneiro, H. Lay, J. Li, and L. Nguyen.

The selection itself relies upon many shared elements in different analyses across the SBN program. In particular, the cosmic ray rejection was largely pioneered by T. Brooks, with the flash matching developed by I. de Icaza and M. Stancari. The muon track rejection builds upon the developments for a ν_μ selection by T. Brooks and G. Putnam[9]. Similarly, the shower rejection cuts build upon the work performed by D. Barker in selecting ν_e interactions for a sterile neutrino induced ν_e appearance search[6]. Finally, the cross section measurement performed in Chapter 8 was implemented entirely by the author, within the CAFAna framework[8]. The author would like to thank D. Brailsford, R. Jones, and A. Smith for their advice on the cross section extraction procedure and M. Del Tutto and J. Zennamo for their help in understanding the flux prediction and associated systematic uncertainties.

Acknowledgements

First and foremost, my utmost gratitude goes out to my amazing partner Hannah for her unwavering support throughout my entire PhD, especially during the writing period. Similarly, I would like to extend thanks to my family, especially my parents, for their encouragement throughout my years of study. This combined support kept me going throughout difficult times and ultimately enabled the completion of this thesis.

My supervisor Neil has remained a guiding hand throughout my PhD, keeping me focused on telling a clear, coherent, and compelling story for which I am extremely grateful. Moreover, the advice and help received from Andrew, Anthony, Viktor, and Vitaly further helped to guide me through difficulties encountered along the way. I am eternally grateful to Dom Barker for both mentoring me up the steep learning curve at the start of my PhD, not least for answering a plethora of stupid questions, alongside making me feel so welcome in Sheffield, neither of which he could have been more supportive in. In that vein, I would like to further extend my gratitude to everyone else in Sheffield that has made the PhD such a positive experience, especially Andy, Pete, Warren, and Matt.

Working on SBND has been nothing but a positive experience, and there are countless people to whom I owe thanks. The support I have received from the conveners of the various working groups within SBND has been vital for the completion of this thesis. In particular, the help from Ornella and Michelle have immeasurably improved its quality. Their support during my, ill-timed, move to Fermilab is especially appreciated.

Without both the foundational work and comradeship of other PhD students working on SBND this thesis would not be possible. In particular, without the work, advice, and friendship of Dom Barker, Tom, Rhiannon, and Gray this thesis would be unrecognisable.

And last, but certainly not least, a special thanks goes out to Dom Brailsford who has supported me throughout. His mentorship permeates this entire thesis such is both the depth of his knowledge and his willingness to share it, for which no level of gushing acknowledgements can ever suffice.

Table of contents

List of figures

List of tables

Chapter 1

Introduction

Since the initial proposal and discovery of the neutrino, there have been many fascinating discoveries and developments in the field of neutrino physics. Chapter 2 provides a brief historical overview of the field of neutrino physics. A particular focus is given to the phenomenon of neutrino oscillations and the tension surrounding sterile neutrino oscillation measurements that motivates the Short-Baseline Neutrino (SBN) program. A brief overview of neutrino nucleus interactions is additionally given covering both theoretical models and existing measurements, focusing on the ν_e Charged Current (CC) interactions that are relevant for this thesis.

A historical overview of Liquid Argon Time Projection Chamber (LAr TPC) detector technology follows in Chapter 3, aiming to demonstrate their suitability as neutrino detectors alongside a review of notable experiments. The working principles of these detectors is subsequently presented, identifying the key physical processes that underpin the performance of such detectors.

The SBN program, which consists of three such LAr TPCs, is then discussed in Chapter 4. This initially discusses the physics goals of the program to search for sterile neutrino induced oscillations and subsequently the goals of Short-Baseline Near Detector (SBND), with a focus on the cross section measurements that are of particular relevance to this thesis. The Booster Neutrino Beam (BNB) in which the SBN program is situated is next discussed, outlining the design of the beam and presenting the flux through SBND. The design of SBND is next presented, discussing each of the individual subsystems that comprise the detector. Finally, the simulation and low level reconstruction that the later chapters rely upon is presented.

The Pandora pattern recognition package used by SBND is outlined in Chapter 5. An assessment of the performance of the unambiguous cosmic removal is performed, with multiple configurations considered. The implementation of Multi Variate Analyses (MVAs) to improve the decision making at key points within the Pandora workflow is then presented. Finally, a new algorithm developed to improve the clustering of showers is discussed.

The proceeding high level reconstruction chain is then discussed in Chapter 6, with dedicated pathways for tracks and showers each presented. For showers, an overview new reconstruction framework that has been developed is given, with the performance of multiple approaches analysed for each characteristic. Finally, the impact of the improvements made to the pattern recognition, as presented in the preceding chapter, on the high level reconstruction is assessed.

A selection has been developed to select ν_e CC interactions and reject backgrounds, which will be presented in Chapter 7. This begins with a discussion of the signal definition and cuts placed to ensure the reconstruction quality of the selected events. Next, the rejection of cosmogenic backgrounds is discussed, utilising each of the subsystems available within SBND. The rejection of ν_μ CC interactions via the identification of muon tracks is then presented. Finally, cuts placed on the electron candidate shower to reject photon and misidentified tracks are presented with the final results of the selection analysed.

Finally, the capabilities of SBND to perform a ν_e CC inclusive cross section measurement are assessed in Chapter 8. This begins with an overview of the cross section extraction procedure used. Next, the systematic uncertainties included in this analysis are discussed, covering both the sources of uncertainty and the formalism of their treatment. The result of the measurement is then presented, alongside a quantitative assessment of the ability of the measurement to differentiate between given interaction models. A discussion of the results, including future developments, is given finally.

Chapter 2

Neutrino Theory

Neutrinos are some of the least well understood but most interesting particles within the standard model of particle physics. In particular, they exhibit the phenomenon of neutrino oscillation, where neutrinos can change between flavours, which is one of the first observed displays of beyond the standard model behaviour. However, much is still not known about neutrinos, from their absolute mass to the parameters which underpin their interactions and oscillations.

Outlined in this chapter is a brief overview of the history of the neutrino, looking at the proposal and discovery of the different neutrino flavours. Next, neutrino oscillations will be discussed, covering the experimental anomalies, theoretical formulation, experimental verification, and current measurements of the phenomenon. The evidence for, and against, the existence of sterile neutrinos is subsequently discussed. Finally, an overview of neutrino interactions will be given, looking at the interaction mechanisms, nuclear models and existing measurements. This final section additionally motivates the continued study of interaction cross sections, the main topic of this thesis, both to probe the understanding of the interactions and nuclear physics but also to facilitate improved oscillation measurements.

2.1 The Proposal and Discovery of the Neutrino

In 1927 Ellis and Wooster showed that the energy spectra of electrons produced in “Radium E” (Bismouth-210) β decays was shown to be continuous, in stark contrast with the sharp peak expected from a 2-body decay[10]. Faced with the prospect of abandoning the fundamental principle of conservation of energy, Wolfgang Pauli took the radical step of proposing a new particle: the neutrino[11]. This neutrino would, invisibly, carry away some of the energy produced in the decay thus explaining the continuum of the electron energy spectra. Enrico Fermi subsequently published his theory of β decay in 1934, in which the newly discovered neutron would decay to a proton, an electron, and a neutrino[12].

The experimental confirmation of the existence of the neutrino then came in 1956 by Cowan, Reines, et al.[13]. Neutrinos produced in the Savannah River nuclear reactor were observed interacting in a liquid scintillator detector via inverse beta decay, described by Equation 2.1:

$$\overline{\nu}_e + p \rightarrow \beta^+ + n \tag{2.1}$$

The signature of this signal was a pair of scintillation pulses: the first pulse corresponding to the β^+ and the second, delayed, pulse corresponding to the capture of the neutron on Cadmium dissolved in the scintillator. The detection of this signal, including correlations between the observed rate and the detector power, confirmed the existence of the neutrino[13]. Furthermore, the number of neutrinos measured was consistent with the theoretical cross section predictions of the time.

In 1959, Pontecorvo questioned whether the neutrinos produced in the decay of pions to muons, ν_μ, was the same as those emitted in β decay, ν_e[14]. This was addressed by the observation of ν_μ interactions by Danby et al. in 1962[15]. This experiment used a spark chamber located in a neutrino beam generated by the decay in flight of pions to measure the number of “Tracks” and electromagnetic “Showers” produced by ν_μ and ν_e interactions respectively. An excess of muon “Tracks” was observed compared to the number of expected electron “Showers”, disproving the $\nu_\mu = \nu_e$ hypothesis.

Following the discovery of the tau lepton in 1975 the existence of ν_τ was predicted [16, 17]. This prediction was reinforced by the measurement of the Z boson decay width performed by the Apparatus for LEP PHysics (ALEPH) experiment at the Large Electron-Positron Collider (LEP) to determine the number of light, weakly interacting, non-sterile neutrino flavours[18, 19]. The results were consistent with only 3 flavours of weakly interacting neutrino existing, with a mass less than half of that of the Z boson. The tau neutrino was experimentally confirmed in 2001 when the Direct Observation of the NU Tau (DONUT) experiment used nuclear emulsion targets to identify kinked tracks indicative of tau leptons in a neutrino beam produced from the Tevatron[20].

2.2 Neutrino Oscillations

One key omission from the historical overview of neutrinos was the phenomenon of neutrino oscillation, where neutrinos are produced in a given flavour and observed in another. This phenomenon was the first observed beyond the standard model behaviour in particle physics, and is thus deserving of a dedicated section. Firstly, an overview of the experimental anomalies that hinted towards the existence of neutrino oscillations will be given. Next, the theoretical proposal and mathematical formalism of neutrino oscillations will be presented. Subsequently, the experimental verification of neutrino oscillations will be reviewed. Finally, an overview of current, and future, neutrino oscillation experiments will be given alongside the global best fits for the oscillation parameters at the time of writing.

2.2.1 Experimental Anomalies

After the observation of neutrinos from nuclear reactors and neutrino beams, efforts were made to observe the neutrinos produced by nuclear fusion in the sun. The first detection of such solar neutrinos was in 1968 by Davis et al. in the Homestake experiment[21]. This utilised the neutrino capture on Cl^{37} producing Ar^{37}, a radioactive isotope, as shown in equation 2.2:

$$\nu_e + n(Cl^{37}) \rightarrow e^- + p(Ar^{37}) \tag{2.2}$$

Subsequently, the number of argon atoms produced was counted, by observing the number of Ar^{37} decays and the flux of neutrinos was calculated. However, this measurement was in tension with the predicted flux from the contemporaneous solar fusion models.

This solar anomaly was reinforced by subsequent measurements made by the Kamioka Neutrino Detection Experiment (Kamiokande) water Cherenkov detector and GALLEX and the Soviet–American Gallium Experiment (SAGE)[22–24]. A comparison between the experimentally observed rate to the theoretical predictions for different detector technologies, each sensitive to different energy ranges, is shown in Figure 2.1 where this deficit can be observed. This compares the experimental measurements, shown as the solid black bars, compared to the theoretical predictions, broken down by production mechanism, for each detector technology. The persistence of the anomaly across different detector technologies, and thus energy ranges, strengthened the tension between experimental observations and theoretical predictions.

Meanwhile, proton decay experiments had begun to study atmospheric neutrinos as a possible background. These neutrinos are produced by the interactions of cosmic rays protons with the atmosphere, and the decays of the secondary particles produced[26]. Both the Irvine Michigan Brookhaven (IMB) and Kamiokande detectors reported a deficit in the ratio of muon neutrinos to electron neutrinos[27–30].

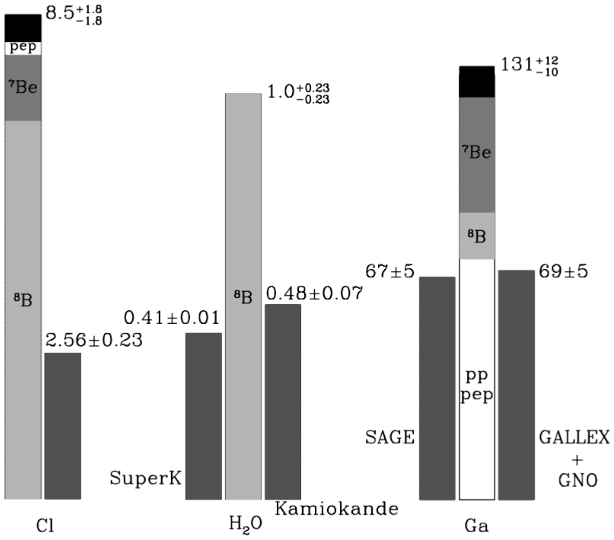


Figure 2.1: A comparison of solar neutrino rate measurements with theoretical expectations for three different detector technologies. The theoretical predictions are broken down by the fusion production mechanism each detector technology is sensitive to. Figure from [25].

The combination of the solar and atmospheric anomalies suggested that, rather than miscalculations of both neutrino fluxes, there was new behaviour being exhibited by neutrinos: the phenomenon of neutrino oscillation.

2.2.2 Theoretical Proposal and Formalism

Prior to the experimental observation of the solar and atmospheric neutrino deficit the idea of neutrinos oscillating between flavours had been proposed by Pontecorvo[31]. This was later mathematically formalised by Maki, Nakagawa and Sakata into the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix[32]. This proposed that there are two distinct forms of neutrinos: those that interact via the weak force (weak eigenstates) and those that describe the propagation of neutrinos (the mass eigenstates)

The mass eigenstates are defined such that they satisfy the Schrödinger equation, defined in Equation 2.3:

$$i\hbar\frac{d}{dt}|\nu\rangle = \hat{\mathcal{H}}|\nu\rangle = E|\nu\rangle \tag{2.3}$$

The weak eigenstates that are observed in neutrino interactions are formed from a superposition of the mass eigenstates, as shown in equation 2.4:

$$|\nu_\alpha\rangle = \sum_i U^*_{\alpha i}|\nu_i\rangle \tag{2.4}$$

The Greek and Latin letters representing weak and mass eigenstates respectively, that is $[\alpha, \beta, \gamma] \in [e, \mu, \tau]$ and $[i, j, k] \in [1, 2, 3]$. U is the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix, given in Equation 2.5:

$$\begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & C_{23} & S_{23} \\ 0 & -S_{23} & C_{23} \end{pmatrix} \begin{pmatrix} C_{13} & 0 & S_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -S_{13}e^{i\delta} & 0 & C_{13} \end{pmatrix} \begin{pmatrix} C_{12} & S_{12} & 0 \\ -S_{12} & C_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{2.5}$$

The matrix is decomposed into three separate mixing matrices each characterised by a single mixing angle θ_{ij}, with $C_{ij} \equiv cos(\theta_{ij})$ and $S_{ij} \equiv sin(\theta_{ij})$. There is an additional CP violating phase δ_{cp} that describes the difference between neutrino and antineutrino oscillations. An additional Majorana phase may be present, depending on the nature of neutrinos, but this is omitted as it does not contribute to oscillations[33].

The Schrödinger equation can be solved using a plane wave solution as described in equation 2.6, where $|\nu\rangle \equiv |\nu(t = 0)\rangle$:

$$|\nu(t)\rangle = e^{-iEt}|\nu\rangle = \sum_i U^*_{\alpha i} e^{-iE_i t}|\nu_i\rangle \tag{2.6}$$

Therefore, the probability of a neutrino of flavour α being detected as flavour β after time t is given by Equation 2.7:

$$P_{\alpha\to\beta}(t) = |\langle\nu_\beta(t)|\nu_\alpha\rangle|^2 = |\sum_{i,j} e^{-iE_j t} U^*_{\alpha i} U_{\beta j} \langle\nu_j|\nu_i\rangle|^2 \tag{2.7}$$

As $\langle\nu_j|\nu_i\rangle \equiv \delta_{ij}$, this can be reformed into Equation 2.8:

$$P_{\alpha\to\beta}(t) = |\sum_i e^{-iE_i t} U^*_{\alpha i} U_{\beta i}|^2 = \sum_{i,j} e^{-i(E_i - E_j)t} U^*_{\alpha i} U_{\beta i} U_{\alpha j} U^*_{\beta j} \tag{2.8}$$

The energy of a given mass eigenstate of a neutrino, E_i, can be given by the formula presented in 2.9, assuming $m \ll E$ to justify Taylor expansion and make the approximations $|\mathbf{p}_i| = E$:

$$E_i = \sqrt{\mathbf{P}_i^2 + m_i^2} = |\mathbf{p}_i| \cdot \left[1 + \left(\frac{m_i}{\mathbf{p}_i}\right)^2\right]^{1/2} \to E + \frac{m_i^2}{2E} \tag{2.9}$$

This results in the oscillation probability, where $\Delta m^2_{ij} \equiv m_i^2 - m_j^2$ and time and length are equivalent up to factors of c ($t = L$):

$$P_{\alpha\to\beta}(t) = \sum_{i,j} U^*_{\alpha i} U_{\beta i} U_{\alpha j} U^*_{\beta j} e^{\frac{-i\Delta m^2_{ij} L}{2E}} \tag{2.10}$$

Which can be rewritten as:

$$\begin{aligned} P_{\alpha\rightarrow\beta}(t) = \delta_{\alpha\beta} &- 4\sum_{i,j}^{i\neq j}\Re(U^{*}_{\alpha i}U_{\beta i}U_{\alpha j}U^{*}_{\beta j})sin^{2}\left(\frac{\Delta m^{2}_{ij}L}{4E}\right) \\ &+ 2\sum_{i,j}^{i\neq j}\Im(U^{*}_{\alpha i}U_{\beta i}U_{\alpha j}U^{*}_{\beta j})sin\left(\frac{\Delta m^{2}_{ij}L}{2E}\right) \end{aligned} \tag{2.11}$$

In the simplified case of two flavour oscillation the mixing matrix can be parameterised by a single angle as shown in Equation 2.12:

$$U = \begin{pmatrix} cos(\theta) & sin(\theta) \\ -sin(\theta) & cos(\theta) \end{pmatrix} \tag{2.12}$$

Thus the oscillation probability simplifies to Equation 2.13:

$$P_{\alpha\rightarrow\beta} = \sin^{2}(2\theta)\sin^{2}\left(\frac{\Delta m^{2}L}{4E}\right) \tag{2.13}$$

The maximum amount of mixing that will occur is thus determined by the mixing angle θ and the distance at which this maximum mixing occurs, referred to as an oscillation maximum, is determined by the mass splitting Δm^2, for a given neutrino energy.

Additional matter effects must be taken into account when the neutrinos are passing through matter, rather than a vacuum as previously assumed, due to the coherent scattering between the neutrinos and matter[34]. This effect enhances, or suppresses, the oscillation probability for neutrinos depending on the sign of δ_{cp}, whilst having the opposite effect on anti-neutrinos. Indeed, studying the difference between the oscillations of neutrinos and anti-neutrinos induced by this effect modern, and future, oscillation experiments can measure δ_{cp}, as will be discussed the forthcoming sections[35].

2.2.3 Experimental Verification

The first experimental observation of neutrino oscillations came from the Super-Kamiokande (SK) experiment, the successor to Kamiokande that had previously reported a deficit of atmospheric neutrinos[36]. SK relied upon the same water Cherenkov technology as its predecessor which allows for discrimination between electrons and muons based on the shape of the light cone produced. This demonstrated the zenith angular and energy dependence of the deficit of atmospheric muon neutrinos consistent with neutrino oscillations, as shown in Figure 2.2. This shows the Up/Down asymmetry of the observed rate of neutrinos as a function of the momentum for both electrons and muons. The electrons exhibit no asymmetry con-

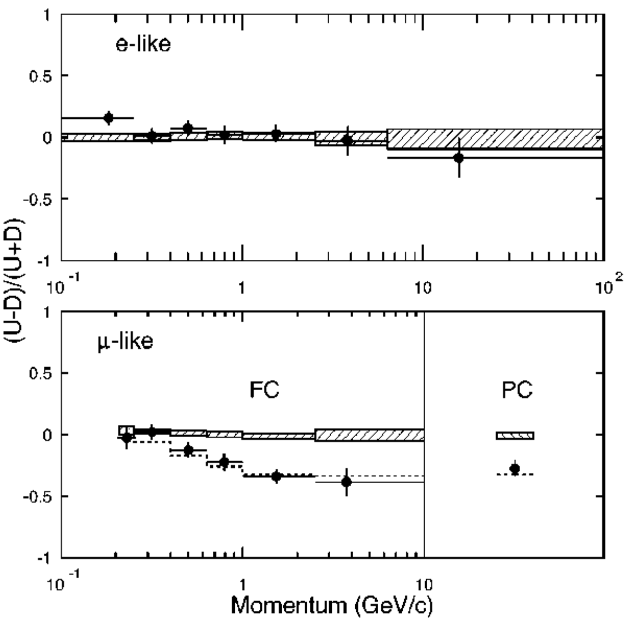


Figure 2.2: Up/Down asymmetry of atmospheric neutrinos measured by SK as a function of reconstructed energy for Fully Contained (FC) events, the integrated Partially Contained (PC) rate is also shown. Shown for both electron-like events (top) and muon-like events (bottom) compared to simulation without oscillations (hashed area) and to the best fit oscillation parameters (dashed line) of $\sin^2(2\theta) = 1.0$ and $\Delta m^2 = 2.2 \times 10^{-3}$ eV^2. Figure from [36].

sistent with Monte Carlo (MC) expectations without oscillations shown in the shaded boxes. In contrast, the muons demonstrate a strong asymmetry in disagreement with the no oscillation hypothesis with the best oscillation parameter fit shown in the dashed line ($\sin^2(2\theta) = 1.0$, $\Delta m^2 = 2.2 \times 10^{-3}$ eV^2).

The Sudbury Neutrino Observatory (SNO) subsequently provided further confirmation of neutrino oscillations in solar neutrinos[37]. This experiment also utilised a water Cherenkov detector but with added Deuterium to facilitate the measurement the rate of Charged Current (CC), Neutral Current (NC) and Electron Scatter (ES) interactions. Each of these interactions has a different sensitivity to the ν_e vs $\nu_{\mu,\tau}$ flux: CC is sensitive only to ν_e, NC is sensitive to all flavours equally and ES sensitive to all flavours but with enhanced sensitivity to ν_e. By combining the measurements from these three channels the total flux of neutrinos could be found, separated into ϕ_e and $\phi_{\mu\tau}$ as shown in Figure 2.3. The three interaction modes are shown in red (CC), blue (NC), and green (ES) alongside the theoretical prediction of the SSM shown as the dashed diagonal line. The best fit between these channels found that a third of the total solar flux arriving was ν_e whilst the remainder had oscillated into ν_μ and ν_τ.

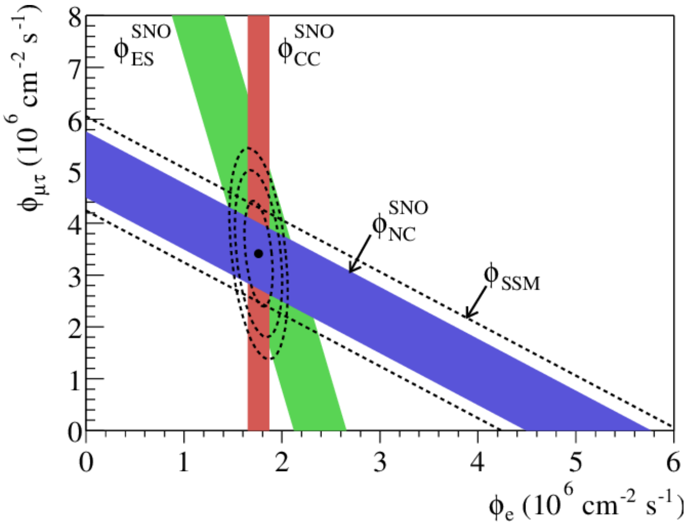


Figure 2.3: Solar neutrino flux measurements by Sudbury Neutrino Observatory (SNO) parameterised in the ν_e flux (ϕ_e) and the ν_μ and ν_τ flux ($\phi_{\mu\tau}$). The red, blue and green lines show the allowed regions for the Charged Current (CC), Neutral Current (NC) and Electron Scatter (ES) channels respectively compared to theoretical predictions of the Standard Solar Model (SSM) shown as diagonal dashed lines. Figure from [37].

2.2.4 Current Oscillation Parameter Measurements

Since the discovery of neutrino oscillation, many experiments have attempted to measure the underlying parameters which characterise the phenomenon; namely the mass splittings, the mixing angles, and δ_{cp}. Experiments typically measure the appearance or disappearance, or even both simultaneously, of a particular weak eigenstate from a neutrino source, including solar, atmospheric, reactor, and neutrino beams. The oscillation parameters that a given experiment is sensitive to depends on the energy and baseline of the observed neutrinos. Global fits subsequently combine the results from multiple experiments across these various sources to extract the entire ensemble of mixing parameters simultaneously.

Despite recent efforts, the mass of the neutrino has not yet been measured, with the current direct detection limit being $m_\nu < 0.8$ eV from the Karlsruhe Tritium Neutrino Experiment (KATRIN)[38]. Whilst neutrino oscillation experiments cannot measure the mass directly they are able to measure the mass splitting, the difference between the mass eigenstates squared.

Recent measurements from oscillation experiments, which will be discussed in more detail in the next section, have demonstrated that the "atmospheric mass splitting" $|\Delta m^2_{32}| \sim |\Delta m^2_{31}| \sim 10^{-3}$ and the "solar mass splitting" $\Delta m^2_{21} \sim 10^{-5}$. However, the sign of the atmospheric mass splitting is currently unknown and thus the hierarchy could either be "Normal", $m_3 > m_2 > m_1$, or "Inverted", $m_2 > m_1 > m_3$, as shown in Figure 2.4. There is some degeneracy between the neutrino mixing parameters (the mixing angles and δ_{cp}) and the mass hierarchy so many of the measurements of the mixing parameters are given for both hierarchies.

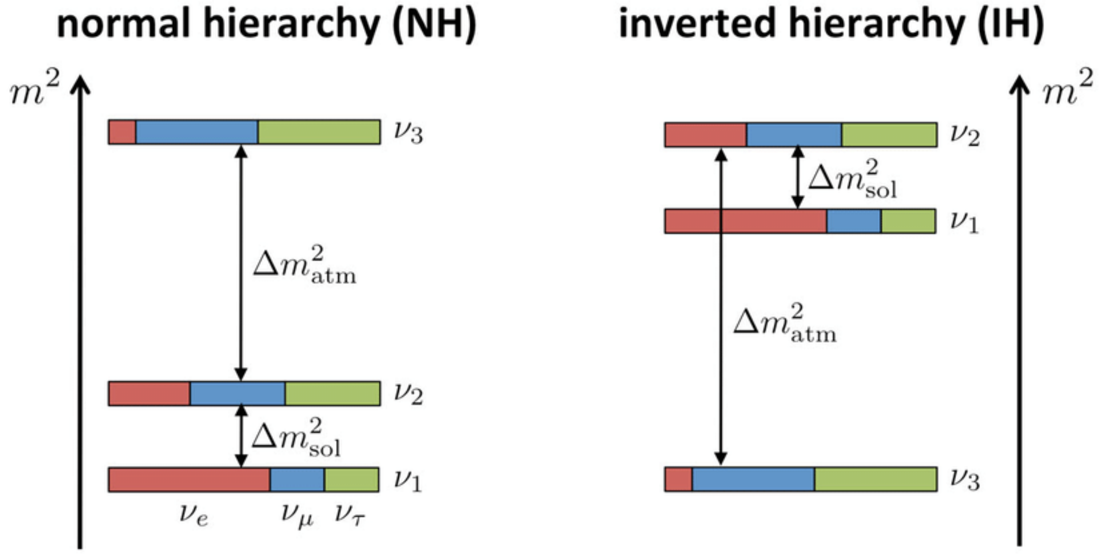


Figure 2.4: Diagram demonstrating the mass splitting between the neutrino mass eigenstates and the relative contributions of each weak eigenstates. The normal (inverted) hierarchy is shown where ν_3 is the heaviest (lightest) of the neutrinos depending on the sign of Δm^2_{atm}. Figure from [39].

Solar neutrino experiments are typically sensitive to θ_{12} and Δm^2_{12}. Recent results from SK, SNO and Borexino provide the strongest constraints in the latest global analysis[40–42]. KamLAND studies reactor neutrinos over a long baseline of 180 km and as such is sensitive to the same oscillation parameters as the solar experiments, θ_{12} and Δm^2_{12}[43]. These complimentary measurements of the solar oscillation parameters are in general agreement and thus strengthen the combined constraint on these parameters. Shorter baseline reactor experiments, such as Daya Bay, Reno and Double Chooz, are typically more sensitive to θ_{13} and a combination of Δm^2_{32} and Δm^2_{31}[44–46].

Long baseline neutrino experiments that make use of accelerator based neutrino beams are typically sensitive to a number of the oscillation parameters, including both the solar and reactor mixing angles. Matter effects are particularly important due to the long baseline of the neutrinos passing through the ground and thus these experiments have additional sensitivity to the neutrino mass hierarchy. By comparing oscillation measurements between neutrino and anti-neutrino enhanced beam modes these experiments can also search for matter-antimatter asymmetry, quantified by δ_{cp}. Current experiments, Tokai to Super-Kamiokande (T2K) and the NuMI Off-axis

ν_e Appearance experiment (NOνA), show some hints and preferences for normal ordering but tensions exist in the measurement of δ_{cp}[35, 47].

The oscillation parameters θ_{23} and Δm^2_{32} are primarily determined by experiments measuring atmospheric neutrinos, produced by cosmic rays interacting in the atmosphere. Specifically, SK and the DeepCore detector of the IceCube experiment are included in the latest global analyses[48, 49].

The results from these experiments are then combined into a global fit by the NuFIT group, the results of which are shown in Figure 2.5. This shows the allowed regions for each of the mixing angles, mass splittings and δ_{cp}. This demonstrates the preference for "Normal" mass ordering over "Inverted" by lower χ^2 for $\Delta m^2_{31} > 0$ compared to $\Delta m^2_{32} < 0$. The determination of the $\sin^2\theta_{23}$ octant (> 0.5 or < 0.5) and value of δ_{cp} are poorly constrained, largely due to tensions between T2K and NOνA.

Next generation long baseline experiments, specifically Tokai to Hyper-Kamiokande (T2HK) and the Deep Underground Neutrino Experiment (DUNE), aim to conclusively measure these parameters[51, 52]. Meanwhile, the Jiangmen Underground Neutrino Observatory (JUNO) is a next generation reactor neutrino experiment that will study neutrinos across a variety of baselines and measure the mass hierarchy and reactor oscillation parameters[53]. As such, the precision era of neutrino oscillation measurements is rapidly approaching.

2.3 Sterile Neutrinos

Whilst data from the LEP experiments indicates that there are three flavours of light weakly interacting neutrinos, some experiments have hinted towards a fourth type of neutrino: the sterile neutrino [18, 19]. This sterile neutrino would not interact via the weak force like the other neutrino flavours and thus would not be directly observable. However, neutrinos produced in weak interactions could oscillate to these sterile neutrinos via an extension to the PMNS matrix to include this fourth neutrino, that is extending the 3×3 matrix to 4×4. The effect of this additional neutrino could then be measured by oscillation experiments, manifesting as either enhanced disappearance or appearance[54, 55].

The first experiment to report observation of such oscillations was the Liquid Scintillator Neutrino Detector (LSND) that measured $\overline{\nu}_e$ appearance from a stopped muon source producing $\overline{\nu}_\mu$[56]. LSND observed a 3.8σ excess of $\overline{\nu}_e$ interactions within the detector consistent with sterile neutrinos with a mass splitting Δm^2 in the range of $0.2 - 10$ eV^2/c^4.

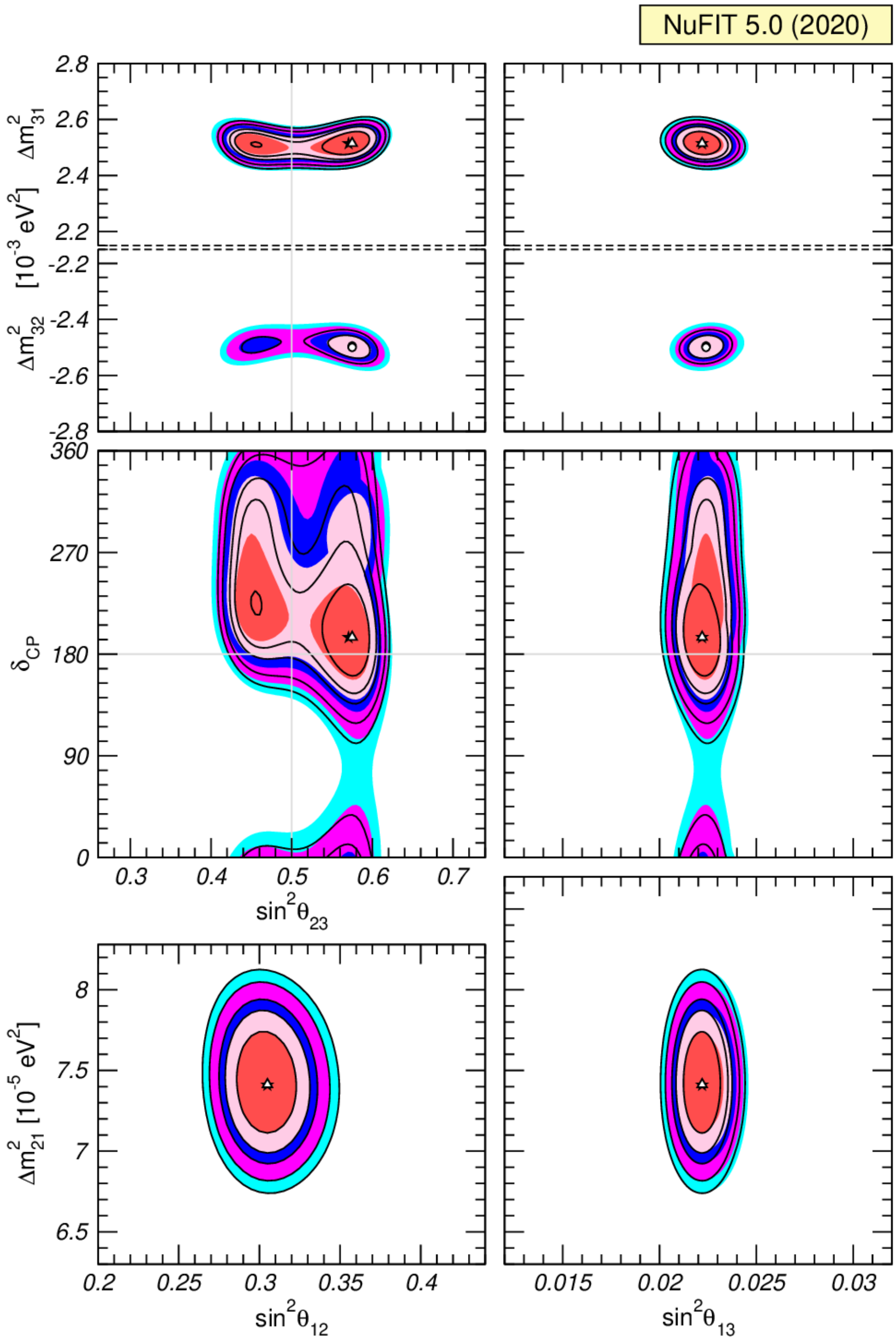
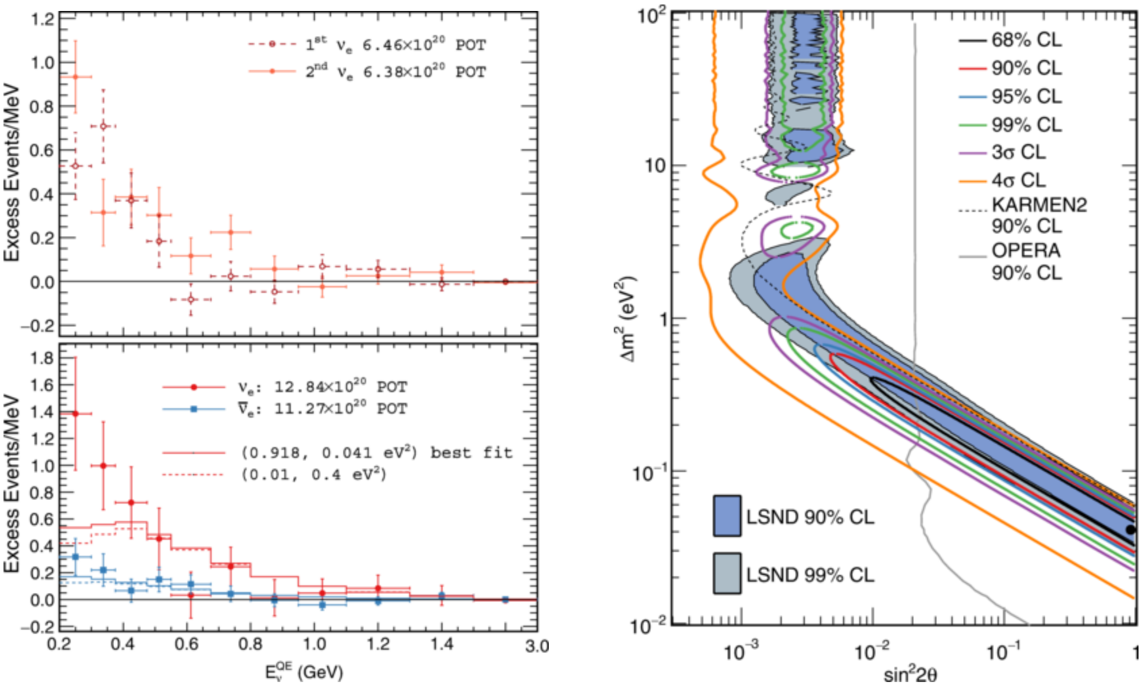


Figure 2.5: Global best fit contours of neutrino oscillation parameters from NuFIT 5.0. Δm^2_{31} and Δm^2_{32} are shown for the normal and inverted hierarchy respectively, with normal assumed for the remaining parameters. The black (coloured) contours represent to 1σ, 90%, 2σ, 99% and 3 σ confidence limits with (without) SK atmospheric data. Figure from [50].

MiniBooNE subsequently investigated this anomalous excess in the Booster Neutrino Beam (BNB)[57, 58]. This beam primarily produces neutrinos by decay in flight of pions produced from protons incident on a beryllium target. The MiniBooNE detector uses mineral oil as a target and detects both the scintillation and Cherenkov light produced by the neutrino interactions. An excess of low energy ν_e ($\overline{\nu}_e$) interactions was observed at low energy when the BNB was run in ν ($\overline{\nu}$) mode, although a larger excess was observed in ν mode.

The excess rate from MiniBooNE is shown in Figure 2.6 as a function of the reconstructed neutrino energy, alongside the best fit prediction from the sterile neutrino hypothesis. Overall, the combined MiniBooNE excesses in both ν and $\overline{\nu}$ mode correspond to an overall significance of 4.8σ with an allowed parameter space for sterile neutrino oscillation shown in Figure 2.6 alongside the LSND allowed region which is in general agreement.

Such sterile neutrinos would not only manifest as a ν_e appearance but additionally corresponding ν_e and ν_μ disappearances would be expected. Some reactor experiments have shown signs of a ν_e disappearance as demonstrated by the global fit shown



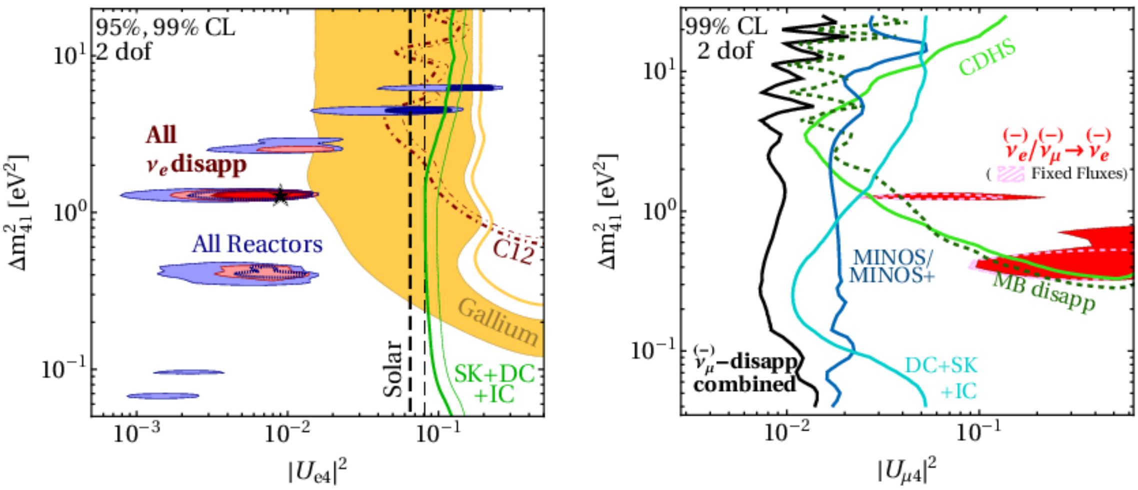
(a) MiniBooNE Low Energy Excess

(b) MiniBooNE NuE Appearance

Figure 2.6: The excess of events obsererved at MiniBooNE is shown on the left consistent across multiple running periods (top) and between both neutrino and anti-neutrino beam modes (bottom) alongside the best fit parameters. Figure from [57]. The ν_e appearance allowed region is shown on the right with imposed limits from KARMEN and OPERA and allowed region from LSND[59, 60]

in Figure 2.7[61]. This compares the allowed regions of phase space compared to exclusion contours from solar and atmospheric experiments.

However, several experiments have found no signs of ν_μ disappearance creating a strong tension with the ν_e appearance and disappearance measurements, as shown in Figure 2.7. This shows the exclusion contours for the ν_μ disappearance alongside the allowed region for the MiniBooNE low energy excess, using the ν_e disappearance best fit point to map between the parameter spaces. The global fit combines ν_μ disappearance data from accelerator based experiments, including Main Injector Neutrino Oscillation Search (MINOS)(+) and MiniBooNE, with atmospheric data, from Super-Kamiokande (SK), IceCube and DeepCore[49, 62–65]. The overlap between the MiniBooNE allowed region and the ν_μ disappearance exclusion contours demonstrates the large tension between the appearance and disappearance measurements.



(a) NuE Disappearance

(b) NuMu Disappearance

Figure 2.7: Global combined ν_e (left) and ν_μ (right) disappearance limits. The MiniBooNE ν_e appearance allowed region is also projected into the disappearance phase space, demonstrating the tension between the appearance and disappearance measurements. Figure from [61]

A series of new experiments are aiming to improve the current experimental limits and resolve the tensions between the different measurements. The J-PARC Sterile Neutrino Search at J-PARC Spallation Neutron Source (JSNS^2) experiment will directly probe the LSND measurement by searching for anomalous $\overline{\nu}_e$ appearance from the same sort of stopped muon source[66]. Similarly, the Short-Baseline Neutrino (SBN) program aims to directly probe the MiniBooNE low energy excess in the same beam, the BNB[1]. The SBN program aims to use multiple detectors of the same technology in the same beam to constrain systematic uncertainties and measure the ν_e appearance and ν_μ disappearance channels simultaneously. Initial results from μBooNE, the first of the SBN detectors to be built, have shown no sign of the excess[67, 68]. The SBN program will be discussed in detail in Chapter 4, with a particular focus on the Short-Baseline Near Detector (SBND).

Similarly, several next generation reactor neutrino experiments are expected to improve the ν_e disappearance sensitivity but are not included in current global fits. This includes movable experiments which can vary the baseline, including DANSS and Neutrino4, which reduces systematic uncertainties arising from detector calibration[69, 70]. Experiments at small research reactors allow for shorter baselines to be achieved, compared to larger commercial reactors, yielding improved sensitivity at high masses, examples include the Precision Reactor Oscillation and Spectrum Experiment (PROSPECT) and Search for Sterile Reactor Neutrino Oscillations (STEREO)[71–73]. Meanwhile, long baseline accelerator experiments such as NOνA and T2K continue to improve the ν_μ disappearance constraints[74, 75]. Experiments, in particular IceCube, also continue to improve sensitivity by searching for disappearance of atmospheric neutrinos[76].

This next generation of experiments aims to conclusively address the tension that exists between the appearance and disappearance channels.

2.4 Neutrino-Nucleus Interactions

Studying neutrino interactions can provide valuable insight to the nature of the weak force and nuclear structure, typically performed by measuring of the cross section which describes how likely a neutrino is to interact. Furthermore, understanding these interactions is vital to unlock the potential of neutrino oscillation measurements, as these constitute the dominant uncertainties in current long baseline analyses[77, 78]. Most oscillation experiments utilise both near and far detectors to constrain the corresponding systematic uncertainties. However differences in neutrino flux and detector acceptance prohibits complete cancellation, particularly in detectors with different nuclear targets in their near and far detectors. Additionally, failing to account for nuclear effects, in particular multi-nucleon processes, can skew estimation of neutrino energy[79].

Furthermore, many ν_e appearance experiments rely on assumptions on the ν_e/ν_μ cross section ratio due to the scarcity of ν_e cross section measurements[80]. Measurements of the ν_e cross section are particularly hard as neutrino beams predominantly produce ν_μ, resulting in measurements with significant background contributions and limited statistics[81, 82]. Thus it is vital for cross sections to be measured for the next generation oscillation experiments to achieve their physics goals.

The modelling of neutrino-nucleus interactions typically factorises the process into nuclear models, primary interactions, and hadron production and transport. These individual models are typically combined by event generators, for example GENIE, which produce comprehensive predictions that can be compared to, and tuned on, experimental data[83]. This modularity allows model building by combining the individual elements to be simple, such that the best combination of nuclear, interaction, and hadron models can be found. The combined cross section of these interaction modes, compared to a compilation of data, is shown in Figure 2.8 for typical accelerator neutrino energies. This demonstrates the often limited statistics of historical measurements, particularly for $\overline{\nu}$.

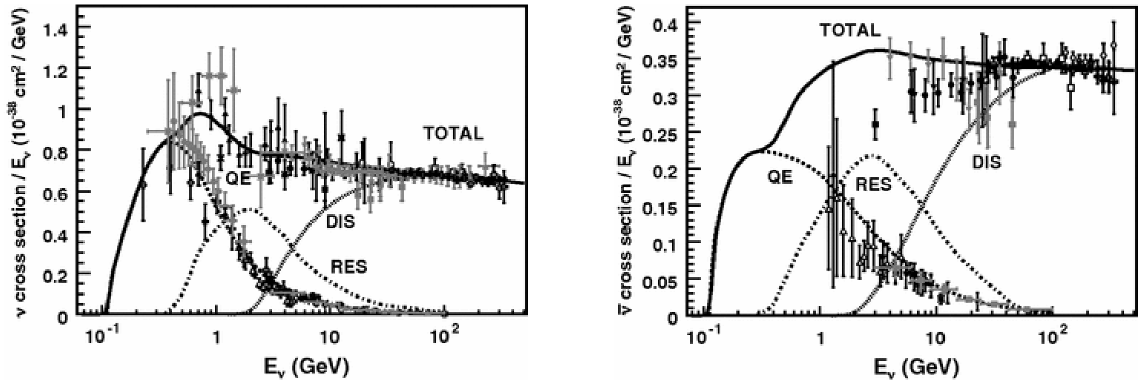


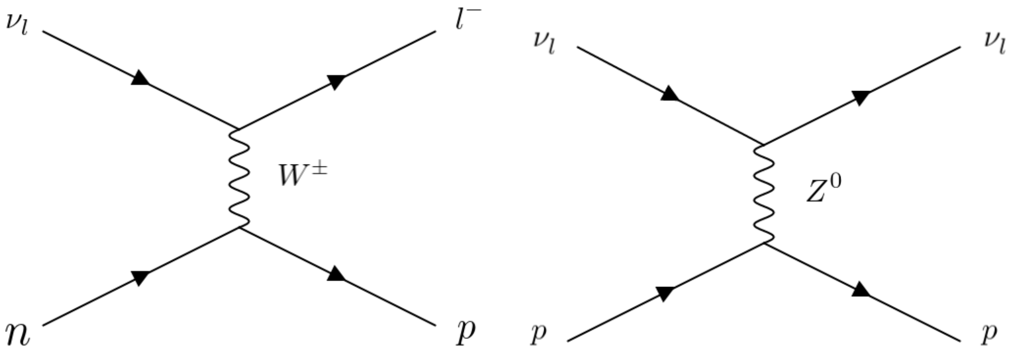
Figure 2.8: Summary of neutrino and anti-neutrino cross section data. The total cross section is shown in addition to the contributions from the individual interaction channels, alongside exclusive channel measurements. Figure from [84]

2.4.1 Interaction Modes

The primary interactions are often separated into discrete processes based on the particles produced in the interaction, each relevant for a different neutrino energy as shown in Figure 2.8. A brief description of each of these interaction modes is subsequently presented.

(Quasi-)Elastic Scattering

Below ~ 1 GeV the dominant process is Charged Current (CC) (Neutral Current (NC)) Quasi-Elastic (QE) (Elastic) scattering where the neutrino interacts with a single nucleon as depicted in Figure 2.9. These are the simplest types of interactions to model and thus can be used to probe fundamental form factors and nuclear models. This process is most commonly modelled using the Llewellyn-Smith formalism which relies on the Axial Mass (M_a) and form factors which describe the charge density within the nucleus[85].



(a) CCQE (b) NCE

Figure 2.9: Charged Current Quasi-Elastic and Neutral Current Elastic Feynman Diagrams.

Multi-Nucleon Interactions

Many early neutrino cross section experiments used hydrogen or deuterium targets where Quasi-Elastic (QE) scattering dominated. However, more recent neutrino experiments typically use heavier, more complicated nuclear targets such as carbon or argon. In these cases, extensions are required to the simple QE models to account for correlations between nucleons, such as Meson Exchange Currents (MEC). In these multi-nucleon interactions the neutrino, rather than interacting with a single nucleon, it interacts with some number, most commonly a pair, of interacting nucleons often referred to as n-particle n-hole (np-nh) interactions. These interactions are particularly important in the few-GeV region. An example 2p-2h interaction is depicted in Figure 2.10.

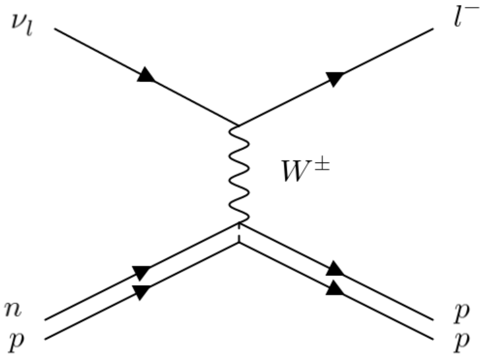


Figure 2.10: 2p-2h Feynman Diagram

These MEC interactions were initially modelled in GENIE by an empirical model tuned to electron scattering data[86]. In contrast, the Valencia model incorporates both QE and multi-nucleon processes whilst additionally including polarisation and coulomb corrections[87].

Pion Production

As more energy is transferred to the nucleus, higher Q^2, the possibility of the nucleus being excited increases. Nucleons are typically excited to a baryon resonance, e.g. a Δ^{++}, that quickly decays to, typically, a pion and a nucleon. These are referred to as Resonant Scattering (RES) interactions, an example of which is shown in Figure 2.11. This is historically modelled by the Rein-Sehgal model which describes both the production and decay of these resonances[88]. The Berger-Sehgal model improves upon this by including the effects of the lepton mass suppression[89].

Similarly, at lower momentum transfers it is possible that the neutrino interacts not with a single nucleon but with the entire nucleus coherently, referred to as Coherent Scattering (COH). The excited nucleus then decays to a lower energy state often producing photon or pions in the final state. These are negligible in GeV scale experiments such as SBND however.

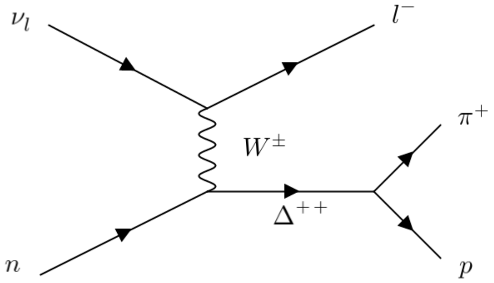


Figure 2.11: Resonant pion production Feynman Diagram

Deep Inelastic Scattering

At the highest energy interactions the neutrinos interact with a single parton within a nucleon, as opposed to the nucleon as a whole as previously. This Deep Inealstic Scattering (DIS) typically breaks apart the nucleon and produced a hadronic shower, denoted by X in Figure 2.12. This is typically modelled by the Bodek-Yang (BY) model which uses Parton Distribution Fuctions (PDFs) to describe the nucleon constituents[90]. The kinematics of the outgoing particles are subsequently modelled using the Andreopoulos-Gallagher-Kehayias-Yang (AGKY) hadronisation model[91].

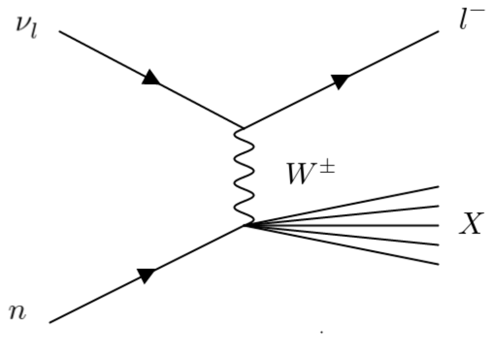


Figure 2.12: Deep Inelastic Scattering Feynman Diagram

2.4.2 Nuclear Effects

Nuclear effects must also be taken into account when simulating and measuring neutrino interaction cross sections, particularly when looking at the particle kinematics. In recent years, electron scattering data has increasingly been used to probe and refine the nuclear models, although these are unable to probe some axial form factors[84, 92, 93]. Many models rely upon the impulse approximation whereby a neutrino interacts with a stationary, free nucleon. However, this is not applicable to the heavy nuclei used in most modern neutrino experiments[84].

The basis for many modern models is the Relativistic Fermi Gas (RFG) model where the nucleons are treated as a series of non-interacting particles[94]. As nucleons obey the Pauli exclusion principle they cannot occupy the same quantum state and thus fill up energy levels from lowest to highest energy within a potential well. This results in “Fermi motion” where the nucleons have non-zero momentum up to a maximum referred to as the “Fermi momentum”, p_f. Additionally, the exclusion principle stipulates that a nucleon is unable to be excited to a state that is already occupied by another nucleon, referred to as Pauli blocking, which reduces the cross section.

This model does not include any correlations or interactions between nucleons however and thus describes data poorly for heavy nuclei. An extension of this model is the Local Fermi Gas (LFG) model where the potential a nucleon experiences is dependent on its radial position which has shown better agreement with electron scattering data[93]. Further extensions to include Short Range Correlations (SRC) between nucleons are included in a Correlated Fermi Gas (CFG) model which introduces a high momentum tail for nucleons above p_f. A comparison of the proton momentum distributions predicted by each of these models is shown in Figure 2.13, demonstrating the dramatic difference between the models.

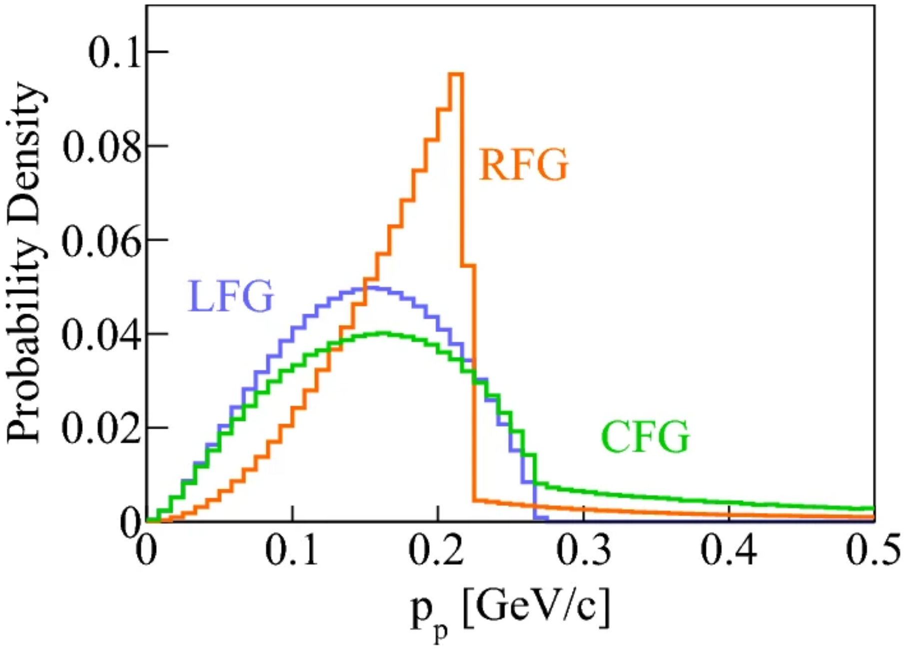


Figure 2.13: Comparison of the expected proton momentum from different GENIE nuclear models. Figure from [94]

As nucleons exist within a potential well, they require some amount of energy to be liberated from the nucleus, typically referred to as the removal energy[95]. This removal energy must be taken into account when attempting to reconstruct the neutrino energy from the particles produced[95].

Correlations between nuclei, both short and long range, are often included via the Random Phase Approximation (RPA) and are particularly important for heavy nuclei[96]. For example MEC interactions, where two nucleons interact via the exchange of a virtual meson, have been shown to contribute significantly to the multi-nucleon 2p-2h interactions[97].

Particles produced by a neutrino interacting within a heavy nucleus must propagate through the nucleus. It is possible for these particles to interact with nucleons during this propagation. Such interactions can dramatically effect the kinematics of outgoing particles and in the most extreme cases some particles may never leave the nucleus. Figure 2.14 depicts some examples of the possible processes where a pion produced in a neutrino interaction could undergo Final State Interactions (FSI)[98]. If a pion elastically scatters before leaving the nucleus the kinematics measured in the detector will be different to the initial kinematics, providing a bias in any measurements made. Similarly, absorption, charge exchange, or pion production will create a disconnect between the true neutrino interaction and observed final state multiplicity. These are currently implemented in GENIE by the INTRANUKE model for intra-nuclear cascades[99].

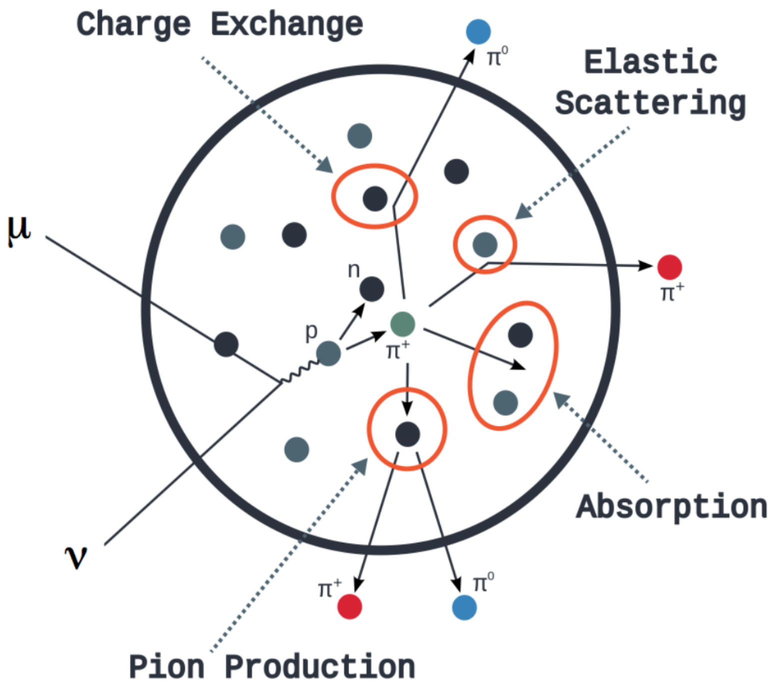


Figure 2.14: Diagram of the potential Final State Interactions (FSI) state interactions that could occur when a pion is produced in a neutrino interaction. The pion could undergo charge exchange, absorption, or pion production, all of which will change the observed kinematics of the interaction. Figure from [98]

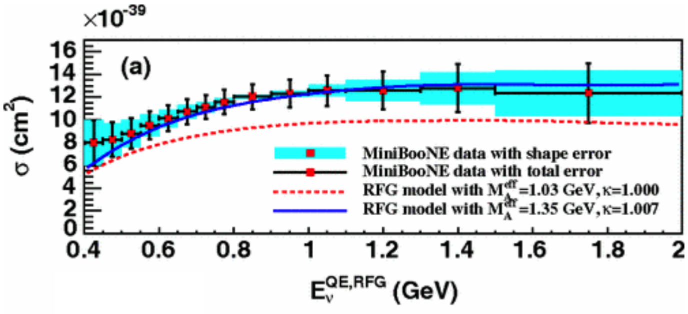
As such, experimental measurements are typically reported in terms of the final state multiplicity rather than the interaction type. For example, experiments often report $1p0\pi$ observed events as “QE-Like” which are expected to contain both true QE interactions and other interaction modes that have undergone such FSI. Details of such measurements, and additional kinematic variables which can distinguish these sorts of events, will be discussed in the forthcoming section.

2.4.3 Existing Measurements

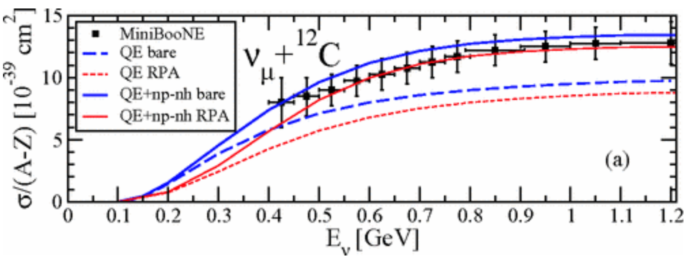
Measurements of the observed interaction cross section have been vital in the development and tuning of the interaction and nuclear models previously discussed. These measurements can either be integrated or differential against a given variable, or even multiple variables, most commonly the energy and angular distribution of the observed particles. Similarly, these can be inclusive, including all interactions, or exclusive, where a requirement is placed on the number of hadrons produced. Over time, the shift has been to more differential and exclusive final state measurements to break degeneracies is the complex parameter space associated with heavier nuclear models. This advancement relies upon modern detectors with improved Particle IDentification (PID), resolution, and neutrino interaction rates.

Early bubble chamber experiments were the first to measure the neutrino cross sections[100, 101]. These typically used hydrogen and deuterium targets and thus were largely unaffected by many of the complications due to nuclear effects previously discussed. As such, QE interactions were largely parameterised by a single interaction parameter: Axial Mass (M_a). A global average measurement of $M_a = 1.026 \pm 0.021$ GeV was found, in good agreement with pion decay data[102].

Subsequent experiments moved away from the light targets to use heavier, more complicated nuclei. For example, the MiniBooNE experiment used mineral oil (CH_2) in the Booster Neutrino Beam (BNB)[58]. When measuring the cross section for ν_μ QE interactions events were selected with a single muon track, and no pions. When fitting M_a to the data, as shown in Figure 2.15a, the best fit value was in strong tension with measurements of the axial mass from the bubble chamber experiments. However, the development of multi-nucleon, np-nh, interaction models was able to provide an alternative explanation, resolving the tension, as can be seen in Figure 2.15b. This shows that the excess above the bare QE cross section is well explained when including np-nh interaction modes.



(a) MiniBooNE data fitted with M_a



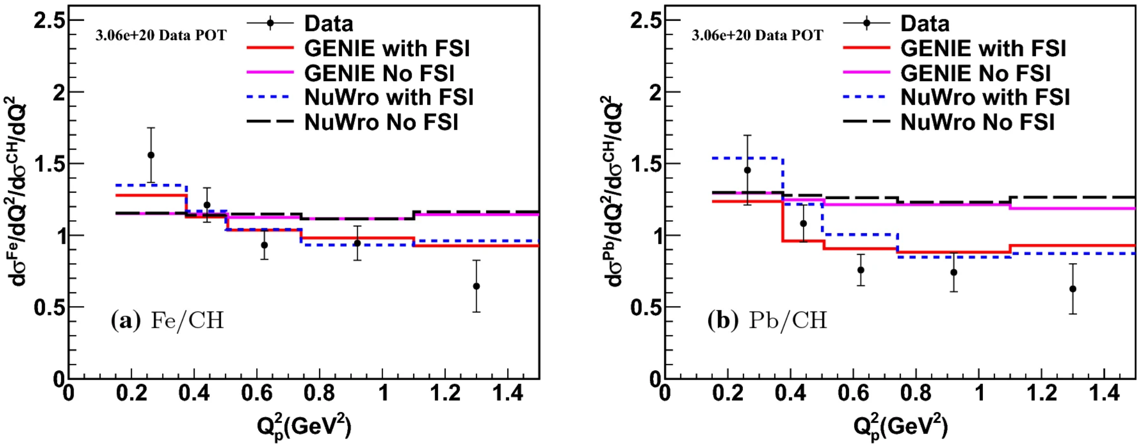
(b) MiniBooNE data fitted with 2p-2h

Figure 2.15: MiniBooNE cross section measurements with a fitted M_a (top) and with the inclusion of np-nh and RPA (bottom). Both explanations show a good agreement with the data demonstrating the degeneracy between models and parameters. Figures from [103, 104]

This demonstrated the importance of modelling interactions between nucleons and FSI state interactions for detectors using complex nuclear targets. In particular, the importance of the distinction between the observed final state from the underlying interaction mode is increasingly important.

Main Injector Neutrino ExpeRiment to study ν-A interactions (MINERνA) is a dedicated cross section experiment capable of running in both "low energy" (3.5 GeV) and "medium energy" (6 GeV) modes[105–107]. The detector consists of a central tracking region of segmented plastic scintillator strips surrounded by calorimeters made of interleaved steel and plastic scintillator layers. A target region containing layers of water, carbon, iron and lead affords the ability to measure cross sections on multiple nuclear targets, directly probing the nuclear dependence of neutrino interaction models[108]. This dependence is shown in Figure 2.16 which shows the ratios of the cross section for iron and lead compared to carbon. The importance of FSI for these heavy nuclei is shown by comparing the predictions of generators with and without these effects. By studying the kinematic imbalance transverse to the beam direction MINERνA was additionally able to probe nuclear effects such as Fermi motion and binding energy[109].

MINERνA was the first experiment to measure the ν_e QE-like cross section on hydrocarbons, as shown in Figure 2.17[82]. This shows the QE cross section differential in momentum transfer, left, finding general agreement with generator prediction. Furthermore, a comparison of the cross sections for ν_e and ν_μ was performed for similarly selected QE interactions providing a direct probe of assumptions made by oscillation experiments, also finding good agreement with predictions. This



(a) Fe/CH (b) Pb/CH

Figure 2.16: MINERνA measurements of the relative cross section between iron (left) and lead (right) compared to carbon. Several neutrino generator predictions are shown with and without FSI to demonstrate the importance for heavy nuclei. Figure from [110].

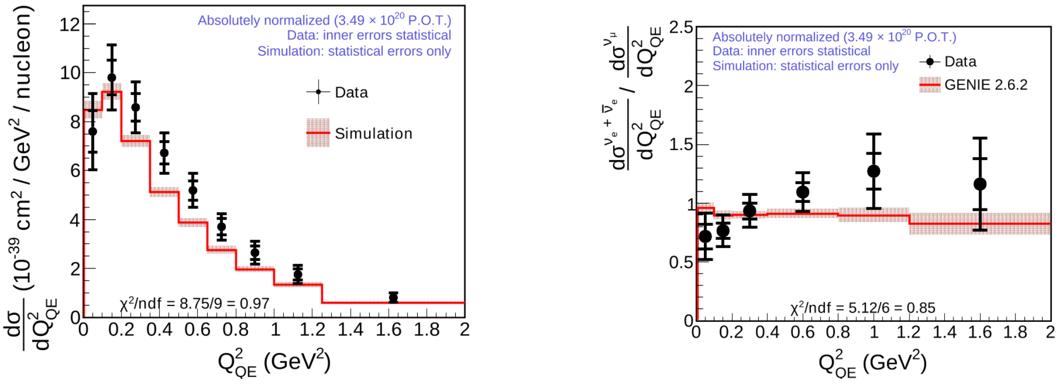


Figure 2.17: MINERνA measurements of the combined ν_e and $\overline{\nu}_e$ cross section against momentum transfer (left) and the ratio compared to ν_μ (right). Both the standalone measurement and the ratio are consistent with the generator predictions. Figure from [82].

measurement was statistically limited compared to ν_μ measurements, containing only 2105 selected events with an expected purity of 52%.

Long baseline experiments often utilise near detectors to constrain flux and interaction systematic uncertainties. These near detectors are, by design, well suited to measuring cross sections, a notable example being the T2K Near Detector at 280 m (ND280). This off-axis detector has a peak energy of 0.6 GeV, comparable to the BNB[111]. Measurements have been made of the ν_μ double-differential cross section, alongside exclusive final state measurements both with and without pions[81, 111, 112]. Similarly to MINERνA, measurements of the transverse kinematic imbalance have also been performed to probe nuclear effects[113].

Measurements of both the ν_e and $\overline{\nu}_e$ cross section were also made as shown in Figure 2.18, exploiting both ν and $\overline{\nu}$ enhanced beam modes (FHC and RHC

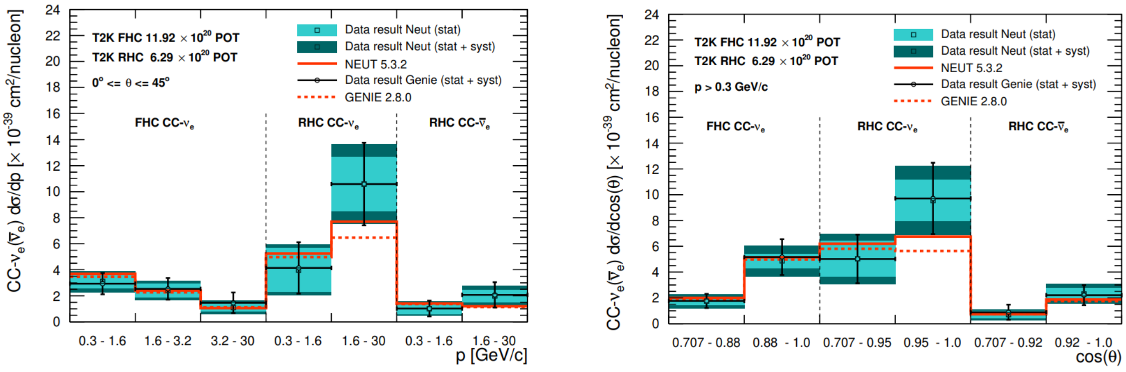


Figure 2.18: T2K ND280 ν_e and $\overline{\nu}_e$ differential cross section in the lepton energy (left) and angle (right), for limited phase space. Overall, agreement is seen between the data and both NEUT and GENIE generators. Figure from [81].

respectively) [114]. Overall, good agreement is found with both the NEUT and GENIE generators. Due in part to the limited phase space of the measurement, this sample consists of only 968 selected events with a purity of around 50%.

Liquid Argon Time Projection Chambers (LAr TPCs) are often referred to as "modern bubble chambers" due to the millimetre level resolution and low energy particle tracking thresholds[115]. As such, they have the capability to make precision measurements of particle kinematics with excellent particle identification capabilities for selecting exclusive final states.

The Argon Neutrino Teststand (ArgoNeuT) experiment was the first LAr TPC to make a cross section measurement on argon in the Neutrinos at the Main Injector (NuMI) beam[116, 117]. Due to its small size, measurements are largely statistically limited and muons were generally not contained in the detector, although the downstream MINOS near detector was used as a spectrometer. Nevertheless, ArgoNeuT demonstrated the potential for LAr TPCs to measure exclusive final states with excellent PID and detection thresholds, crucial for refining cross section and nuclear models[117, 118].

The first measurement of the ν_e cross section on argon was made by ArgoNeuT but this was extremely statistically limited with only 13 selected events[119]. Additionally, due to the small size of the detector these events were typically uncontained thus no energy measurement was available. Nevertheless, this demonstrated the potential of LAr TPCs to effectively differentiate between electrons and photons.

μBooNE, the first of the SBN detectors to take data, was the next LAr TPC to measure cross sections, utilising both the on-axis BNB and off-axis NuMI beams. Several measurements have now been made including both inclusive and exclusive

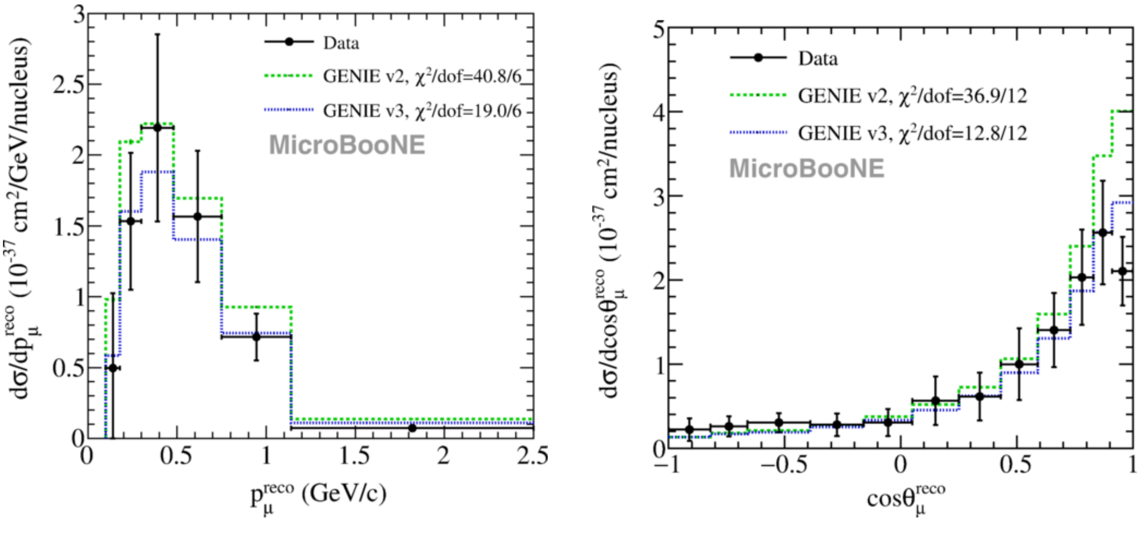
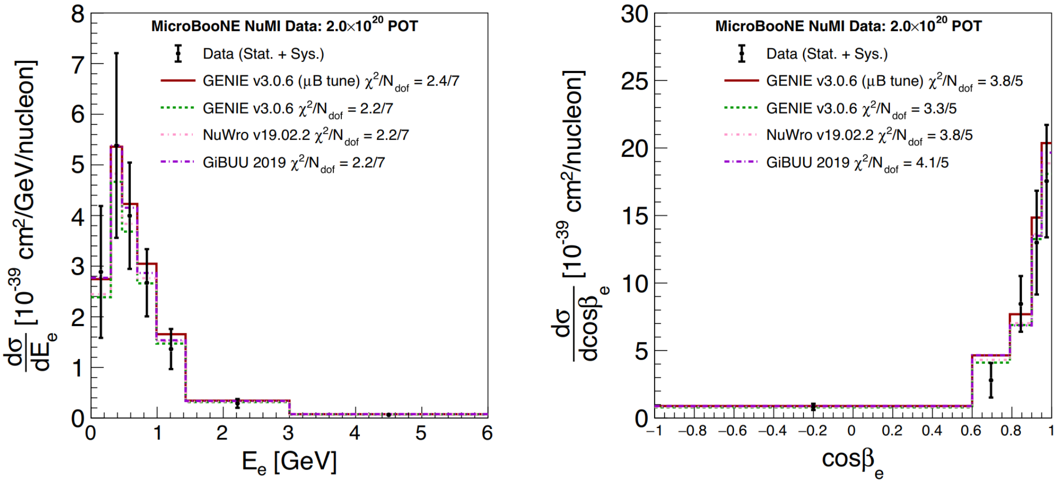


Figure 2.19: μBooNE measurement of the ν_μ CC 0π cross section on argon differential in muon momentum (left) and angle (right). Figure from [120].

final states[121, 122]. Measurements of pionless ν_μ CC interactions are shown in Figure 2.19. This shows significantly improved agreement with the more recent GENIE version 3 compared to version 2 in both muon momentum and angle. However, a suppression of forward going muon tracks is observed in the data and predicted by neither configuration.

μBooNE additionally made a measurement of the ν_e cross section on argon, with the highest statistics to date of 243 selected events, with an estimated purity of 72%[123]. Unlike the ArgoNeuT measurement, this measurement utilised a fully automated reconstruction and demonstrated the ability to reject the cosmic ray background associated with surface detectors, demonstrating vital capabilities for future SBN measurements. Additionally, this was the first measurement of the ν_e cross section on argon differential in energy, as shown in Figure 2.20. The large systematic uncertainties, predominantly driven by flux, and statistical uncertainties make meaningful comparisons between generators difficult.



(a) (b)

Figure 2.20: μBooNE measurement of the ν_e CC cross section on argon differential in electron momentum (left) and angle (right). Figure from [123].

Overall, many developments to both the theoretical models and experimental measurements have been made in recent years. In particular, nuclear effects including removal energy, Fermi momentum and FSI are now significantly more well modelled and constrained. However, scaling these models to heavier nuclei such as argon remains a challenge. In particular, the scarcity and impurity of ν_e data fails to adequately constrain the models and underlying assumptions. Understanding these effects is vital for not only developing the understanding of nuclear physics but also reaching the physics potential of next generation oscillation measurements, in particular for SBN and DUNE.

2.5 Concluding Remarks

Since the original proposal of the neutrino, many experimental and theoretical advances have been made to better understand its properties and behaviour. In particular, the phenomenon of neutrino oscillation was the first observed beyond the standard model behaviour. Efforts to better understand and characterise the oscillation continue, both to measure the parameters that describe it and investigate anomalous results.

Underpinning these efforts is improved modelling of neutrino nucleus interactions. Despite numerous advances, both theoretically and experimentally, continued efforts are required to understand these interactions and unlock the potential of next generation neutrino experiments. This is especially the case for experiments that plan to make use of heavy nuclear targets, specifically DUNE using an Ar target. The work presented in this thesis assesses the capability of SBND to be able to measure ν_e cross sections, improving on the statistically limited measurements made to date.

Chapter 3

Liquid Argon Time Projection Chambers

Often described as modern bubble chambers, Liquid Argon Time Projection Chambers (LAr TPCs) are a rapidly developing detector technology within the field of neutrino physics. As presented in Chapter 2, the Short-Baseline Neutrino (SBN) program and Deep Underground Neutrino Experiment (DUNE) are current and future experiments utilising this technology. Understanding the detector effects is vital to reaching the physics goals of these experiments.

The history and operating principles of LAr TPCs are first reviewed. Next, an overview of how particles deposit energy in the detector is given. Then, the propagation of the electrons and photons created by these energy depositions to the detectors that read them out is described. Finally, the detection of these electrons and photons is then presented, describing some of the key readout technologies used.

3.1 Historical Overview and Introduction

First proposed in 1977, LAr TPCs were presented as a way to maintain the resolution of bubble chambers whilst scaling detectors to larger target masses [115]. A diagram depicting a typical LAr TPC can be seen in Figure 3.1. This shows charged particles, such as those produced in a neutrino interaction, ionising the argon as they traverse the detector. A high, negative, voltage is applied to a cathode plane creating a strong electric field under which the ionised electrons will be drifted to the anode readout plane, typically a series of wire planes. The signals induced on the wires by the drifting electrons yields a high granularity image of the interaction. Scintillation light produced as the particles traverse the argon can additionally be detected in order to gain precise timing information on the interaction.

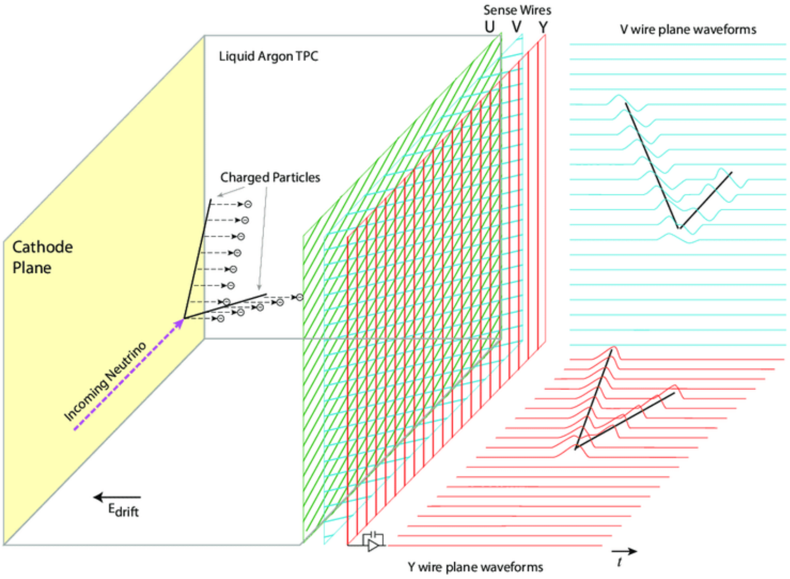


Figure 3.1: Diagram demonstrating the operational principles of a LAr TPC. The charged particles ionise the argon and produce ionisation electrons which are drifted to the wire readout under a constant electric field. Signals are then induced on the wire planes, each giving a unique perspective of the event based on the wire orientation. The first two wire planes have bipolar induction signals whilst the final collection plane has unipolar signals. Figure from [124]

An event display from a simulated ν_e Charged Current (CC) interaction can be seen in Figure 3.2, where the colour represents the amount of charge. The neutrino interaction vertex is clearly visible, accentuated by the charge deposited from low energy hadronic activity, from which an electron shower and proton track emerge. The energy deposited by the proton track increases along its length, showing clear evidence of the Bragg peak expected from a stopping particle. In contrast, the electron shower demonstrates the complicated substructure expected from the stochastic processes associated with showering particles. This event display highlights the capabilities of LAr TPCs to provide high granularity information, both calorimetrically and topologically, for neutrino interactions.

Argon has several qualities that make it an ideal choice for a neutrino detector. The high density of argon, 1.39 g/cm^3, increases the probability of a neutrino interacting in the detector[125]. The inert nature of argon yields high electron mobility and low reattachment, making drifting electrons over long distances possible with minimal attenuation given sufficient purity[126]. Similarly, the high scintillation light yield and transparency to the light produced allows for ample light to be

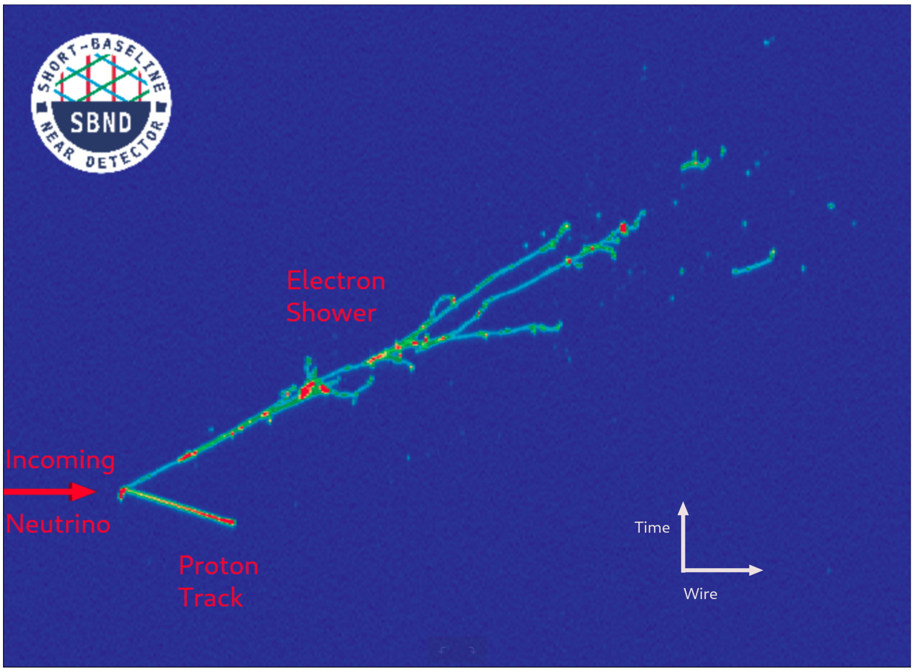


Figure 3.2: Event display of a simulated ν_e CC interaction in SBND, producing an electron shower and proton track.

collected[127]. The relative availability of argon coupled with the ability to be liquefied by liquid nitrogen are key economic considerations.

The first demonstration of a LAr TPC was performed in 1990 using a prototype for the Imaging Cosmic And Rare Underground Signals (ICARUS) detector[128]. Subsequent developments included scaling up to the full ICARUS T600 detector, made of of two cryostats each with with a active target mass of 235 tons[129]. Whilst running at the underground Gran Sasso laboratory, ICARUS developed demonstrated the both hardware and reconstruction capabilities of LAr TPCs[126, 130]. Using the CERN Neutrinos to Gran Sasso (CNGS) beam, ICARUS performed a search for Liquid Scintillator Neutrino Detector (LSND) anomaly, limiting the allowed phase space[131].

Meanwhile, the Argon Neutrino Teststand (ArgoNeuT) detector was built in the Neutrinos at the Main Injector (NuMI) beam at Fermilab which performed the first cross section measurements on argon[132]. Due to the small size of ArgoNeuT, at only 24 kg active mass, the detector was placed upstream of the Main Injector Neutrino Oscillation Search (MINOS) near detector which provided energy estimation for uncontained tracks. ArgoNeuT further demonstrated the potential for LAr TPCs to make precision measurements of neutrino interactions with low thresholds and excellent Particle IDentification (PID) capabilities, albeit with only semi-autonomous reconstruction[118]. Particularly relevant for this thesis, ArgoNeuT made the first measurement of the electron neutrino cross section on argon, although this was statistically limited with only 13 observed events and no energy measurement due to the small size of the detector[119].

The cryostat from ArgoNeuT was later repurposed for the Liquid Argon In A Testbeam (LArIAT) experiment[133]. Rather than measuring neutrino interaction cross sections, LArIAT measured the hadronic interaction cross sections from particles produced in a testbeam[134]. Such testbeam experiments are vital to ensuring particle interactions are correctly modelled in LAr TPCs.

The next LAr TPC to be constructed was the 85 ton μBooNE, placed in the Booster Neutrino Beam (BNB) upstream of MiniBooNE to investigate the observed low energy excess[124]. Continued developments were made in improving the understanding of the detector technology and reconstruction, in particular addressing the challenges associated with the high cosmic ray rate of a surface detector. Several cross section measurements have been made using μBooNE, leaveraging both the BNB and the off-axis NuMI beam, further demonstrating the capabilities of LAr TPCs[121, 122]. Notably, this includes the first differential measurement in energy of the electron neutrino cross section in argon using the NuMI beam, improving the observed number of events by an order of magnitude to 243[123]. Whilst searching for the low energy excess observed in MiniBooNE, μBooNE found no evidence of such an excess and thus further increased tensions within the field[67, 68].

In order to conclusively address the low energy excesses from LSND and MiniBooNE the SBN program at Fermilab was proposed[1]. This will make use of three detectors at various baselines along the BNB to search for sterile neutrino induced oscillations that could cause the excess. This program will consist of a newly built 112 ton near detector, the Short-Baseline Near Detector (SBND), μBooNE as the intermediate, and a refurbished ICARUS as the far detectors. The physics goals of this program, and the standalone goals for SBND, will be discussed further in Chapter 4.

The next-generation DUNE detector will perform a precision long baseline measurements of neutrino oscillations, ultimately aiming to measure δ_{cp}[51]. The far detector will consist of four 10 kt LAr TPCs, over an order of magnitude larger than previous detectors. The prototypes for these LAr TPCs, named protoDUNE, demonstrated the scalability of the detector technology at a mass of 530 tons[112]. New detector readout technologies that improve performance in high-multiplicity environments have additionally been demonstrated, driven partially by the requirements of the near detector complex[135, 136].

3.2 Ionisation and Scintillation

The general energy loss profile for particles is shown in Figure 3.3, specifically showing muons on copper but the mechanisms are applicable to argon[137]. This shows the stopping power, the energy loss per unit length divided by the density of the target material, against the momentum of a particle. For highly relativistic particles, such as > 100 MeV electrons and photons in argon, the energy is lost primarily through radiative effects. As such, these particles form electromagnetic showers within the detector. Heavier particles, such as muons, pions and protons, are described by the Bethe-Bloch formalism and produce tracks within the detector. Examples of these topologies can be seen in Figure 3.2.

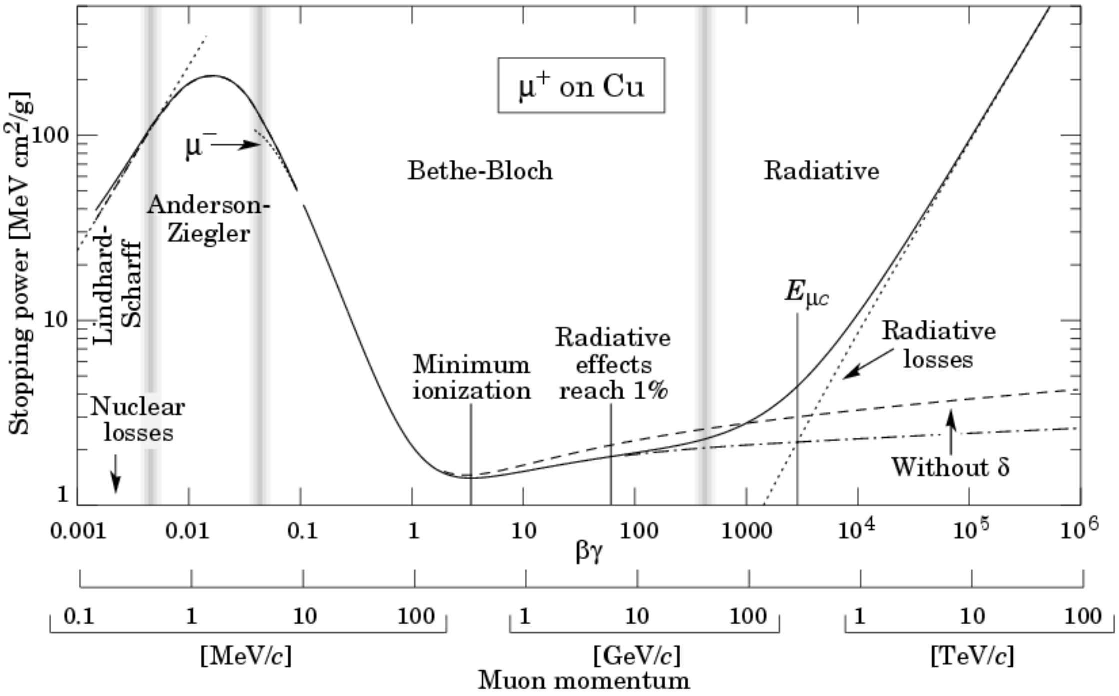


Figure 3.3: Particle energy loss in matter, exemplified by a muon in copper. The Bethe-Bloch and radiative regimes are shown which describe the energy loss for tracks and showers respectively. The rise in stopping power at lower energies represents the Bragg peaks observed as particles come to rest in a detector. Figure from [137]

Tracks typically deposit energy via ionisation in a straight line through the detector, although undergo Multiple Coulomb Scattering (MCS) dependent on the particle momentum. The energy deposited per unit length, dE/dx, is typically fairly constant in the Minimum Ionising Particle (MIP) region but rises as the particle comes to a stop, known as a Bragg peak. The profile of this energy loss is dependent on the particle mass and thus can be used to perform Particle IDentification (PID)[132]. The energy depositions in for MIP tracks is described by a Langau-Gaussian Convolution (LGC), with some large forays from the Most Probable Value (MPV).

Electrons above a critical energy, $E_{Crit} = 39$ MeV in argon, typically form showers which lose the majority of their energy via radiative effects and thus create a cone of electromagnetic activity. Photons travel some distance before interacting with the argon, as they are neutral no ionisation occurs prior to this interaction, and thus a gap is observed, referred to as a conversion gap. The distance this photon travels before interacting is shown in Figure 3.4[138]. At low energies the photons typically Compton scatter but at higher energies they generally pair produce into an e^+e^- pair. Both the conversion gaps and energy loss profiles of these showers can be used to differentiate between electrons and photons, discussed further in Chapter 6.

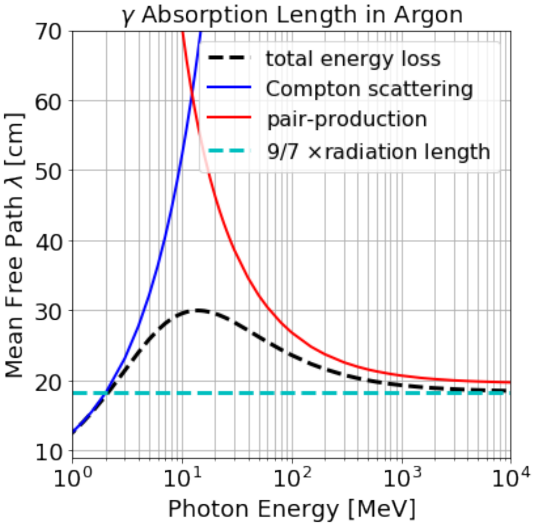


Figure 3.4: Plot showing mean free path of photons as a function of their energy. The contributions from different interaction modes are shown alongside the total and asymptotic limits. Figure from [138].

When a charged particle passes through argon it loses energy by both ionising and exciting the argon, as shown in Figure 3.5[139]. The liberated electrons can then recombine with the argon ions, a process known as recombination, or escape to be drifted towards the anode. Excited argon atoms (Ar^*), produced either via collisions or recombination, will typically form excimers (Ar_2^*) that subsequently decay via the emission of a 128 nm Vacuum Ultra-Violet (VUV) photon[139].

The probability of recombination occurring depends on both the electric field and local charge density, and is defined as:

$$R = \frac{W_{ion}\, dE/dx}{dQ/dx} \tag{3.1}$$

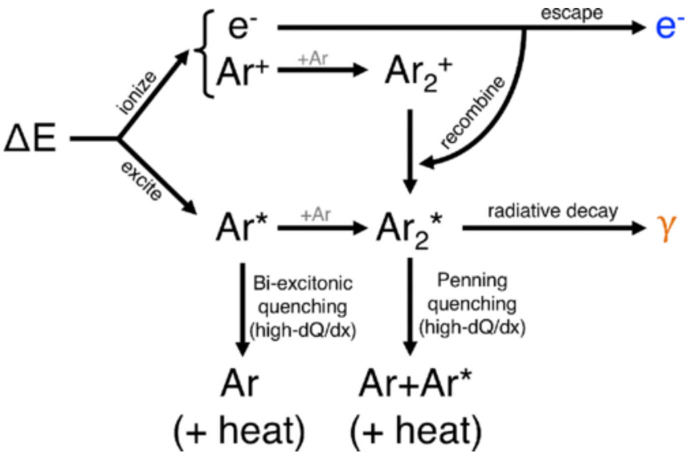


Figure 3.5: Diagram depicting the ionisation, recombination and scintillation in liquid argon. Figure from [139].

Where dQ/dx and dE/dx represent the energy and charge loss per unit length respectively and $W_{ion} = 23.6$ eV is the energy required to ionise an argon atom[141]. The recombination factor is typically modelled using either the modified box or Birks model, although the latter is disfavoured due to spurious values at high charge densities[140]. The original box model is based on columnar theory around the charge deposition which is then modified by the addition of, experimentally derived, parameters which increase agreement with data at low charge densities[140]. As such, the recombination factor depends on both the density of deposited charge in the track and the electric field present. A comparison of these models can be seen in Figure 3.6, demonstrating the non-linear dependence on dE/dx. Understanding

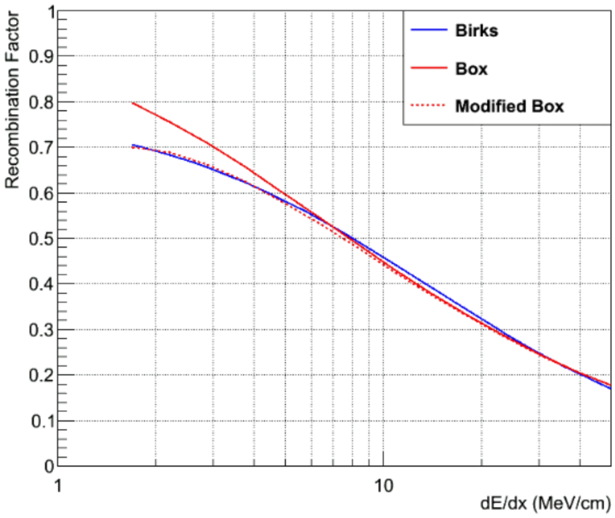


Figure 3.6: Recombination, the fraction of electrons not recaptured by the argon ion, as a function of the deposited energy density at 500 V/cm. Figure from [140].

recombination is vital to being able to reconstruct the deposited energy from the observed charge in the detector, discussed further in Chapter 6.

The argon exitons, and excimers, that are produced are subject to quenching where heat is produced rather than light. The rate of this quenching can be effected by the presence of impurities such as oxygen or nitrogen[142]. Excimers can exist in either a singlet or triplet state, which have fast (6 ns) and slow (1300 ns) decay times respectively. Cherenkov light is also produced for particles travelling at a sufficient velocity, although the light yield is significantly below that of the scintillation light[143].

3.3 Electron Drift and Photon Propagation

Electrons that do not recombine with the argon ions are then drifted under the presence of an electric field, created by applying a high negative voltage to a cathode as shown in Figure 3.1. At typical LAr TPC conditions, drift field of 500 V/cm and temperature of 87 K, the electrons drift away from the cathode at a velocity of 0.16 cm/μs[144].

Electrons produced in a point like energy deposition will diffuse both longitudinally and transversely, defined with respect to the drift direction. The width of a pulse is modelled as:

$$\sigma^2(t) = \sigma_0^2 + \left(\frac{2D}{v_d^2}\right) t \tag{3.2}$$

The observed width σ depends on the initial width σ_0, the drift velocity v_d, the drift time t, and the diffusion coefficient D[145]. The diffusion coefficients, under the same typical conditions, have been measured to be $D_L = 7.2$ cm^2/s and $D_T = 12.0$ cm^2/s[144].

The argon ions produced drift towards the cathode at a velocity of 8×10^{-7} cm/μs, several orders of magnitude slower than that of the electrons[146]. With a sufficiently high rate of ionisation, for example from cosmic rays in a surface detector, these argon ions can build up causing a distortion to the electric field, referred to as the Space Charge Effect (SCE)[120]. These electric field distortions can effect the energy depositions both calorimetrically and spatially. The calorimetric effects come via the dependence of recombination on the local electric field. Spatial offsets appear due to electrons drifting along a distorted field, leading to bowed tracks and end points being pulled away from the detector edges. These effects, if uncorrected, can lead to calorimetric distortions, effecting PID and energy resolution, and incorrect track lengths and positions. This effect can be measured by using a combination of cosmic ray muon tracks and dedicated laser calibration systems[120]. Once this effect has been measured, corrections can be applied both spatially and calorimetrically as will be discussed in Chapter 6.

Drifting electrons can additionally be captured by electronegative impurities in the argon, most commonly oxygen and water[112]. This attenuates the charge that arrives at the wires, subsequently reducing signal amplitude, as:

$$Q_{Wire} = Q_{Dep} \cdot e^{-t/\tau} \tag{3.3}$$

Here, Q_{Wire} is the charge collected on the wire, Q_{Dep} is the deposited charge, t is the drift time, and τ is the electron lifetime which characterises the level of attenuation. ProtoDUNE, which utilises the membrane cryostat technology as SBND as will be discussed in Chapter 4, has recently reported lifetimes in excess of 100 ms, corresponding to an oxygen-equivalent purity of 3.4 ppt[112]. This is several orders of magnitude larger than the drift time of SBND (1.25 ms) making this effect almost negligible. Nevertheless, corrections are applied to account for this attenuation in during the calorimetric reconstruction, as discussed in Chapter 6.

Scintillation light produced in the argon subsequently propagates through the detector, with some chance to be detected based on the solid angle of the detectors. This light may also undergo Rayleigh scattering, whereby the photons elastically scatter off the nuclei, which must be taken into account as it can reduce the light yield over long distances. The distance a photon can propagate before scattering is dependent on the wavelength of the light but measurements range from 66-90 cm for 128 nm light[147, 148]. TetraPhenyl Butadiene (TPB) coated wavelength-shifting foils can also be placed on the cathode which reflect the indecent light back towards the light detectors placed behind the anode, as is planned for SBND[149]. By shifting the light to a longer wavelength the Rayleigh scattering is reduced, increasing the number of photons detected[1].

3.4 Detection of Charge and Light

Once the electrons have propagated through the argon to the anode they then induce a signal on a readout. Traditionally, this is performed using a series of wire planes, each consisting of a set of parallel wires typically separated by a few mm. A bias voltage is applied to each of the planes to ensure transparency for the drifting electrons in the initial planes and collection on the final plane. These are typically referred to as induction and collection planes respectively due to shapes of the signals induced. This can be seen in Figure 3.1, where the induction (labelled V) plane shows bipolar signals induced by the passage of electrons and the collection (labelled Y) plane shows the unipolar pulses from the collection. There are most commonly three planes orientated 60° apart, minimising the impact of a track travelling parallel to a wire plane. Combining the signals between these planes allows for reconstructing the position of an energy deposition in the plane perpendicular to the drift direction.

The time a signal is induced on the wire plane, T_{Wire} is described by:

$$T_{Wire} = T_0 + T_{Drift} = T_0 + \frac{X}{v_{drift}} \tag{3.4}$$

Here, T_0 is the time the ionisation occurred and T_{Drift} is the time taken for the charge to drift to the wires which can also be expressed as the product of the position in the drift direction, X, and the drift velocity v_{drift}. As such, there is a degeneracy between the time and drift co-ordinate of ionisation, although this can be disambiguated by detector subsystems, e.g. using the light signal. If the T_0 is not known, the X position is typically reconstructed by assuming that the ionisation occurred at the trigger time, defined as $T_0 = 0$.

The signals induced on these wires are subsequently shaped, amplified and digitised by cold electronics before being recorded by the Data AcQuisition (DAQ). The data is typically recorded for a drift window, the time required for charge to drift the full length from the cathode to the anode, plus some padding for recording cosmic rays that may enter the detector outside the trigger time. Placing the electronics within the Liquid Argon (LAr), referred to as cold electronics, greatly reduce the noise compared to warm electrons, placed outside the LAr, due to both the lower thermal noise and shorter cable lengths, resulting in lower capacitance[150]. As such, a large portion of the noise present in modern LAr TPCs is due to the inherent capacitance of the Time Projection Chamber (TPC) wires, as opposed to external wiring or electronics. Filtering, via both hardware and software, has additionally been shown to further reduce this noise and raise the Signal-to-Noise Ratio (SNR)[150].

Several additional considerations must be made to take into account including the effect of the drifting electrons on distant wires, the electronics response during shaping and amplification, and the spatial distribution of charge, both due to diffusion and the topology of the particles. The processes of both simulating and correcting for these effects will be discussed in Chapter 4.

Recent efforts have begun to explore alternative readout technologies. Pixelated readouts have been demonstrated the ability to provide full 3D reconstruction whilst avoiding ambiguities present with wires[135, 136]. This is particularly important for high occupancy detectors like the proposed DUNE near detector. Dual-phase readouts that extract electrons to a gas phase and amplify the signal before being read out using a Printed Circuit Board (PCB) has also recently been demonstrated, potentially allowing for increased Signal-to-Noise Ratio (SNR)[151].

In order to detect the Vacuum Ultra-Violet (VUV) light produced in a LAr TPC a wavelength shifter, most commonly TPB, is used to shift the light to a longer wavelength that can then be detected, typically by a Photo-Multiplier Tube (PMT). This is typically performed by coating the PMTs with TPB, although TPB coated reflected foils can also be utilised. TPB has been shown to have an efficiency close to 100% but can introduce a time delay due to the de-excitation[152].

Alternatively, the Argon R&D Advanced Program at UniCAmp (ARAPUCA) is designed to trap light by utilising a combination of wavelength shifters and dichroic filters, the latter are transparent to a narrow range of wavelengths and otherwise reflective[153]. The dichroic filter is coated in wavelength shifter such that incident VUV light passes through the filter and is subsequently shifted to a longer wavelength and thus cannot exit through the filter, as shown in Figure 3.7. The light is then internally reflected in the trap until it is detected by a series of Silicon Photo-Multipliers (SiPMs), guided by a wavelength shifting light guide in the refined X-ARAPUCA design[154]. These detectors have been developed for the DUNE detector where there is insufficient room for traditional PMTs between the anode planes. The ability to function in the presence of an external electric field is an additional advantage over PMTs, extending the potential placement options.

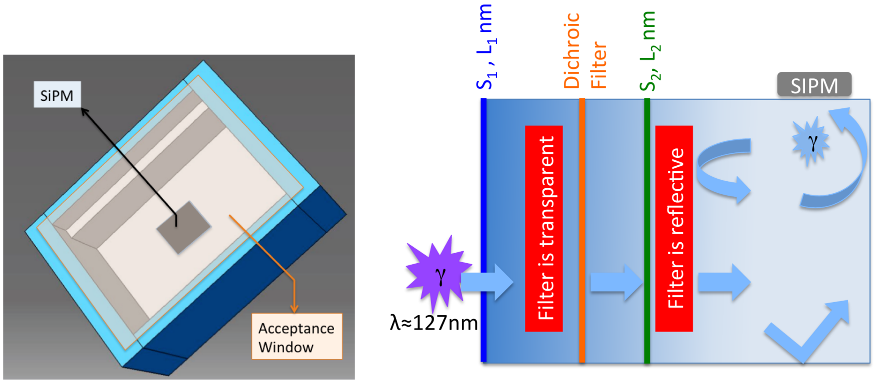


Figure 3.7: Diagram demonstrating the operational principles of a ARAPUCA. Incident photons are wavelength shifted such that they can pass through a dichroic filter. They are subsequently shifted a second time so that they are unable to pass though the filter and are trapped. Figure from [153].

3.5 Concluding Remarks

Since their proposal in 1977, LAr TPCs have come to fruition and demonstrated their potential as a high acceptance, low threshold particle detector. The behaviour of particles interacting within the detector and of the detector response to these interactions are increasingly understood. Technical advancements to these detectors continue to broaden their application, from scaling to ever larger detectors to readout technologies better suited to high multiplicity environments. Overall, the next generation of LAr TPCs aims to deepen the collective understanding of neutrino interactions and oscillations alike, including the SBN program discussed in the following chapter.

Chapter 4

The Short-Baseline Near Detector

In order to investigate the Liquid Scintillator Neutrino Detector (LSND) and MiniBooNE anomalies, as outlined in Chapter 2, the Short-Baseline Neutrino (SBN) program was proposed to search for evidence of sterile neutrino induced oscillations. This programme will utilise three Liquid Argon Time Projection Chambers (LAr TPCs), the Short-Baseline Near Detector (SBND) at 110 m, μBooNE at 470 m and ICARUS at 600 m, in the Booster Neutrino Beam (BNB) that MiniBooNE observed its excesses in. Of particular relevance to this thesis, the Short-Baseline Near Detector (SBND) not only plays a key role in constraining systematic uncertainties for the oscillation measurements, but has a rich physics program of its own.

This chapter begins by discussing the SBN program in section 4.1, including a discussion of the detectors that comprise this program and the physics goals. Next, an overview of the BNB is given in section 4.2 that describes the both the design and flux prediction at SBND. Details of the SBND detector subsystems are then presented, including the unique hardware features of the detector, in section 4.3. The simulation of SBND is next discussed in section 4.4, covering event generation, particle propagation, and detector simulation. Finally in section 4.5, the reconstruction of the low level Time Projection Chamber (TPC) and the detector subsystems is presented, covering both the Cosmic Ray Tagger (CRT) and Photon Detection System (PDS).

4.1 The Short-Baseline Neutrino Program

In order to investigate the LSND and MiniBooNE anomalies discussed in Chapter 2, a new physics program was proposed: the SBN program. The program consists of three LAr TPCs as depicted in Figure 4.1. This shows the placement of the three experiments and their along the BNB (the same beam as MiniBooNE): the SBND at 110 m, μBooNE at 470 m and the Imaging Cosmic And Rare Underground Signals (ICARUS) detector at 600 m.

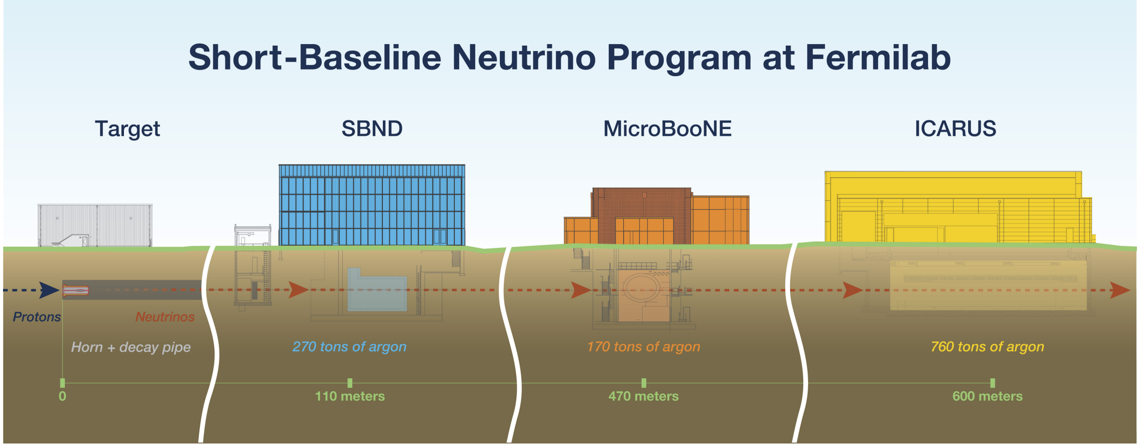


Figure 4.1: Overview of the SBN program showing the BNB target and the three detectors: SBND, μBooNE, Imaging Cosmic And Rare Underground Signals (ICARUS). The respective masses of the experiments are also shown alongside the distance from the BNB target. Figure from [1].

By utilising the same detector technology across the three experiments, the program aims to achieve significant cancellation of systematic uncertainties in the search for sterile neutrino induced oscillations. In addition to this search for sterile neutrinos, the individual experiments also have rich cross section and Beyond Standard Model (BSM) capabilities, with both μBooNE and ICARUS also receiving neutrinos from the off-axis Neutrinos at the Main Injector (NuMI) beam[155].

4.1.1 SBN Physics Goals

The excesses that have been observed by LSND and MiniBooNE now stand at 3.8σ and 4.8σ respectively[56, 58]. The SBN program was designed to definitively search for the excess and, if found, investigate its nature. The ability for LAr TPCs to effectively differentiate between electron and photon showers, something that MiniBooNE was not able to do, will be crucial for both rejecting backgrounds and revealing the nature of the excess[119]. Early searches from μBooNE have found no evidence of any excess, in either the electron or photon channels, but the full SBN program is required to address these tensions definitively[67, 68].

The most common explanation for the excesses observed by LSND and MiniBooNE are the existence of light sterile neutrinos that enhance neutrino oscillations. The excess of observed ν_e interactions however exists in strong tension with the lack of associated ν_e and ν_μ disappearance. The SBN program will be able to search for sterile neutrinos in all of these channels and aims to definitively address these anomalies and tensions. Having a near detector, SBND, the SBN program is designed to constrain the systematic uncertainties associated with the neutrino flux and interaction model.

This enables the program to differentiate between an excess caused by oscillations, appearing differently in each detector, or some other physical process, manifesting in all detectors identically.

The exclusion sensitivity of the SBN program to both the ν_μ disappearance and ν_e appearance are shown on the left and right of Figure 4.2 respectively. For ν_e appearance, the shaded LSND allowed region is excluded at almost 5σ by the SBN program, demonstrating the improved sensitivity over existing constraints showed in the dashed lines. Similarly for ν_μ disappearance, the program extends the excluded region beyond the existing measurements. Utilising this ability to probe both the appearance and disappearance channels simultaneously the SBN program aims to conclusively address the low energy excesses presented by LSND and MiniBooNE.

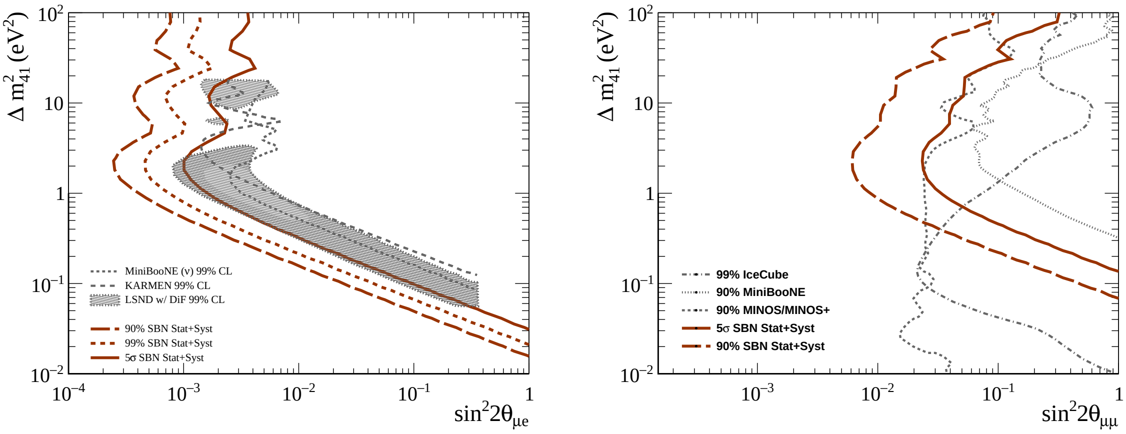


Figure 4.2: Sensitivity of the SBN program to ν_μ disappearance (left) and ν_e appearance (right). Figure modified from [156] with external data from [61–63, 76].

4.1.2 SBND Physics Goals

In addition to investigating the low energy excess, the individual experiments that comprise the SBN program each have rich physics programs of their own. As the near detector, SBND is exposed to the highest neutrino flux of the three detectors. As such, over 5.5 million neutrino interactions are expected over three years of running, broken down by hadronic final state in Table 4.1. These numbers represent a multiple order of magnitude increase over the current argon cross section measurements made by μBooNE, which contained only 11,528 events, even after selection inefficiencies are included[157].

The large number of interactions yields low statistical uncertainties, allowing for exclusive final states and multi-dimensional differential cross section measurements to be performed. The low thresholds and excellent Particle IDentification (PID) capabilities of LAr TPCs allow for this selection of exclusive final states and precise measurements of particle kinematics. Of particular interest to this thesis is the expected rate of 29,841 ν_e Charged Current (CC) interactions, assuming a nominal Protons On Target (POT) of 6.6×10^{20} corresponding to three years of data collection. This represents another multiple order of magnitude increase over existing measurement, also made by μBooNE with 243 selected ν_e CC interactions[123]. However, the large backgrounds represent a challenge due to the intrinsically low fraction of ν_e in the beam. Specifically, the 640,000 ν_μ CC and ν Neutral Current (NC) interactions that produce a π^0 represent a particularly difficult background to reduce which dramatically outnumbers the signal.

Additionally, SBND can search for rare phenomena including neutrino-electron elastic scatters, where the neutrino interacts with an electron rather than a nuclei. As a purely electroweak process with a well understood cross section, these events can be used to constrain the flux prediction for the first time using a LAr TPC[158]. These factors, combined with novel detector features that will be discussed later in this chapter, grant SBND the ability to make world leading cross section measurements and BSM searches alike[155].

Hadronic Final State	Expected Rate
ν_μ CC Inclusive	3,998,310
→ 0p	175,219
→ 1p	2,107,538
→ 2p	746,517
→ ≥3p	969,051
→ $0\pi^\pm$	3,224,194
→ $1\pi^\pm$	665,977
→ $2\pi^\pm$	90,737
→ $1\pi^0$	405,174
ν_e CC Inclusive	29,841
→ 0p	2,377
→ 1p	14,524
→ ≥2p	12,939
→ $0\pi^\pm$	21,630
→ $1\pi^\pm$	6,401
→ $1\pi^0$	4,316
ν NC Inclusive	1,570,640
→ $0\pi^\pm$	1,304,762
→ $1\pi^\pm$	221,173
→ $1\pi^0$	250,805
→ e^-	393
Total	5,598,791

Table 4.1: Expected neutrino interaction rates in SBND for 6.6×10^{20} POT, corresponding to 3 years of running, predicted using GENIE tune G18_10a_02_11a.

4.2 The Booster Neutrino Beam

Neutrinos are provided to the SBN program by the Booster Neutrino Beam (BNB), a tertiary neutrino beam with an average neutrino energy of 800 MeV discussed extensively in Reference [2]. The process begins with extracting 8 GeV protons from the booster synchrotron in a “spill” at a rate of up to 5 Hz. Each spill consists of 5×10^{12} protons delivered in 81 individual “buckets”, each with characteristic width of 2 ns and separated by 19 ns, over a total spill time of 1.6 µs. Experiments with sufficient, ns level, timing resolution can utilise the bunch structure to reject cosmic activity in outside the buckets[1].

The protons removed from the booster are measured by two toroids, measuring the intensity, and beam position monitors and a resistive wall monitor, measuring the timing and position of the beam, which are used to calculate the systematic uncertainties. These protons are then incident on an air cooled beryllium cylinder 71.1 cm long and 0.51 cm wide shown in black on Figure 4.3. Secondary hadrons produced when the protons interact with the target, such as those shown in green, are then focused by a 170 kA electromagnetic horn, indicated by the blue region. The polarity of the horn can be chosen to focus either positive or negative hadrons for running in neutrino or antineutrino enhanced modes respectively. Focused particles then propagate to a 50 m long decay region which is terminated by a steel and concrete absorber, shown as the grey area. This decay region allows the hadrons to decay and produce neutrinos, shown in red, whilst the absorber reduces the decay in flight of long lived muons, reducing the production of ν_e. The neutrinos then propagate through a dirt region before reaching the SBND detector.

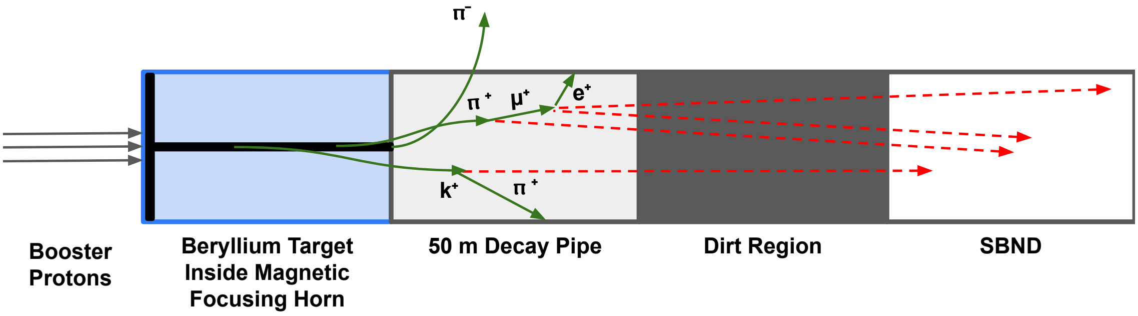


Figure 4.3: Illustrative diagram of neutrino production in the BNB. Incoming protons interact on the beryllium target producing secondary hadrons. These hadrons are focused or defocused by a magnetic horn depending on the chosen polarity. Focused hadrons decay within the decay pipe to produce neutrinos that can propagate through the dirt to SBND.

The beam is simulated using GEANT4 with the hadronic interaction cross sections tuned to HARP data[2, 159]. Systematic uncertainties are calculated by a reweighting process, which will be presented in Chapter 8. The simulated flux at the front face of SBND is shown in Figure 4.4, broken down by the neutrino flavour and as a function of the neutrino energy. The beam is dominated by ν_μ (90.2%) with the second largest contribution from $\overline{\nu}_\mu$ (9.1%) and small contributions from ν_e (0.6%) and $\overline{\nu}_e$ (0.1%).

The flux for each flavour is further broken down by production mechanism in Figure 4.5. These show that the pion decay is the dominant production mechanism for both ν_μ and $\overline{\nu}_\mu$. The decay of the muons produced by these pion decays is the primary source of ν_e, particularly at low energy. Kaon production is significantly lower than that of pions, so the kaon decays contribute less to the overall neutrino flux. Nevertheless, the kaon decays contribute significantly at higher energies, especially for ν_e, due to semi-leptonic decays to leptons and pions. A peak can be seen at around 235 MeV in the ν_μ flux corresponding to kaons that decay at rest[160].

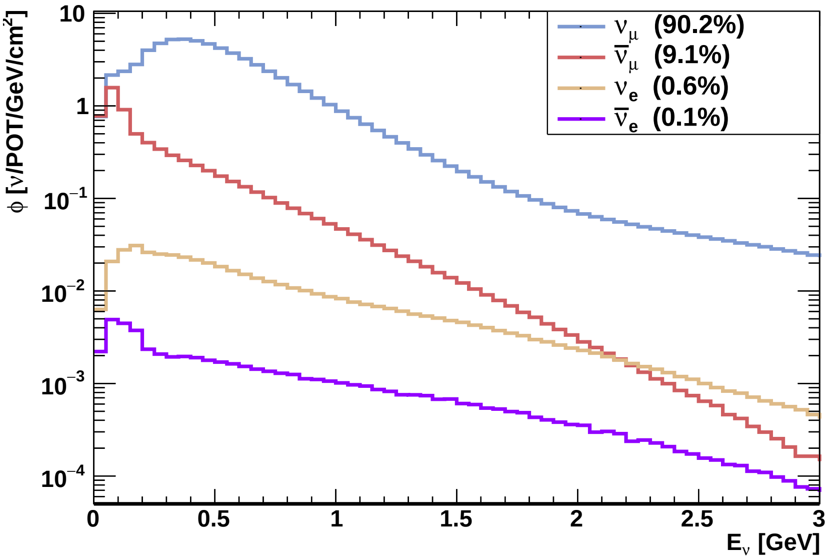
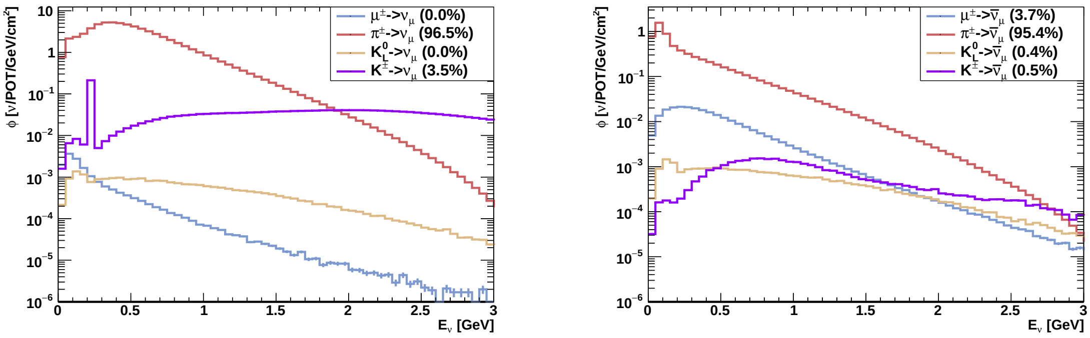
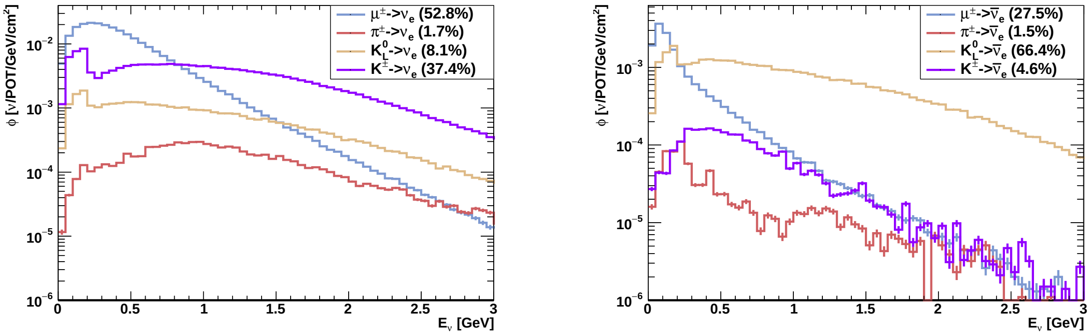


Figure 4.4: Simulated flux of different neutrino flavours at SBND from the BNB as a function of the true neutrino energy.



(a) ν_μ (b) $\overline{\nu}_\mu$



(c) ν_e (d) $\overline{\nu}_e$

Figure 4.5: Simulated flux of different neutrino flavours at SBND and the production mechanism in the BNB.

4.3 The Short-Baseline Near Detector

SBND is a 112 tonne active volume LAr TPC located 110 m from the BNB target, shown in Figure 4.6. The detector is 5 m long and 4 m high and consists of two separate TPCs, sharing a common Cathode Plane Assembly (CPA), each with a 2 m drift length. A comprehensive Photon Detection System (PDS) is located behind each of the Anode Plane Assemblies (APAs) with TetraPhenyl Butadiene (TPB) coated reflective foils placed within the cathode. The membrane cryostat that houses the TPC is surrounded by seven planes of Cosmic Ray Tagger (CRT) to give almost full coverage.

4.3.1 Time Projection Chamber

The SBND APA contains three sets of wire planes: two induction planes (referred to as U and V) at ±60° from the vertical collection plane (referred to as Y), as shown in green, blue and red respectively in Figure 4.7. Each of these wire planes consists of 150 µm diameter copper-beryllium wires spaced at 3 mm intervals and tensioned

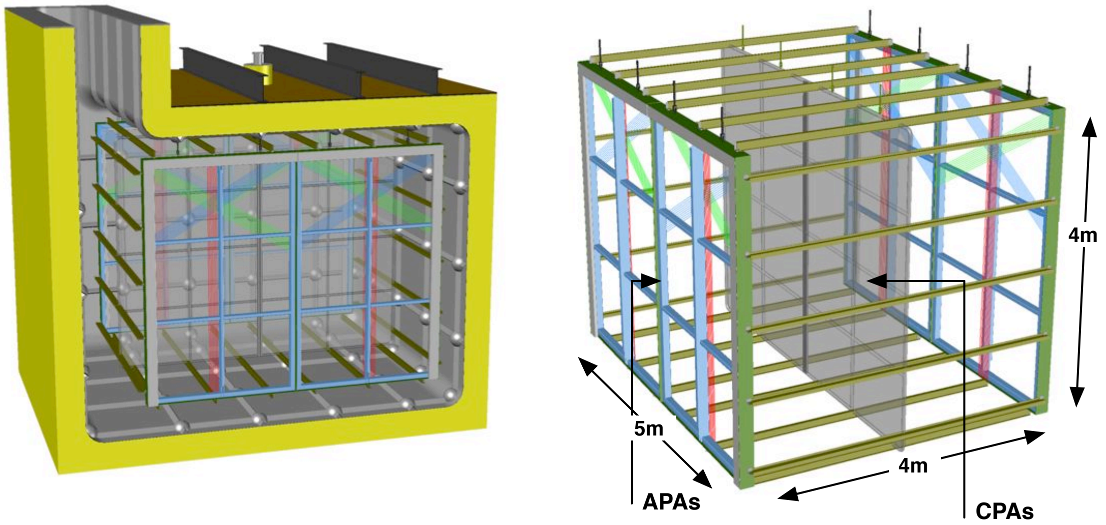


Figure 4.6: Schematic showing the SBND detector and cryostat (left) and the dimensions, APA and CPA (right). The detector is formed of two TPCs seperated by the central cathode. Figure from [161]

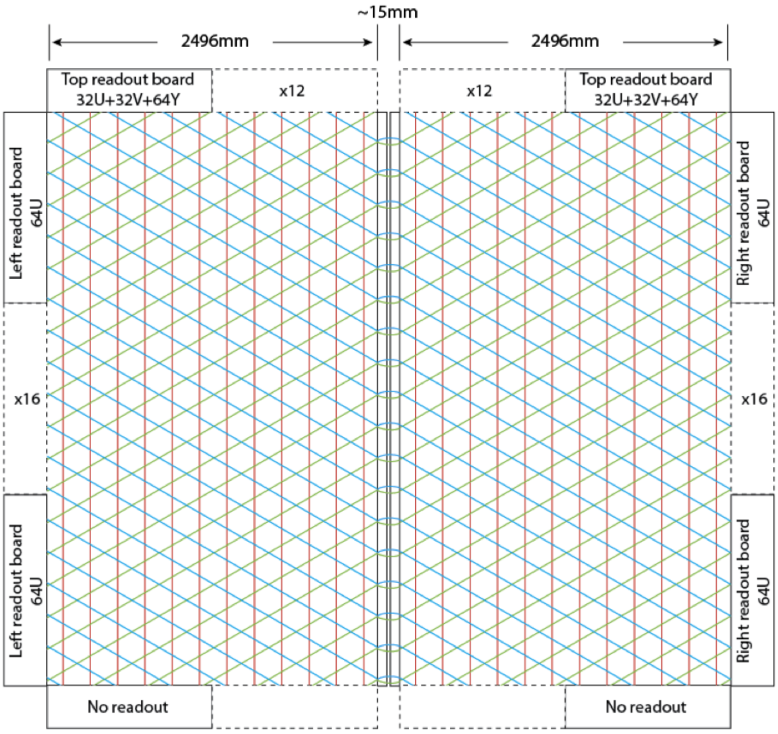


Figure 4.7: Diagram depicting a pair of coupled APA which together form a APA plane. The U, V, and Y wires are shown in blue, green and red respectively alongside the electronic readout boards. The APAs are coupled in the centre of the image with the induction (U and V) channels bridged across the gap. Figure from [1]

to 7 N in order to prevent sagging when cooled by the liquid argon[161]. The wire planes are separated by 3 mm and biased to voltages of -200 V, 0 V, and 500 V for the U, V, and Y respectively to ensure electrical transparency for the induction planes and collection efficiency for the collection plane. Each TPC consists of two connected APAs with jumper cables used to connect the induction planes across the 15 mm gap to form a single electronic channel. This can be seen in Figure 4.7 which depicts an pair of coupled APAs; the U, V, and Y planes are shown in blue, green and red respectively. In total there are 5,632 wires in each TPC: 1664 collection and 1984 in each induction plane.

The wires are connected to the cold electronics, shown in Figure 4.8, at the top and sides of each APA, these electronics are located in the Liquid Argon (LAr) to reduce both thermal noise and cable distances. The first stage of the readout electronics is a 16-channel front end Application-Specific Integrated Circuit (ASIC) that both shapes and amplifies the signal at a rate of 2 MHz. This signal is then passed to an Analogue Digital Converter (ADC) ASIC and subsequently a mezzanine

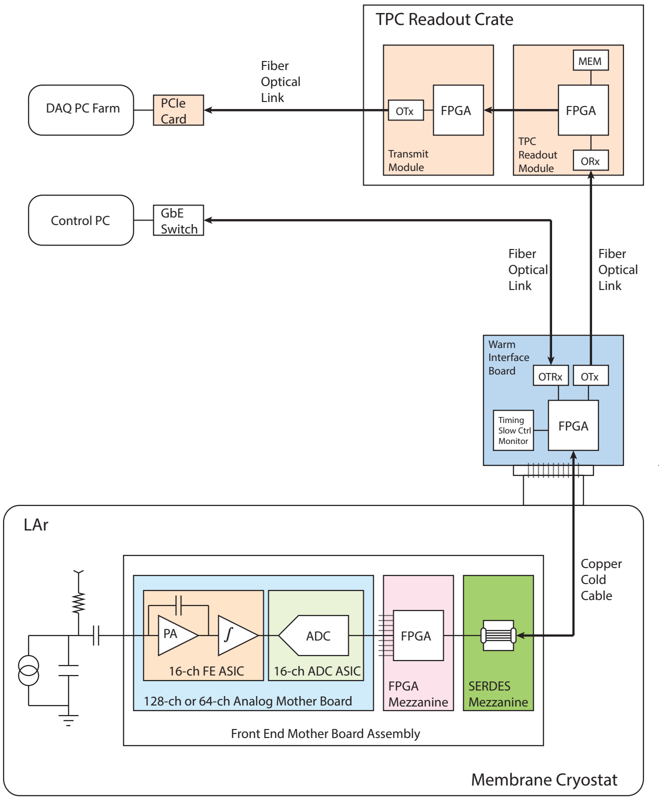


Figure 4.8: Schematic of the SBND readout electronics chain. The ASICs and FPGAs within the cryostat form the cold electronics which shape, digitise and transmit the signal to the WIBs outside the cryostat. Figure from [1].

Field Programmable Gate Arrays (FPGA) to multiplex the data from 128 channels, 8 ASICs. The data is subsequently sent out of the cryostat to a Warm Interface Board (WIB) via a SERializer/DESerializer (SERDES) mezzanine. Once outside the cryostat, data is sent from the WIB to the Data AcQuisition (DAQ) which will be discussed later in this section.

The cathode is made up of two CPAs, each consisting of a central frame and 8 sub-frames, as can be seen in Figure 4.9. The sub-frames consist of a TPB coated reflective foil covered by a wire mesh, covered by black protective shields in Figure 4.9. The cathode will be charged to a voltage of -100 kV by a high voltage feedthrough from outside the cryostat.

The field cage ensures a uniform 500 V/cm electric field over the drift volume, depicted in Figure 4.10 with the bottom section visible in Figure 4.9. This field cage consists of a series of electrodes orientated perpendicularly to the drift direction to step down the voltage in 3 kV increments.

At the time of writing, the TPC is undergoing assembly where the components are mounted to the Assembly Transport Fixture (ATF). This allows for the TPC and cryostats to be constructed in parallel, as opposed to other experiments where the TPC was assembled within the cryostat. Once fully mounted to the ATF, the TPC will then be connected to the cryostat lid and subsequently lowered into the cryostat. SBND is currently scheduled to begin taking data in 2023.

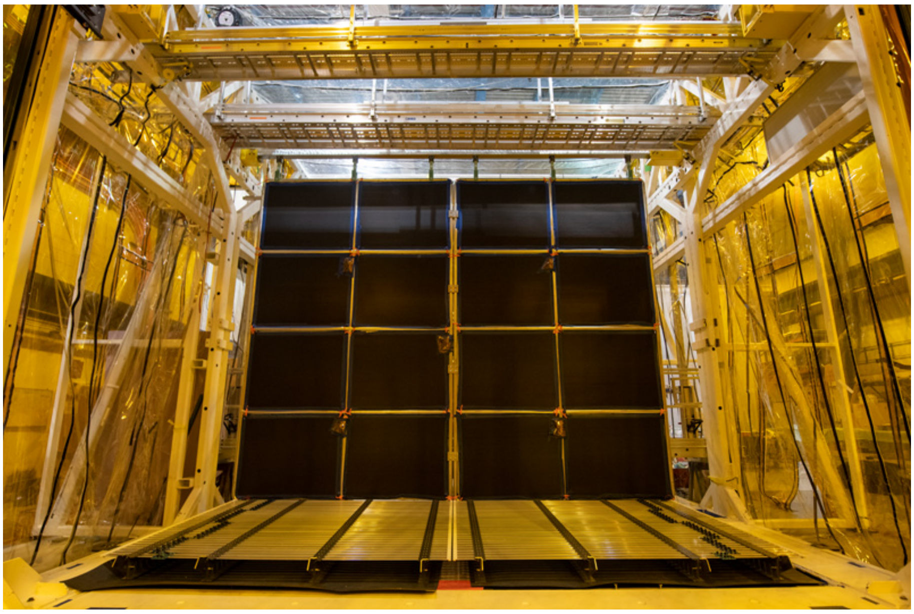


Figure 4.9: Picture of the SBND detector under construction at FermiLab. The central CPA is shown with the black covers protecting the reflective foils. The bottom section of the field cage can also be seen as the series of metal bars at the bottom of the image.

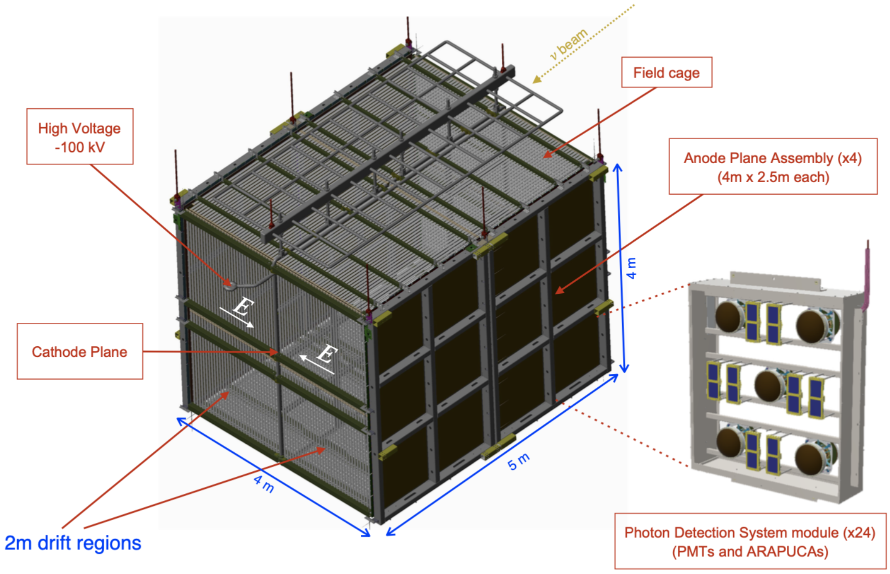


Figure 4.10: Schematic of the SBND detector showing the high voltage feedthrough and field cage used to create the electric field required to drift the electrons. The location of the 12 PDS boxes behind each APA plane can also be seen, with each consisting of 5 Photo-Multiplier Tubes (PMTs) and 8 Argon R&D Advanced Program at UniCAmps (ARAPUCAs). Figure from [1].

4.3.2 Photon Detection System

Detecting scintillation light is vital to achieve the timing resolution required to enable identify interactions in coincidence with the beam spill and thus reject cosmogenic backgrounds. SBND has a comprehensive light detection system to provide R&D for future experiments, in addition to facilitating the physics goals. This system is comprised of TPB coated reflective foils mounted to the cathode to both reflect the Vacuum Ultra-Violet (VUV) light towards the Photon Detection System (PDS) mounted behind the anode and shift it to visible (VIS) wavelengths.

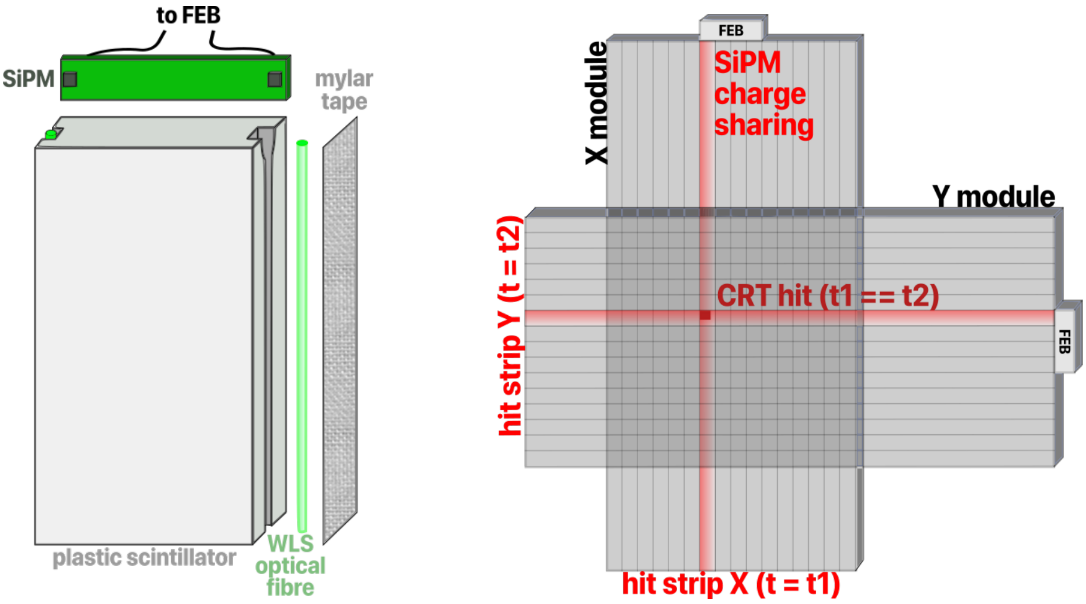
The PDS in SBND is made up of a series of modular "PDS boxes", shown on the right of Figure 4.10. Each box contains five PMTs, four TPB coated and a central uncoated, and 8 ARAPUCAs, arranged in four pairs of VUV and VIS detectors. The TPB coated PMT is capable of detecting the VUV light directly whilst the uncoated PMTs detect the wavelength shifted light reflected from the cathode foils. Similarly, ARAPUCAs with different wavelength shifters are paired to be sensitive to the direct (VUV) and reflected (VIS) light respectively. Detection of both direct and reflected light can be used to not only increase the light yield uniformity across the drift distance but additionally locate the position of an interaction in the drift direction, which will be presented in Chapter 7. The combination of traditional PMTs and novel ARAPUCAs allows for a direct comparison of the performance between the systems, providing vital research and development data for the Deep Underground Neutrino Experiment (DUNE).

The PMTs are read out at 500 MHz allowing for a 2 ns level timing resolution, in comparison to 80 MHz and 12.5 ns for the ARAPUCAs. The ns-level timing is a requirement for the ability to match interactions to individual buckets within the BNB spill, increasing cosmic rejection by an order of magnitude relative to matching to the entire spill.

4.3.3 Cosmic Ray Taggers

In order to reject the cosmic ray background associated with a surface detector, SBND is almost completely surrounded by a Cosmic Ray Tagger (CRT) system. Tracks from the TPC can be matched to hits in the CRT to provide ns-level timing resolution, allowing rejection of tracks that occur outside of the beam spill window.

This system is comprised of 12 cm wide, 1 cm thick scintillator strips connected to a pair of SiPMs via wavelength shifting optical fibres, shown in Figure 4.11a[162]. Each strip is read out if both SiPMs pass above threshold in coincidence, achieving ns-level time resolution. The relative amount of charge detected by each SiPM

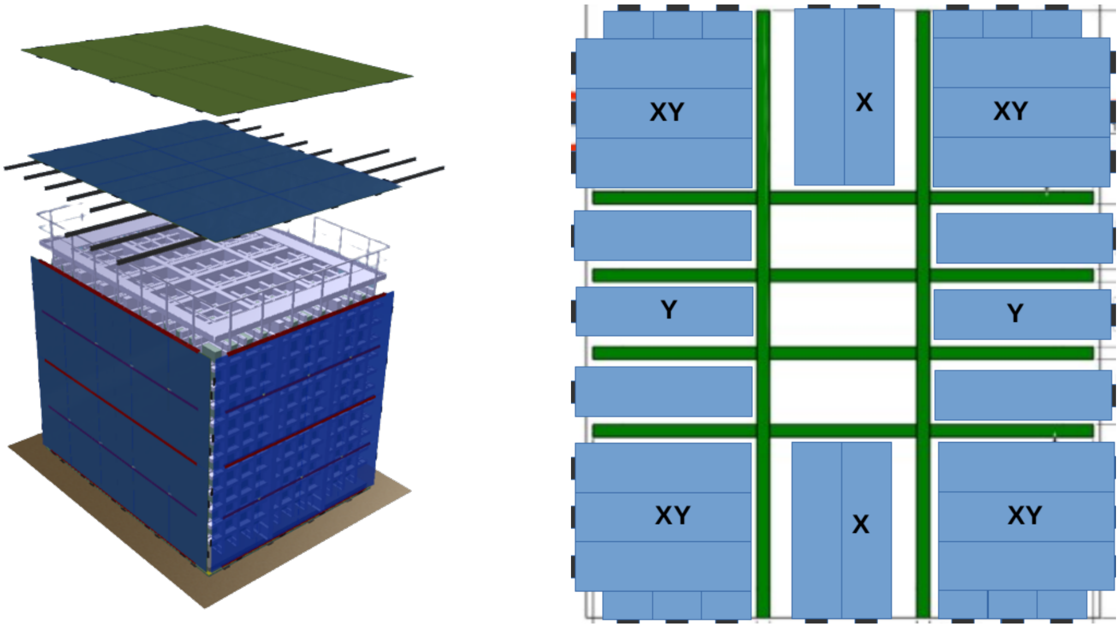


(a) CRT Strip (b) CRT Module

Figure 4.11: Diagram of a CRT strip (left) constructed of a plastic scintillator coupled to a pair or readout Silicon Photo-Multiplier (SiPM) via wavelength shifting optical fibres. A series parallel strips are combined to form a plane and perpendicular planes form a module (right). CRT hits are formed by finding coincident and overlapping hits across the planes within a module. Figure from [1]

enables reconstruction of the hit location within the 12 cm width of the strip. Sets of perpendicularly orientated strips are combined to form a single CRT plane, as can be seen in Figure 4.11b. The coincidence of hits in these overlapping strips allows for 2D determination of the particle passage through the plane.

There are CRT planes on each sides of the detector, as shown in Figure 4.12a. The bottom of the detector has limited coverage due to the placement of the cryostat feet, including some areas with only 1D coverage shown in Figure 4.12b. The top two planes form a telescopic array allowing for the tagging of vertical stopping cosmic rays that can be used for calibration.



(a) CRT Geometry (b) CRT Bottom Geometry

Figure 4.12: Schematic of the SBND CRT geometry (left) consisting of 7 planes, one on each side and a telescopic array above the detector. The reduced coverage of the bottom plane (right) is constrained by the cryostat feet, depicted in green. Figure from [1]

4.3.4 Trigger and DAQ

The decision to record activity within the detector, referred to as triggering, heavily utilises the subsystems in SBND due to their timing resolution. Many types of trigger can be defined to contribute to the physics signals, background estimation and calibration.

The main physics trigger will require a number of PMTs to pass above a given threshold in coincidence with a signal from the BNB. The same condition can also be applied in anti-coincidence with the BNB to estimate the rate of cosmic triggers during the beam spill window, vital for understanding backgrounds. Calibration triggers will utilise the CRT in coincidence with the PMTs to select event topologies of interest, for example stopping vertical cosmics and through-going cosmics parallel to the beam can be used for energy and electron lifetime calibrations respectively. External calibration devices, such as pulsers on the readout electronics or calibration laser systems, can also be used to trigger the readout for dedicated studies[1].

Once the decision is made to record an event, the data from all of the subsystems is passed to the DAQ which decodes and records the event. Alongside the trigger readout window, the trigger time plus the time required for charge to drift from the cathode to the anode, an additional 15% buffer is recorded either side. This buffer allows for identification of cosmic rays that enter the detector outside of the trigger time, as will be discussed in Chapter 5.

4.4 Simulating SBND

Many modern physics experiments heavily rely on simulations to develop and understand reconstruction, selection and physics reach. These simulations generally rely on Monte Carlo (MC) methods based on random sampling of predefined distributions. For SBND, the simulation and reconstruction stages discussed are provided via the LArSoft framework[5]. This framework allows for sharing common simulation, reconstruction, and analysis code between LAr TPCs experiments, including Argon Neutrino Teststand (ArgoNeuT), μBooNE, DUNE, SBND, and ICARUS.

The overall workflow of simulating SBND is shown in Figure 4.13. It begins with a generator, whether neutrino, cosmogenic, or BSM, generating primary particles that enter the detector. These particles are subsequently propagated through the detector, simulating the energy loss and daughter particles produced. The detector response is then simulated for both the CRT and TPC subsystems, initially calculating the yield of light and charge from the energy depositions. The propagation and detection of both the charge and light is subsequently simulated such that the waveforms produced, ideally, represent real data.

The ability to accurately predict the interactions that will occur within a detector is vital to understand the performance of the detector and reconstruction. As such, specialist event generators are used which model the underlying processes occurring, often tuned to global data fits. In SBND, two event generators are used: CORSIKA to simulate cosmic rays and GENIE to simulate neutrino interactions. Alternative generators can additionally be utilised for comparison in the future, but at the time of writing only these two are currently interfaced. An overview of these generators used in SBND will now be given.

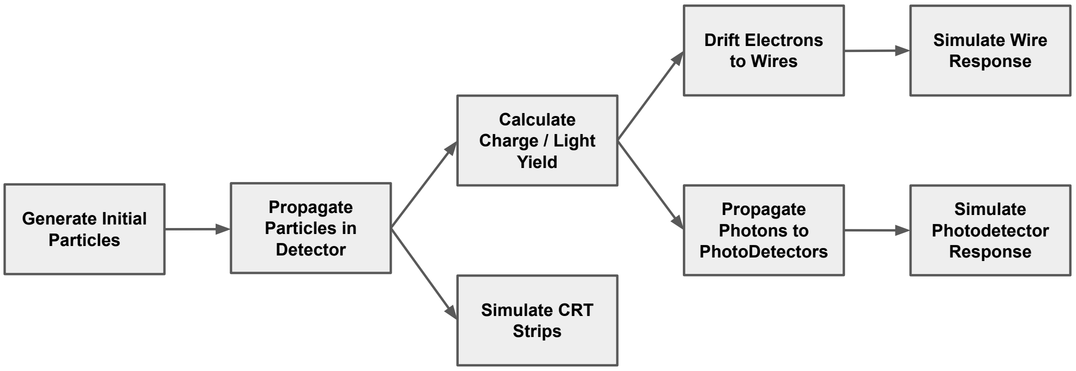


Figure 4.13: Diagram of the detector simulation workflow within SBND.

4.4.1 GENIE

The ability to accurately model neutrino interactions on argon is crucial to achieving SBNDs physics goals. The scarcity of neutrino scattering data, especially on heavy targets like argon, results in large uncertainties on the cross sections which can hamper any prospective physics measurements. The GENIE generator provides numerous theoretical and empirical models which are combined into given “tunes” by comparing the predictions to data, from both neutrino and electron scattering experiments. These tunes describe data spanning an energy range from MeV to PeV, with an emphasis on the transition between non-perturbative and perturbative regimes in the few-GeV energy range. This energy range is particularly relevant for accelerator based neutrino experiments such as SBND.

GENIE begins by selecting a nuclear model which describes the momenta and potential energy of the nucleons, modelling the effects discussed in Chapter 2. The cross section model is next included with the integrated cross section providing a probability for a neutrino interaction to occur, when combined with the neutrino flux. The differential cross section is then used to determine the interaction type and kinematics. These models include Quasi-Elastic (QE), Baryonic Resonant Scattering (RES), Coherent Scattering (COH), Deep Inealstic Scattering (DIS) and ν-e elastic scattering. Together these encapsulate scattering off the nucleus, individual nucleons, quarks and atomic electrons, depending on the relevant energy of the neutrino.

Hadronization describes the production of hadrons on free targets within the nuclei, particularly important for DIS interactions. Intra-nuclear rescattering may additionally occur as these hadrons propagate through the nucleus, modifying the observed kinematics of the interaction. This modelling of Final State Interactions (FSI) is crucial for understanding the observed event topologies of neutrino interactions, particularly for heavy nuclear targets such as argon.

In addition to the many available models and tunes available, a comprehensive reweighting scheme is provided to enable evaluation of systematic uncertainties within the models. This reweighting will be discussed in Chapter 8.

Neutrino interactions that occur outside of the active volume of the detector, referred to as “Dirt” neutrinos, can form a significant background, as will be discussed in Chapter 7. Generating these backgrounds has historically taken the, inefficient, approach to generate interactions in the “world” volume surrounding the SBND detector. Using this large volume ensures that all potential events are encapsulated but with the computational cost of simulating many interactions that will never enter the TPC. After the generation stage has been run, the particles are propagated by GEANT4 with a filter then run to select “interesting” spills. These “interesting” spills are defined as spills where there is a neutrino interaction within the TPC active volume or a dirt interaction, defined as interacting outside the TPC volume but depositing > 100 MeV of energy inside.

In order to improve efficiency, an improved workflow was implemented by the author utilising the “RockBox” functionality within GENIE. Rather than generate in the entire world volume, a restricted volume is defined for the detector volume in which all neutrino interactions are kept, as depicted in Figure 4.14. A buffer is then placed around the detector volume in which all neutrino interactions are additionally kept, this was tuned to 5 m in SBND. This buffer volume encapsulates neutrino interactions whose products, particularly neutrons and photons, could potentially deposit energy in the detector and thus pose as a background.

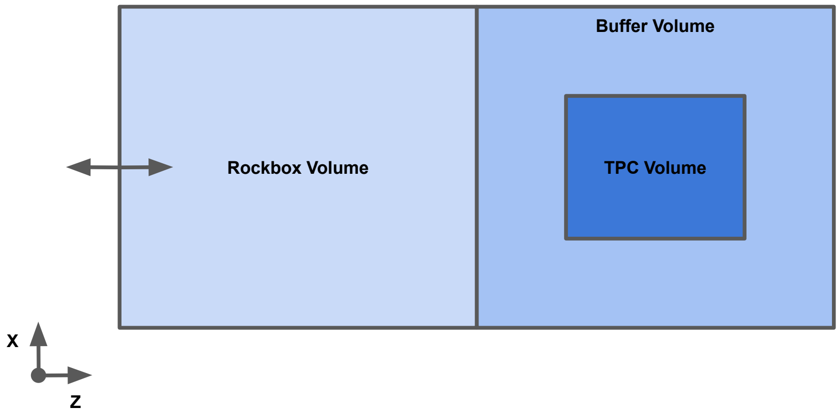


Figure 4.14: Schematic diagram demonstrating the regions used within the GENIE rockbox. All interactions within the TPC active volume and buffer zone are kept as signal and prospective backgrounds respectively. The rockbox volume is extended in the beam (Z) direction based on the lepton kinematics of the interaction.

However, this buffer is insufficient to encapsulate interactions which produce high energy leptons, specifically muons. As such, the rockbox dynamically increases the volume based on the muon kinematics, typically this extension is along the beam (Z) direction. This is performed by estimating a constant energy loss, dE/dx, for the muons produced and propagating them along their initial momentum vector. If the muon has sufficient energy, and the correct orientation, to intercept with the detector volume the interaction is kept, and it is otherwise discarded.

The position of dirt neutrino interactions, those that occur outside the active volume but deposit $>$ 100 MeV of energy inside, are shown in Figure 4.15. This demonstrates that the 5 m buffer is sufficient to encapsulate the dirt interactions in the vertical (Y) and drift (X) directions, where the active volume limits are at -200 cm and 200 cm. The beam (Z) direction has a significantly longer tail, reaching out to over 15 m before the front face of the detector, z=0 cm, demonstrating the need for the dynamic extension.

Overall, these plots show that the rockbox sufficiently encapsulates all of the dirt interactions generated in the entire world volume. The bottom right plot shows the

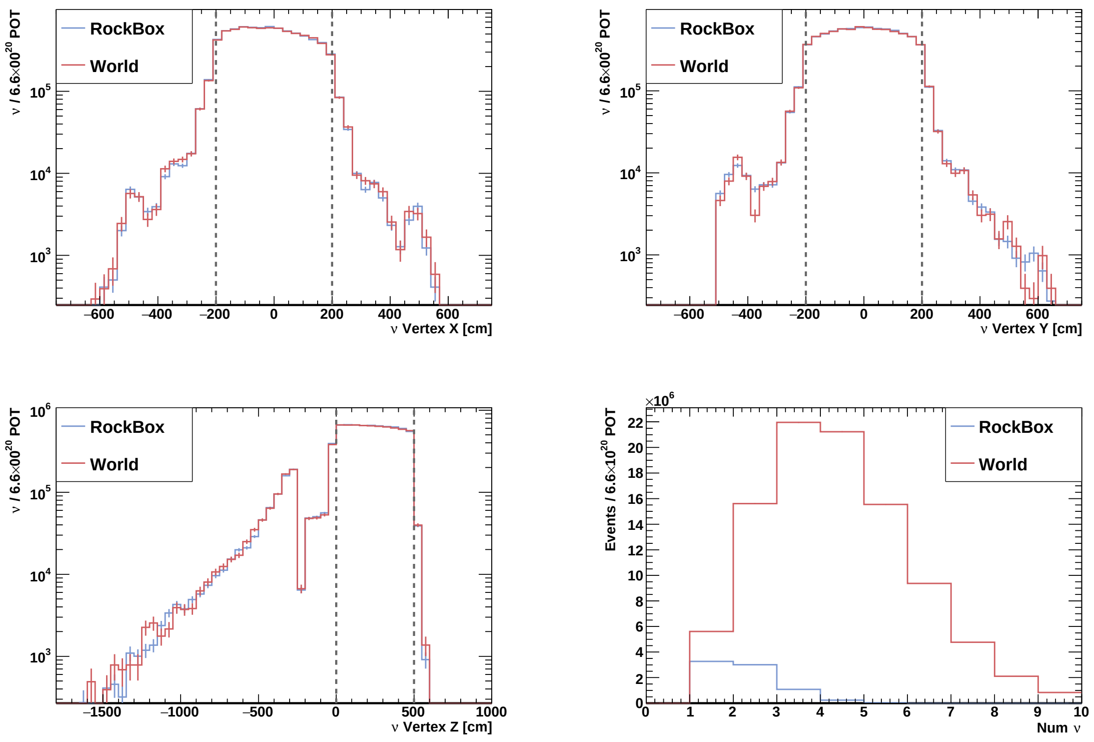


Figure 4.15: Comparison of the positions of neutrino interactions simulated in the world volume and rockbox, with the detector limits indicated by the dashed lines. The total number of interactions per spill is also shown to demonstrate the number of “uninteresting” interactions per spill.

number of neutrino interactions simulated per spill in both the world and rockbox cases. This demonstrates the reduction in “uninteresting” neutrino interactions that do not deposit any energy within the detector volume, hence showing the efficiency gain of the rockbox workflow. In addition to reducing the number of neutrinos per spill, the number of “uninteresting” spills is also reduced by a factor of 300, reducing the number of spills that are generated and subsequently discarded. These combined efficiency gains allow for dirt interactions to be included in the nominal sample for the first time in the production used for this thesis.

4.4.2 CORSIKA

As a surface level detector, it is vitally important to understand the cosmic ray background expected at SBND. In contrast to neutrino interactions, it is possible to study the cosmic ray background in anti-coincidence with the beam. This allows for a pure sample of cosmic rays to validate the simulation and constrain the expected backgrounds, once SBND has data. It is not only important to understand cosmic rays in the context of backgrounds but also for their use in calibration studies.

The CORSIKA generator used in SBND begins by simulating high energy primary particles incident on the Earth’s atmosphere, only proton primaries are used in SBND due to better agreement with μBooNE data but primaries up to iron nuclei are available. The particles are then propagated through the atmosphere where interactions with the air and secondary decays are simulated. This propagation continues until the particles produced reach the surface, in the case of SBND these are modelled to a surface just above the near detector building roof. The surviving particles are saved such that they can later be retrieved and propagated to the detector using the GEANT4, which will be discussed in the upcoming section.

The simulated flux of primary cosmic ray muons is expected to be 3.5kHz, with kinematics shown in Figure 4.16. These muons are generally downwards going, a zenith angle of 180° represents vertically downwards, with a typical energy in the few GeV range, albeit with a high energy tail.

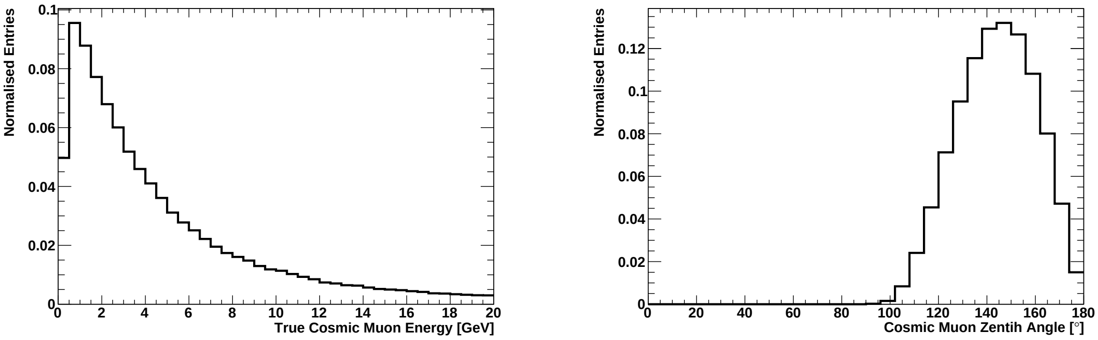


Figure 4.16: CORSIKA simulated primary cosmic ray muon energy (left) and zenith angle (right), where 0° is vertically upwards.

4.4.3 Detector Simulation

Once particles have been generated, they must next be propagated through the detector, typically this is performed by using GEANT4[163]. GEANT4 is a MC particle transport toolkit that propagates particles through the detector, including the subsystems, applying scattering, interaction, and decay physics to the particles. The result is a complete set of particle trajectories through the detector, including energy depositions in the various materials and a hierarchy of particles produced in interactions and decays.

These energy depositions are then processed through the detector simulation, applying the relevant detector effects discussed in Chapter 3. This begins with a recombination calculation using the modified box model to determine the amount of charge and light produced, dependent on both the electric field and the deposited energy density[140]. Distortions to the electric field, caused by the Space Charge Effect (SCE), are additionally applied which affect both the recombination calculation and the apparent position of the energy depositions.

The electrons that survive recombination are subsequently propagated to the readout wires of the detector, assumed to travel directly along the drift field. The charge arriving at the wire is attenuated to simulate the capture of electrons on impurities, using a characteristic electron lifetime of 10 ms. Longitudinal diffusion is applied to smear the arrival time of the electrons on the wires whilst transverse diffusion spreads a portion of the charge to adjacent wires. These energy depositions on each wire are subsequently convolved with the detector electric field and readout electronics response[164]. Noise is subsequently added based on test bench results from the detector electronics taking into account the dependence on wire length[150].

Unlike electrons, photons cannot be similarly approximated as travelling in a straight line along the drift direction due to their isotropic emission, scattering, and reflection. Analytically tracing the path of every photon is computationally expensive and not viable for producing large statistics samples. Instead, a semi-analytic approach is used to estimate the visibility of each photon detector for a given position in the detector, taking into account the scattering and reflections within the detector[149]. Sampling this visibility and taking into account the quantum efficiency of the photodetectors, the number of photons arriving at each detector is predicted, along with an estimated arrival time distribution. Waveforms of the photodetectors are subsequently produced adding noise and convoluting with the expected detector response, similarly to wires.

The CRT simulation is somewhat simpler with a direct conversion from the deposited energy to a simulated hit, skipping the simulated waveform stage. This is possible due to the low dark rate and self-triggering of the CRT strips; a CRT strip is triggered when both SiPM exceed a given threshold within a 100 ns coincidence. This is simulated by converting the deposited energy into a light yield within the scintillator and applying corrections for the collection efficiency of each SiPM based on the position of the energy deposition within the strip. The time of the hit is estimated from the time of the energy deposition combined with a light propagation delay based on the distance between the energy deposition and the SiPM.

4.5 Low Level and Subsystem Reconstruction

Reconstruction broadly encompasses the attempt to extract high level quantities about the underlying particle interactions from the raw detector data. The reconstruction workflow used in SBND is shown in Figure 4.17, beginning with waveforms from either simulated or real detector data. The first stage performed is the signal processing which aims to remove noise and other detector detector effects to recover the true energy depositions. Next, hit finding is performed to extract the individual energy depositions from the wires, this is typically the last step classified as “low level” reconstruction. Pattern recognition, dedicated reconstruction for tracks and showers, and PID is typically classified as “high level” reconstruction and will be

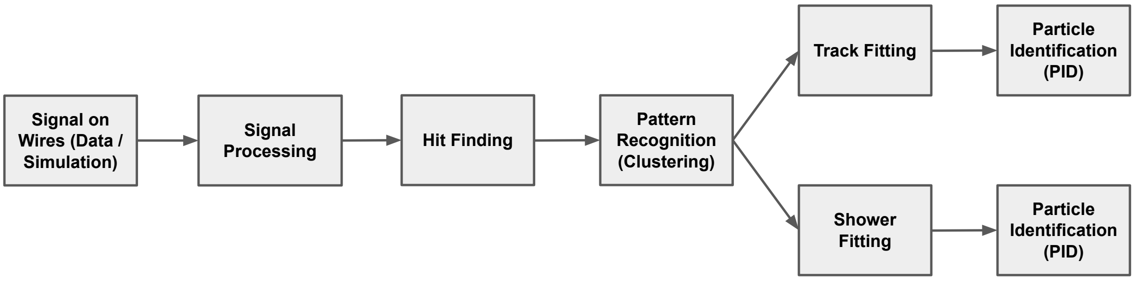


Figure 4.17: Diagram of the reconstruction workflow within SBND.

discussed extensively in Chapters 5, 6 and 7; The subsystem reconstruction, namely the CRT and PDS, follow a similar prescription but typically ends at the clustering stage. This subsystem information is then matched to TPC where the improved timing resolution can aid in the removal of cosmogenic activity, as will be presented in Chapter 7.

4.5.1 Signal Processing

In order to recover the original charge deposited on the wire, signal processing is performed to remove detector effects such as noise and deconvolve the electronics and electric field response[165]. The deconvolution is particularly important for the induction planes to convert the bipolar signal into a unipolar signal such that the integral of the waveform can be used to estimate the charge deposited. The deconvolution can be either 1-D, where only the response on a single wire is considered, or 2-D, where the response of neighbouring wires is additionally considered.

The impact of the deconvolution can be seen in Figure 4.18 which shows a neutrino interaction in μBooNE at various stages of the signal processing[165]. Before the deconvolution, shown on the left, the waveforms are biploar induction signals with large amounts of smearing and artifacting. In particular for tracks travelling vertically along a single wire large cancellations in the bipolar waveforms can be seen, leaving

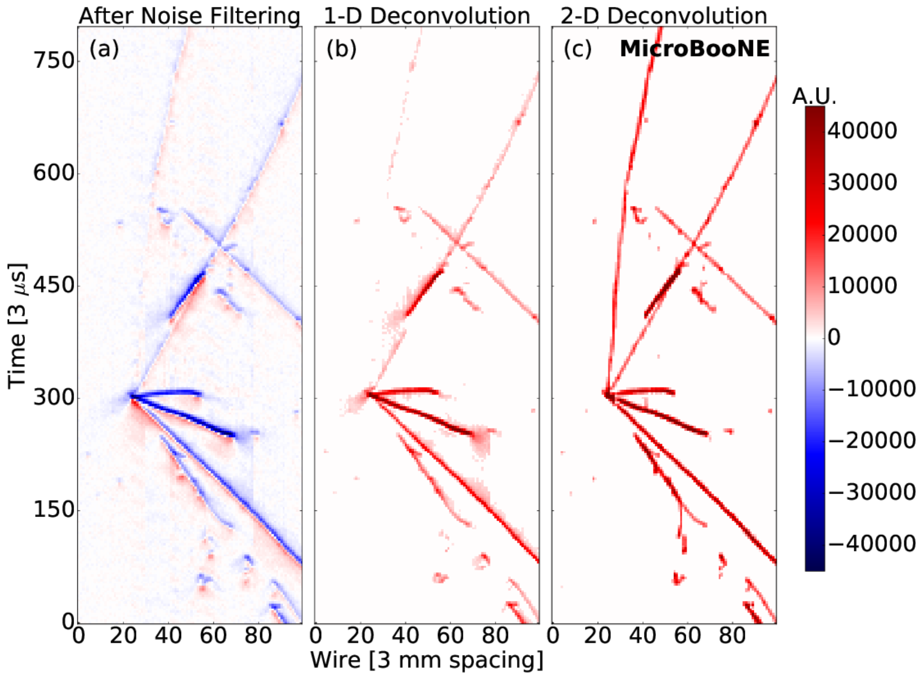


Figure 4.18: Event display from μBooNE showing the signal with only noise filtering (left) after the 1-D (middle) and 2-D (right) deconvolutions. Figure from [165]

.

tracks barely visible. The 1-D deconvolution, shown in the middle, removes the bipolar nature of the signals, but significant artifacting remains. However, the 2-D deconvolution, on the right, is able to remove the artifacting and recover the lost track travelling parallel to the wire.

In SBND, both the convolution, used in the simulation, and deconvolution, used in the signal processing, are limited to 1-D at the time of writing, although work is ongoing to upgrade this to 2-D. The deconvolution is performed using a Gaussian filter as it was found to better preserve the shape of the signal compared to a Wiener filter, despite the better Signal-to-Noise Ratio (SNR) of the latter[164].

After the deconvolution has been performed, Region Of Interest (ROI) finding is performed on the wires to zero suppress the waveforms without signal and reduce the file size by several orders of magnitude. This is performed by finding regions of the wire where the signal exceeds a given threshold, tuned such that only 10% of the ROI are noise induced. Each region then adds an additional 5 µs before and after the threshold is passed to ensure the energy deposition is fully encapsulated.

4.5.2 Hit Finding

Once the ROI have been formed, hit finding is then run to extract information on the energy depositions, specifically creating a “Hit” that encapsulates the information regarding the energy deposition. Specifically, this is performed by “GausHitFinder” module within LArSoft which attempts to fit a series of Gaussians to the waveform[166]. This begins by finding a region where the wire exceeds a given threshold and subsequently identifying the number of peaks present in this “snippet”. The number of peaks is found by analysing the differential of the waveform to find the number of maxima, that number of Gaussians are then fitted to find the centre, height and width of each.

This can be seen in Figure 4.19 which shows an electron shower and an example waveform with fitted Gaussians for the selected wire. The wire shown has four contiguous peaks above threshold, so four Gaussians are fitted and grouped into a single snippet. Track like particles typically have only a single peak and are thus represented by a single Gaussian, in the absence of delta rays.

Once the hits have been fitted, the parameters that describe the hits can be extracted and used by the downstream pattern recognition and reconstruction, as will be discussed in Chapters 5 and 6. The peak time represents the time that the charge arrived at the wires, used for determining the drift position of a hit and matching coincident hits between planes. The height and width of the Gaussian are used to calculate the integral of the pulse which represents the charge deposited on the wire, subsequently used for calorimetry as discussed in Chapter 6.

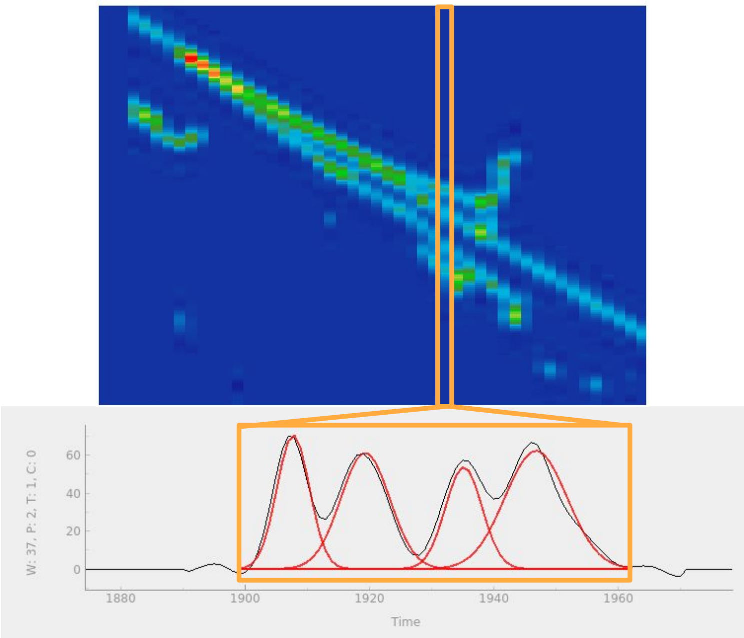


Figure 4.19: Example event display showing an electron interaction where the colour represents the charge collected on a given wire and the y axis represents time. The highlighted wire is shown on the bottom with the waveform shown in black and four fitted Gaussians shown in red.

4.5.3 Light Reconstruction

The reconstruction of the light follows a similar prescription to that of the charge. It begins by deconvoluting the observed waveform using a test bench measurement of the single Photo-Electrons (PE) response. The subsequent deconvolved waveform thus represents the number of PEs incident on the photodetector.

Optical hits are subsequently formed by finding points of the waveform above a given threshold, with the time of the hit taken as the time the threshold was passed as opposed to the central time used in the TPC. These hits sum all PEs that arrive within some time of the initial hit to provide an estimate of how much light was incident on the photodetector. This procedure is targeted to capturing the fast component of the light, as opposed to the slow component as discussed in Chapter 3, and thus does not encapsulate all of the light produced.

Optical flashes are then formed by combining coincident optical hits from optical detectors of a given type, e.g. coated or uncoated PMTs. The PE weighted centre of the optical detectors used in the hit is then used to calculate the centre of the flash in the Y-Z plane. The X position of the flash can be estimated using either the spread of the light or the ratio of the coated to uncoated PMTs, which will be discussed in more detail in Chapter 7. The ARAPUCAs are not included in the optical flashes at the time of writing, but future work will rectify this. Corrections to timing can additionally be applied to correct for the photon propagation time based on the relative position of the flash compared to the photodetectors. These flashes can subsequently be matched to TPC activity where they can provide timing information to identify activity in coincidence with the beam spill, as will be discussed in Chapter 7.

4.5.4 Cosmic Ray Tagger Reconstruction

Reconstruction in the CRTs is somewhat different to that of the TPC and PDS as the waveforms are not saved, either in simulation or data. Instead, only the time and peak height of a pulse is saved due to the high SNR, low dark rate, and self triggering of the detectors. As such, the reconstruction starts with hits on each SiPM within a CRT strip. The position within the strip is subsequently estimated from the relative PE measured on each SiPM, alongside an estimate for the total PE in the strip.

2-D hits are subsequently formed by finding coincident hits in overlapping strips within a given plane, as shown in Figure 4.11. Corrections to the PE are then made to correct for the attenuation of light along the strip. CRT tracks can additionally be made by finding coincident 2-D hits in opposing planes. The rate of tracks is significantly lower than hits due to gaps in coverage and cosmic rays that stop within the TPC.

Once CRT hits and tracks have been formed, matching to the TPC is subsequently performed, relying on TPC tracks which will be presented in Chapter 6. The matching procedure begins by applying a shift to the TPC tracks in the drift direction to account for the timing offset between the CRT hit and the trigger time. Any matches where the TPC track is shifted outside of the active volume are invalid and thus rejected. Next, a metric is calculated to determine the quality of each potential match. For CRT hits this is performed by extrapolating the TPC track to the CRT plane and calculating the Distance of Closest Approach (DCA). For CRT tracks the DCA can be calculated for both CRT planes and the angular agreement between the CRT track and TPC track can additionally be considered. The extra information afforded by CRT tracks results in a higher purity compared to CRT hits, albeit with reduced coverage.

This matching between CRT and TPC information allows one to combine the ns-level timing resolution of the CRT system with the TPC. This is particularly important for identifying activity within coincidence with the beam spill window, similarly to the flash matching previously discussed. However, tagging cosmic rays is crucial for resolving the drift-time degeneracy of LAr TPC for calibration purposes, as will be discussed in Chapter 6.

4.6 Concluding Remarks

The Short-Baseline Neutrino (SBN) program aims to conclusively address the MiniBooNE low energy excess and search for sterile neutrino induced oscillations in both the ν_e appearance and ν_μ disappearance channels. As the near detector for this program, SBND plays a key role in constraining systematic uncertainties. Alongside this role, it additionally has a rich independent physics program of cross section measurements and BSM searches. Of particular relevance to this thesis, SBND expects to take the worlds largest collection of ν-Ar interactions and thus make world leading cross section measurements.

The Booster Neutrino Beam (BNB) provides the SBN program with a high purity flux of ν_μ, with only a 0.5% ν_e contamination, at an average energy of 800 MeV. The SBND detector comprises of three key subsystems: the Time Projection Chamber (TPC), the Photon Detection System (PDS), and the Cosmic Ray Tagger (CRT). These subsystems each contain novel features to help SBND achieve its physics goals.

As with many modern physics experiments, SBND relies heavily upon simulations to understand the detector performance and ultimate ability to achieve physics goals. Many event generation, particle propagation, and detector simulation tools are shared between experiments allows for efficient and coherent development. Similarly, the low level reconstruction, namely signal processing and hit finding, that is performed for SBND is based on foundations shared between LAr TPCs. This low level reconstruction provides a foundation for which the pattern recognition and high level reconstruction build upon, as will be discussed in the forthcoming chapters.

Chapter 5

Pandora Pattern Recognition

Pattern recognition is an intermediate stage in the reconstruction pathway which takes the reconstructed hits, discussed in the preceeding chapter, as an input. This stage ultimately yields 3D reconstructed objects, each ideally representing a single true particle. This stage is performed by the Pandora pattern recognition package widely used across many Liquid Argon Time Projection Chambers (LAr TPCs)[4].

Work performed to improve the performance of Pandora in Short-Baseline Near Detector (SBND) is presented, specifically analysing the current performance, introducing Multi Variate Analysis (MVA) techniques, and writing new algorithms. These improvements target the removal of cosmic ray backgrounds, vertexing, track vs shower classification, and shower growing. This is a vitally important part of the reconstruction chain as all of the high level reconstruction and particle identification is based upon the outputs of this stage, as discussed in Chapter 6. There is currently no procedure to rectify mistakes made during the pattern recognition, so any impurity or incompleteness introduced at this stage will directly impact later physics analyses.

An overview of Pandora is first given in Section 5.1, with Section 5.2 focused on an assessment of the performance of the unambiguous cosmic removal. The impact of introducing MVA techniques into the reconstruction is discussed in Section 5.3 with 5.3.1 providing an introduction to the techniques used. The application of these MVAs is presented in sections 5.3.2, 5.3.3, and 5.3.4 for vertex finding, track vs shower classification, and neutrino slice ID respectively. Additional algorithms developed to make use of the information afforded by the MVAs to improve shower reconstruction are discussed in Section 5.4.

These developments culminate in significantly improved reconstruction efficiency and completeness for shower-like objects in turn allowing for enhanced shower energy reconstruction, as will be discussed in Chapter 6. The improved characterisation and reduced shower segmentation simultaneously increases both the efficiency and purity for selecting electron showers from ν_e interactions, as discussed in Chapter 7.

5.1 Pandora Overview

Pandora is a multi-algorithm approach to pattern recognition first developed for the International Linear Collider (ILC) and later extended to LAr TPCs, initially μBooNE[3][4]. It currently consists of over 100 individual algorithms that can be utilised, with each performing a specific role in the chain. The output is a particle flow representation of the interaction, either cosmic or neutrino, creating a hierarchy of produced particles.

There are two distinct pathways that run within Pandora: PandoraCosmic and PandoraNu, depicted in Figure 5.1. The former focuses on reconstructing cosmic rays and thus assumes that all particles will leave long tracks in the detector, with any residual electromagnetic activity being assumed to be resultant from delta rays. Conversely, PandoraNu is designed around growing interactions out of a neutrino vertex, with more complicated particle hierarchies and dedicated reconstruction for electromagnetic activity.

The reconstruction begins on each wire plane independently, referred to as 2D reconstruction, and begins with simple clustering where hits that form a continuous straight line are grouped together, assuming that everything is track-like. A series of algorithms are then run to merge these clusters together across gaps in the detector

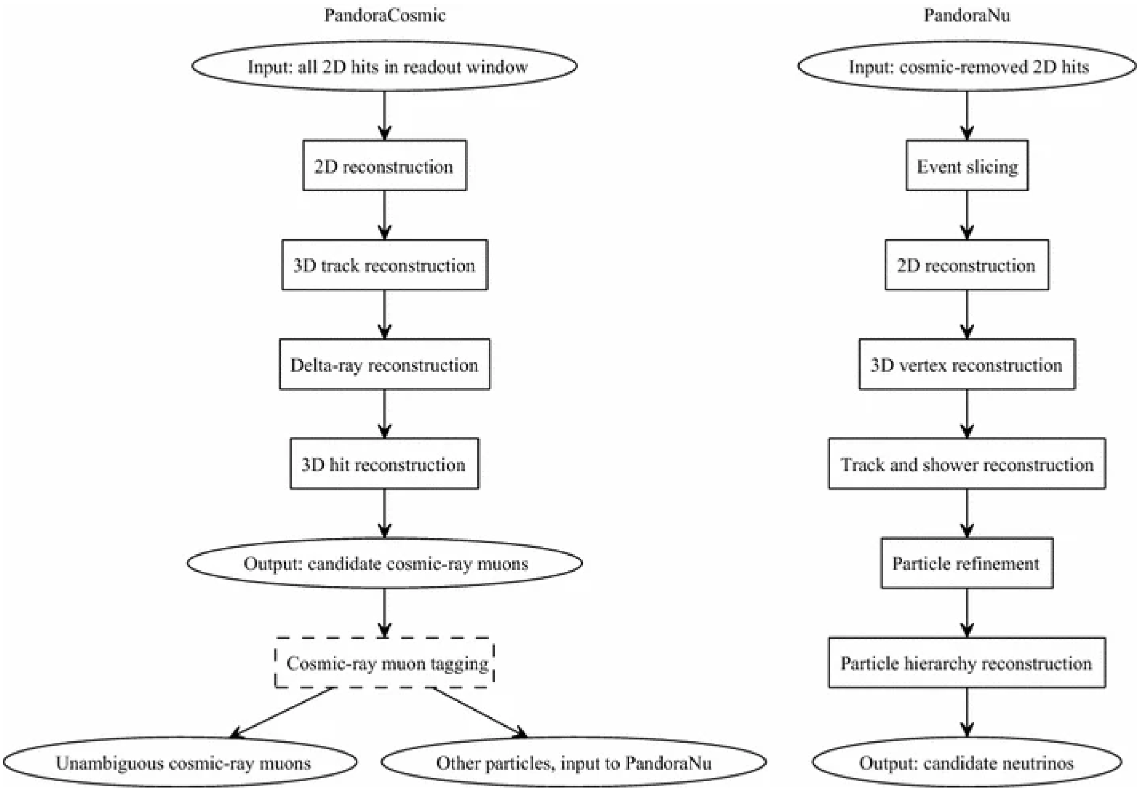


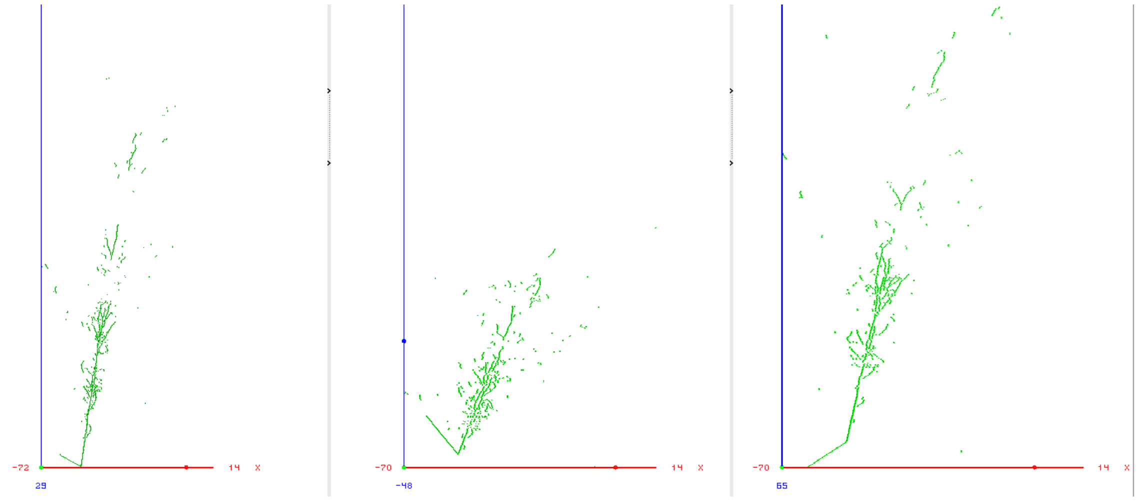
Figure 5.1: Overview of Pandora reconstruction pathways.[4]

and around ambiguities such as intersecting or interacting particles. These clusters are subsequently matched across planes by finding candidate clusters in other planes with equivalent spans in the drift direction. Once a series of these candidates has been found, an overlap between two of these planes can be found and be projected onto the third plane where a goodness-of-fit metric can be computed to quantify the consistency between the clusters. If the clusters are determined to be a good match they are combined together to form a Particle Flow Object (PFO). A series of algorithms that can merge or split clusters as required to improve this goodness-of-fit metric are also run at this stage, along with some algorithms targeting delta rays for cosmics. Pandora then attempts to find a 3D position for each hit, referred to as a space point, by considering sliding linear fits across the matched clusters.

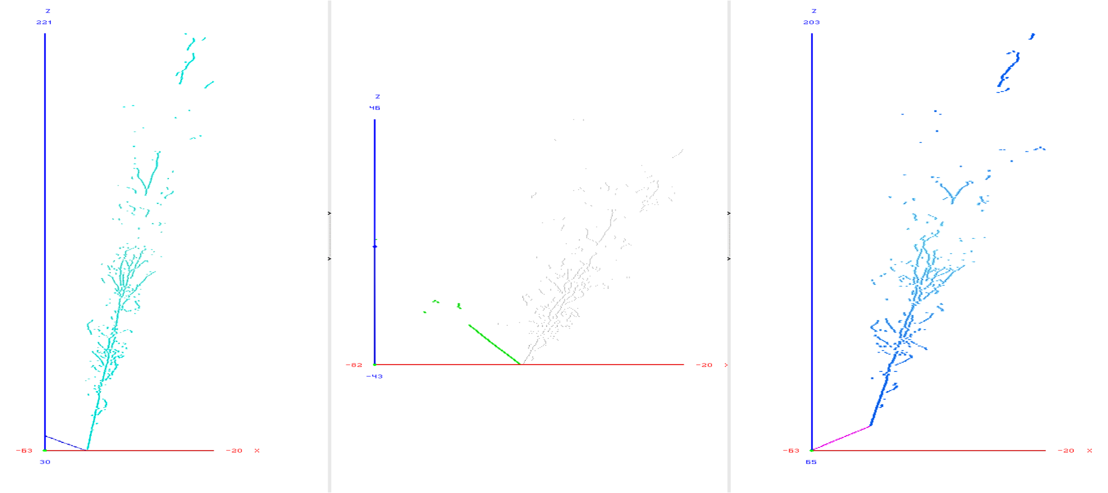
The consolidated approach first runs PandoraCosmic and removes all of the clusters, and associated hits, that are deemed to be unambiguously cosmic in nature, details on this classification are discussed in Section 5.2. The remaining clusters, all deemed to be potential neutrino interactions, are then processed by Pandora's slicing algorithm where an event is divided into sections, referred to as slices. Each slice ideally encapsulates all the hits that come from a single origin (i.e. all particles resultant from a neutrino interaction or primary cosmic ray) and are created based on the proximity and directionality of the constituent clusters.

After repeating some of the 2D clustering algorithms, Pandora next tries to find the neutrino interaction vertex for each slice, discussed in more detail in Section 5.3.2. The event can subsequently be grown outwards from this vertex using separate algorithms for tracks and showers, with characterisation algorithms classifying each cluster as track-like or shower-like at various stages of this procedure, as discussed in Section 5.3.3. Once all of these algorithms have been run, Pandora creates a Particle Flow Object (PFO) output for each 3D cluster with a vertex, a series of 3D space points (and associated 2D hits) and a particles flow hierarchy. The hierarchy starts with a “Neutrino” particle with only a vertex and series of primary daughters, each of which may in turn have distinct hierarchies.

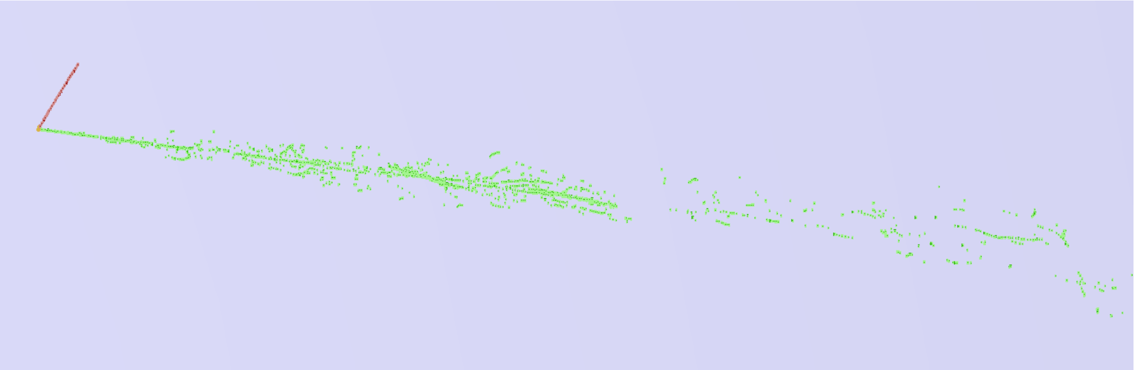
An example neutrino interaction, with a proton and electron produced, is shown in Figure 5.2 at various stages of the Pandora reconstruction. The unclustered hits that form the input to Pandora in each of the three planes are shown in 5.2a. The 2D reconstructed clusters in each plane are shown in 5.2b demonstrating that Pandora has correctly clustered the hits, as each plane has one track-like cluster and one shower-like cluster. The colours shown represent individual clusters in each plane and do not correspond across planes. Combining these planes together, Pandora produces a full 3D reconstruction of the event which is depicted in Figure 5.2c.



(a) Unclustered Hits input to Pandora



(b) 2D Clusters

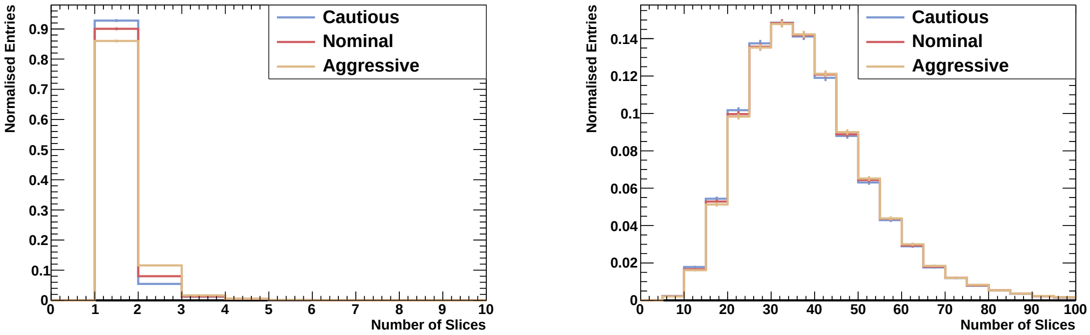


(c) Full 3D reconstruction

Figure 5.2: Example ν_e Charged Current (CC) interaction clustered by Pandora showing the U, V and W views (left, centre and right respectively) before and after clustering, with the full 3D reconstruction shown at the bottom.

5.2 Cosmic Removal

As SBND is a surface level detector, three cosmic ray muons are expected to pass through the detector in each 1.25 ms drift window which poses a significant background for the neutrino interactions. The impact of adding cosmic rays is demonstrated in Figure 5.3 which shows the total number of reconstructed slices per event for a Booster Neutrino Beam (BNB) sample with and without CORSIKA cosmic overlay. Each line shown represents a different configuration of the aggressiveness of the cosmic removal, as will be discussed in more detail later. The neutrino-only events typically have a single slice per event whereas the event with cosmic overlays show an order of magnitude increase. This demonstrates the importance of being able to reject cosmics so the downstream event selection is not overwhelmed.



(a) GENIE BNB

(b) GENIE BNB + CORSIKA overlay

Figure 5.3: Number of slices per event with, left, and without, right, cosmic overlay comparing different cosmic removal modes.

Pandora makes an attempt to remove the "unambiguous cosmic rays" from the detector to both ease the burden on the downstream cosmic rejection and clean up the event for PandoraNu. This process begins by grouping together PFOs into groups that occur from a single origin, performed by comparing the separation and directionality of the PFOs. A series of checks are subsequently performed on the PFOs within each of these hierarchies to determine consistency with a neutrino-like topology:

- **Contained:** Both end points of the PFO are contained within a fiducial volume of the detector, i.e. the distance of the end point to the closest wall is above some threshold. This identifies cosmics that enter through a detector face.
- **In Time:** Assuming that the particle occurred at the trigger time ($T_0 = 0$) the space points should all be contained within the detector volume. If any

points lie outside the volume this suggests the interaction occurred outside the trigger time, and is subsequently tagged as a cosmic. Stitching two halves of a track across the cathode provides a T_0 that can additionally be used to identify out-of-time activity.

- **Top-to-Bottom:** If the PFO has both end points within some margin of the top and bottom of the detector it is assumed to be a through going cosmic ray that does not stop within the detector.
- **Particle Topology:** The direction and deflection of the PFO are checked for consistency with cosmic ray muons, assumed to be predominantly vertical and straight, low deflection, tracks.

If all of the PFOs within a hierarchy are consistent with a cosmic hypothesis the entire hierarchy is tagged as “unambiguously cosmic” and the associated hits are removed from the subsequent PandoraNu pass. Figure 5.4 shows an example event with a single true neutrino interaction (circled) and cosmic overlay where the colour represents if the slice was deemed unambiguously cosmic (red) or a neutrino candidate (blue) by Pandora. The two TPCs are represented by the shaded regions

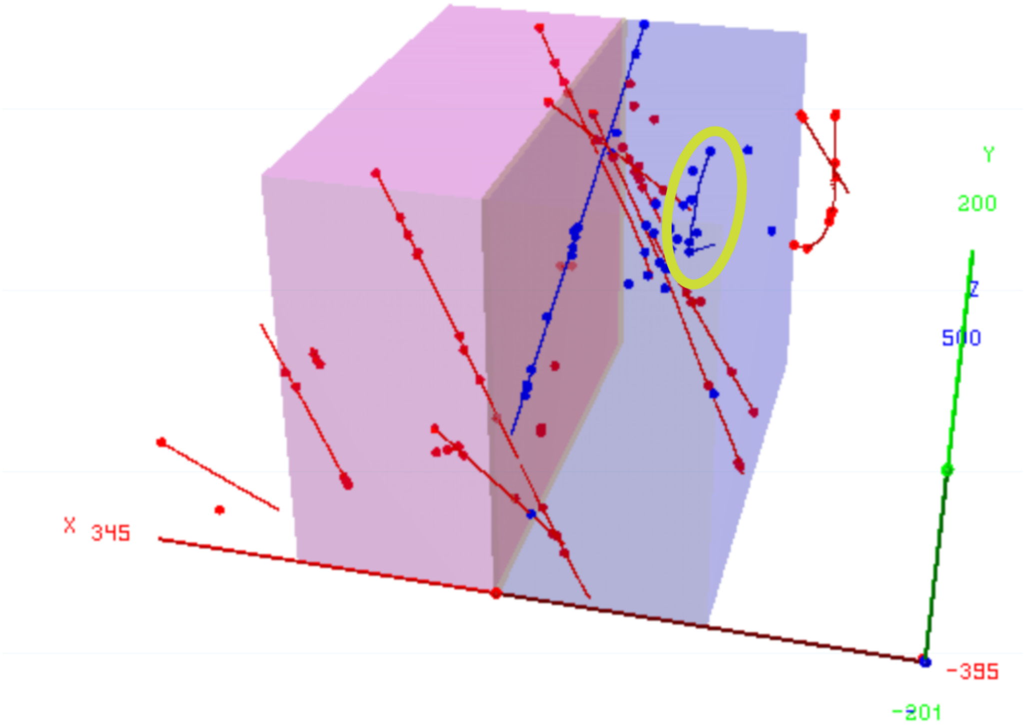


Figure 5.4: Event display of a simulated neutrino interaction (circled) with cosmic overlay showing the particles tagged as unambiguous cosmics (red) and neutrino candidates (blue) by Pandora.

allowing for the uncontained cosmics to be clearly identified. The long cosmic ray in the foreground has been stitched across the cathode, hence the small gap in the track. Several through going top-to-bottom cosmics can be seen as being labelled as cosmics in the background.

There are three available configurations of the particle topology cuts: Aggressive, Nominal and Cautious, each shown in Figure 5.3. These configurations tune the cuts that are placed on the particle topology, with Aggressive tagging more PFOs as cosmic and Cautious disabling the cuts entirely, as shown in Table 5.1. A comparison of these modes was performed to find the optimal configuration for SBND.

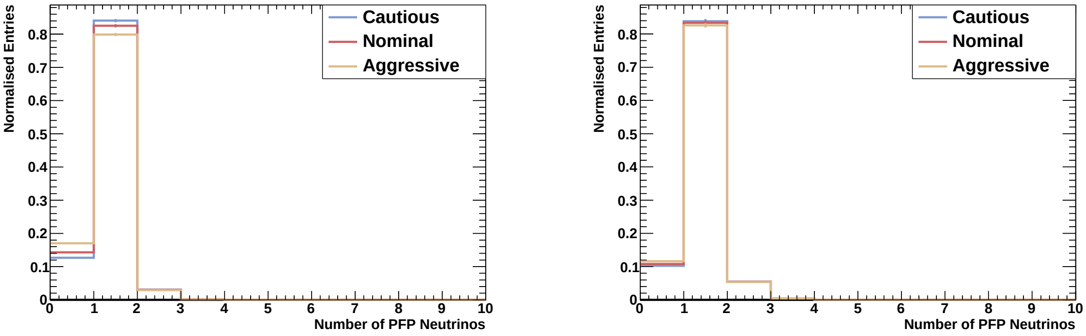
Configuration	Cautious	Nominal	Aggressive
Track cosine to vertical	-	0.6	0.6
Track curvature	-	0.04	0.1

Table 5.1: Parameter configurations for the unambiguous cosmic removal presets.

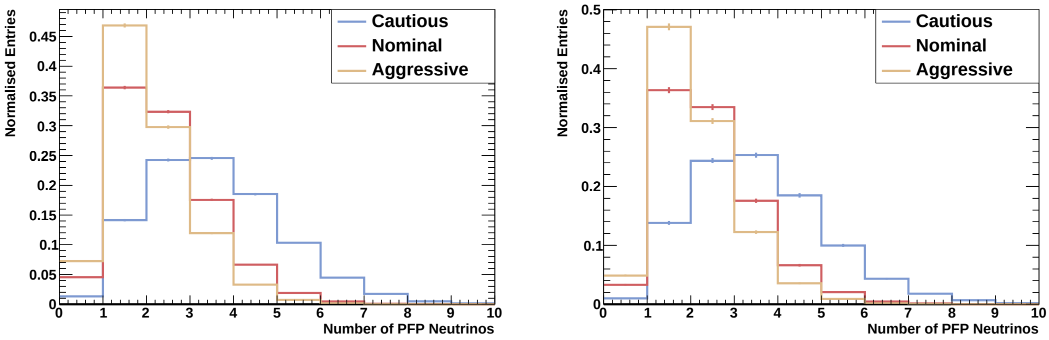
For this study, samples of CC neutrino interactions were simulated both with and without cosmic overlay. This enables assessment of both failure modes where certain topologies of neutrino interactions are removed and the impact of adding cosmic rays that may overlap the interactions. Both the full BNB sample, dominated by ν_μ, and a separate ν_e sample were analysed to identify performance differences.

The number of neutrino candidates per event for each of the configurations is shown in Figure 5.5. A large increase can be seen when adding cosmic overlay in addition to the neutrino interactions, these are cosmic induced slices that were not tagged as unambiguous and thus not removed. Running in aggressive mode reduces the total number of remaining slices compared to the nominal mode whereas cautious increases it, as expected. This difference is significantly larger for the overlay samples than the neutrino samples which demonstrates that the majority of the removed slices are from cosmic rays.

Table 5.2 summarises the efficiency to reconstruct a neutrino as any slice and as a neutrino slice, separating failure modes for not reconstructing interactions and erroneously labelling them as unambiguous cosmic rays. In the cases without cosmic overlay, the efficiency to correctly reconstruct a slice for the neutrinos is 100%. The cosmic removal running cautious mode tags almost all (99.6%) of neutrino interactions as neutrino-like, consistent for both ν_μ and ν_e interactions, demonstrating excellent efficiency. Nominal (Aggressive) mode removes an additional 2% (3.6%) of ν_μ interactions compared to 0.3% (1%) of ν_e interactions, exhibiting the expected trend. The muon tracks produced in the former are more topologically similar to a cosmic ray than the electron showers in the latter so the difference in efficiency is expected.



(a) GENIE BNB (b) GENIE Intrinsic ν_e



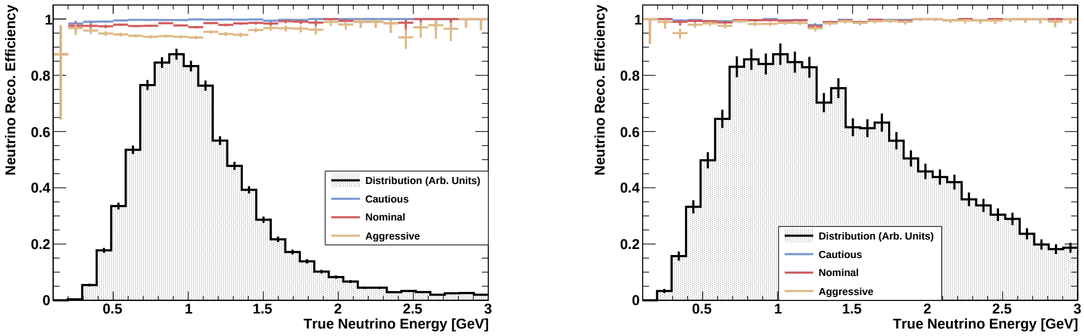
(c) GENIE BNB + CORSIKA overlay (d) GENIE Intrinsic ν_e + CORSIKA overlay

Figure 5.5: Comparing the number of neutrino candidates per event for combinations of GENIE neutrino and CORSIKA cosmic simulations.

Sample	Cautious		Nominal		Aggressive	
	Neutrino	Slice	Neutrino	Slice	Neutrino	Slice
BNB	99.7	100.0	97.6	100.0	94.0	100.0
Intrinsic ν_e	99.6	100.0	99.3	100.0	98.3	100.0
BNB + Overlay	95.5	96.2	93.1	95.9	89.6	96.2
ν_e + Overlay	95.7	96.6	96.1	97.4	95.6	97.8

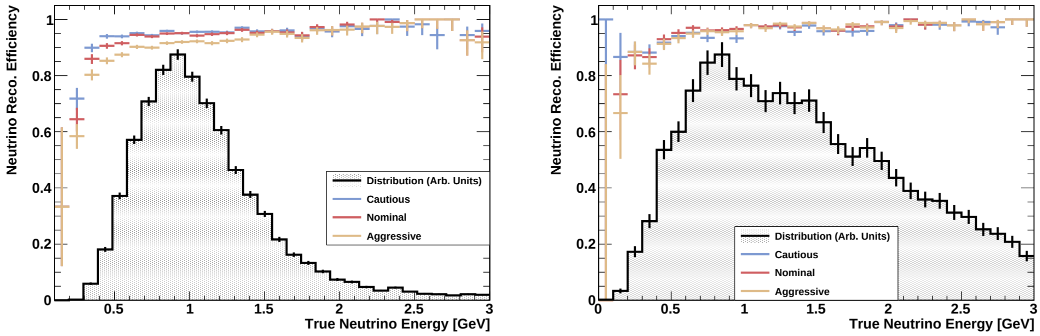
Table 5.2: Reconstruction efficiencies for neutrino interactions for any slice and as a neutrino candidate slice for different simulated samples comparing Pandora unambiguous cosmic removal presets.

A $\sim 4\%$ drop is observed for all modes when adding in cosmic overlay due to Pandora being unable to separate cosmic rays overlapping with neutrino interactions, indicated by the failure to reconstruct any slice. The difference between the cut modes remains consistent nonetheless. In total, 90.4% of cosmics are rejected by Cautious mode, 93.7% by Nominal and 94.7% by Aggressive mode, demonstrating the expected increased rejection for more aggressive modes.



(a) GENIE BNB

(b) GENIE Intrinsic ν_e



(c) GENIE BNB + CORSIKA overlay

(d) GENIE Intrinsic ν_e + CORSIKA overlay

Figure 5.6: Efficiency to reconstruct a neutrino CC interaction as a neutrino candidate as a function of the neutrino energy for combinations of GENIE neutrino and CORSIKA cosmic simulations.

Figure 5.6 profiles the neutrino efficiency of the various modes against the true neutrino energy, alongside the underlying energy distribution. In the absence of cosmics, a small drop in efficiency can be seen for ν_μ interactions around 1 GeV when running in aggressive mode, most likely to be caused by interactions producing only a muon which looks cosmic-like. Adding cosmic overlay disproportionately degrades the reconstruction efficiency at low energy for both ν_μ and ν_e alike. These interactions are absorbed into a slice with a cosmic and subsequently fail the 50% purity requirement to be classified as reconstructed.

Overall, this analysis demonstrated that the unambiguous cosmic removal is performing well, rejecting a over 90% of the cosmic ray slices whilst keeping the majority of the neutrino interactions across all preset modes. Weighing up the neutrino efficiency and cosmic background rejection, it was ultimately decided that the Nominal setting provided the most balanced performance, and was thus used for the remainder of this thesis.

5.3 Multi Variate Analysis

Multi Variate Analyses (MVAs) have been introduced at several key points within Pandora to help guide the reconstruction. Specifically, MVAs are used to help make decisions where complex topologies result in traditional cuts performing poorly, particularly for certain, typically uncommon, topologies. In these cases, using an MVA allows for the combination of many weak predictors together into a single strong predictor that performs well across a wide range of topologies. Additionally, this can be a significantly more computationally efficient approach than tuning cuts on each of the weak predictors individually, especially for correlated variables which require multidimensional tuning with exponentially increasing complexity in the number of parameters. The technical details of Boosted Decision Trees (BDTs), the type of MVA used, is presented before the application to different areas are discussed in the following subsections.

5.3.1 Boosted Decision Trees

A BDT is based upon the idea of a decision tree as depicted in Figure 5.7, this tree is taken from the BDT that will be discussed in Section 5.3.3 used to classify PFOs as track-like or shower like. The tree starts with a node where a cut is placed on a variable which defines which pathway, branch, the particle will progress along, with particles passing the cut being passed to the left daughter node and those failing to the right[167]. Typically the cut will be placed to maximise separation between

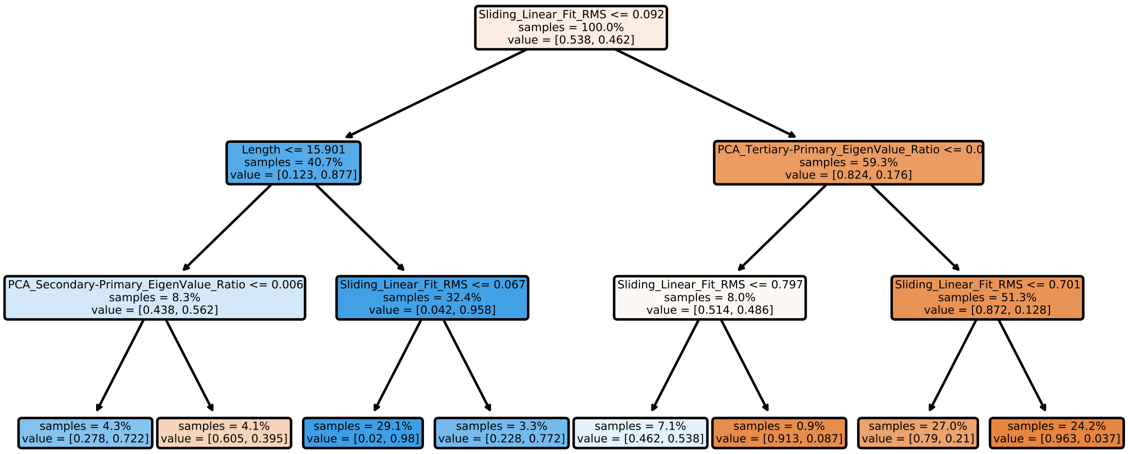


Figure 5.7: Example decision tree from the BDT discussed in Section 5.3.3 used to distinguish between tracks (signal) and showers (background). Samples represent the fraction of events present on a node and values represent the fraction of the node that is background and signal, represented by the shading being more red or blue respectively.

signal and background[168]. This process is repeated until an end node, called a leaf, is reached where the particle is classified as signal or background, depending on which sample dominates the node. A BDT builds upon this by creating a series of these trees, referred to as an ensemble, that are combined to produce a score of how signal-like or background-like a given input is. By combining many trees in this way it is possible to have trees that focus on different features of the sample, targeting different areas of phase space.

Training a BDT begins with first training a single decision tree and applying a weighting to the particles that were classified incorrectly such that the next tree to be trained will give these events more emphasis. This procedure, referred to as boosting, is then repeated for each additional tree that is added to the ensemble. The trees are then combined in a procedure called voting where each tree votes to classify the input as either signal (1) or background (-1) and score is computed from the weighted average. A cut is subsequently placed on this score to determine the classification as signal or background, distilling the information from multiple inputs into a single variable that is more easily digestible.

5.3.2 Neutrino Interaction Vertex Identification

As a neutrino interaction can occur at any point in the detector volume, finding the interaction vertex is a particularly challenging yet important aspect of reconstruction. Once a vertex is found it can be used to grow an event from, providing extremely powerful information about which clusters should be merged or split. However, using an incorrect vertex to grow an event from can lead to clusters being incorrectly merged and split, leading to a large degradation in performance of downstream algorithms. Much of the downstream reconstruction is also heavily relies on having a well reconstructed vertex, for example finding a conversion gap between a neutrino interaction and a photon shower, as will be discussed further in Chapter 6.

In order to find the neutrino interaction vertex, Pandora first creates a list of vertex candidates by matching the end points of each cluster with the end points of clusters in other planes. A series of scores, each focused on a different aspect of the event, are then calculated to score each vertex candidate:

- **Energy Kick Score:** Combines the transverse energy with the distance of closest approach between the vertex candidate and clusters.

- **Asymmetry Score:** Compares the number of upstream and downstream hits, with respect to the beam direction, to suppress candidates placed along tracks.

- **Beam Deweighting Score:** How upstream the vertex is relative to the beam direction as the majority of primary particles produced are expected to travel downstream.

The final score is then calculated from the product of these three individual scores and the candidate with the highest score is chosen.

A MVA can instead be used to choose the best vertex candidate by combining these individual scores with other predictors. The MVA was designed, originally in μBooNE, to compare two candidates directly and choose the best, rather than rank them by score. This process is then repeated comparing the leading candidate to all remaining candidates. The work presented here represents the application, training, and validation, of this MVA in SBND.

This process is split into two distinct parts: region and vertex finding. The region finding aims to find the approximate area of the vertex by collecting all vertex candidates within a given distance together and chooses a representative candidate for the region. The representative candidates from the first two regions, ordered by beam deweighting, are compared with the MVA choosing the best one to be compared to the next candidate, this process is then repeated for all region candidates. The vertex finding then chooses the best candidate within the selected region, using the same repeated comparison method. This not only reduces the combinatorics of comparisons needed but allows for more specialised MVAs to be trained: the first focusing on differentiating between primary and secondary interaction vertices and the second focusing on picking the most appropriate candidate in an area with many low energy particles.

The features given to the MVAs are split into "Event" and "Candidate" features. The former are event wide metrics designed to break up the phase space to allow for the MVA to perform well for different interaction topologies. The latter are specific to each candidate that are being compared. For example, a high "Event Showeryness" may indicate to the MVA that the "Shower Asymmetry" score should be given particular importance. These features are then passed to the MVA in the following form:

$$\langle \textbf{Event Features} \rangle \langle \textbf{Candidate 1 Features} \rangle \langle \textbf{Candidate 2 Features} \rangle$$

The MVA then returns a score from the comparison, with a positive score preferring candidate 1 and a negative score preferring candidate 2.

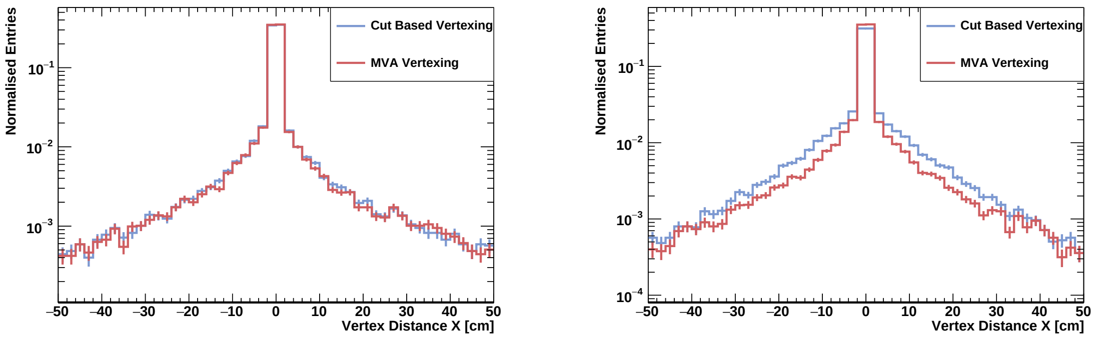
The event features are:

- **Event Showeryness:** Fraction of hits in the slice belonging to shower-like clusters
- **Event Energy:** Total charge in the slice
- **Event Volume:** Volume of cuboid that encompasses all hits in the slice
- **Longitudinality:** Ratio of Z extent of the slice to the larger of the X or Y extents
- **Number of Hits:** Total number of hits in the slice
- **Number of Clusters:** Total number of clusters in the slice
- **Number of Vertex Candidates:** Total number of vertex candidates in the slice

The vertex features are modifications to the scores previously discussed, with the asymmetry score being split into local, global and shower asymmetries. The local and global scores differ only in the distance from the candidate that hits are considered in and the shower score only considers hits from shower-like clusters.

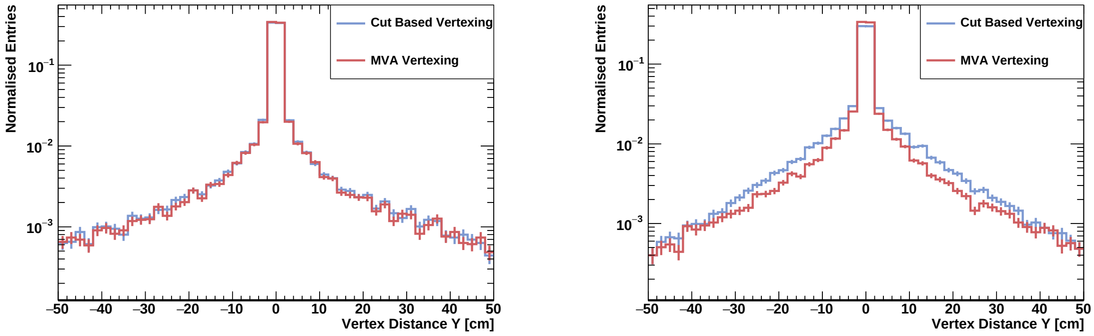
For optimal performance in SBND, the same feature set that was developed for μBooNE was used but the MVA retrained using neutrino interactions simulated in SBND. In order to produce the training samples, the best vertex candidate was defined as the one closest to the true vertex, and then compared to each other candidate in turn. Distributions for each of the input variables can be found in Appendix A.

A comparison of the vertexing performance between the standard, cut based, and the MVA approaches are shown in Figure 5.7. This shows the distance between the true interaction and reconstructed vertices for both ν_μ and ν_e, showing the distance in X, Y, Z, and 3D separately. The mean and standard deviation quantifying the

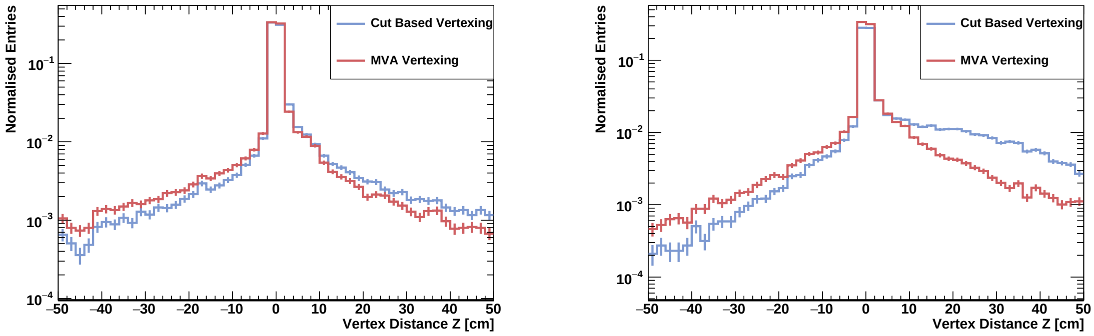


(a) BNB X Displacement

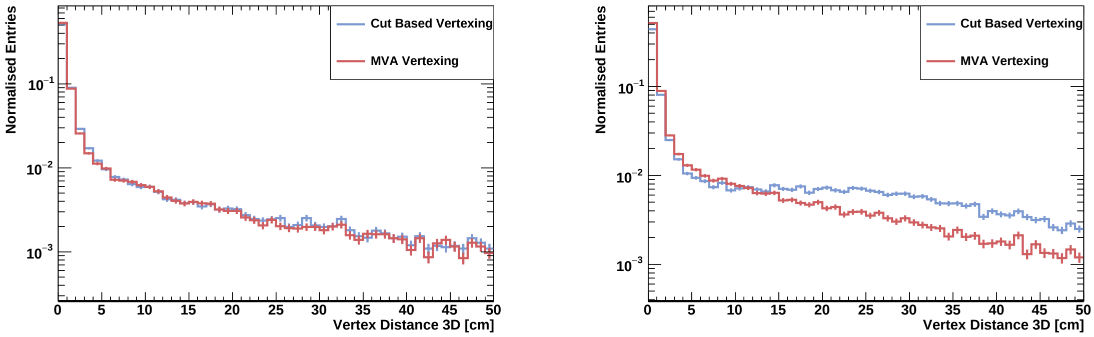
(b) ν_e X Displacement



(c) BNB Y Displacement (d) ν_e Y Displacement



(e) BNB Z Displacement (f) ν_e Z Displacement



(g) BNB 3D Distance (h) ν_e 3D Distance

Figure 5.7: Comparison of the distance between true and reconstructed vertices for the cut based and MVA approaches for BNB (left) and ν_e interactions (right).

performance is shown in Table 5.3. Overall, introducing the vertex MVAs yields a 8% (40%) decrease in the average displacement for BNB (NuE) interactions, somewhat closing the performance gap between the two neutrino flavours. In particular the bias in the beam (Z) direction is reduced by roughly 80% when using the MVA. This is particularly evident ν_e interactions where there was previously a tendency to place the vertex at the end of the initial track of the shower, which has largely been eliminated by introducing the MVA.

Sample	Vertexing	X [cm]	Y [cm]	Z [cm]	3D [cm]
BNB	Standard	-0.0 (7.0)	-0.0 (7.6)	1.0 (8.4)	3.9 (8.8)
	MVA	-0.0 (6.9)	-0.0 (7.4)	0.1 (8.3)	3.7 (8.5)
NuE	Standard	-0.1 (8.4)	0.0 (8.8)	4.6 (11.9)	7.9 (12.6)
	MVA	-0.1 (7.1)	-0.0 (7.6)	0.9 (8.6)	4.8 (9.6)

Table 5.3: Vertexing bias (resolution) quantified by the mean (standard deviation) for BNB and NuE samples comparing between traditional and MVA vertexing.

The ability to correctly identify the vertex is crucial to the subsequent stages of the reconstruction. Implementing the MVA significantly improves the performance for ν_e interactions, with a 40% reduction in average displacement. This, in turn, effects the downstream high level shower reconstruction as will be presented in Chapter 6.

5.3.3 Particle Flow Object Characterisation

The Particle Flow Object (PFO) characterisation defines whether a given PFO is considered track-like or shower-like by Pandora. This classification specifies whether the cluster will be grown as a track, using sliding linear fits, or a shower, using cones. Misclassifying a cluster can lead to tracks with very low purity or showers with low completeness. This classification is repeated at several times within Pandora as more information becomes available, for example moving from 2D clusters to 3D PFOs. The final classification is particularly important as it defines the pathway for higher level reconstruction performed after Pandora, as discussed in Chapter 6.

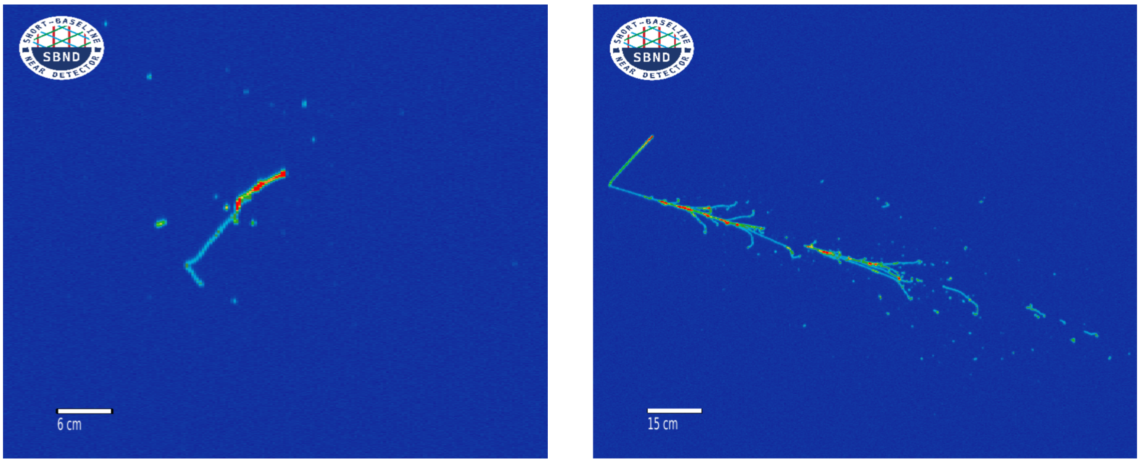
It is important to note that the correct classification of low energy electrons and photons is somewhat ambiguous. Above 100 MeV these particles lose the majority of their energy via radiative effects but below this threshold much of the energy can be lost via ionisation, as discussed in Chapter 4[137]. An example of an ambiguous, low energy electron that looks like a track is shown in Figure 5.8a in comparison to a higher energy electron shown in Figure 5.8b which is unambiguously showering.

The traditional method previously employed by SBND was using a series of cuts to characterise the PFOs:

- **Min Hits:** Any PFO with less than 6 hits is considered a shower.
- **Straight Line Length:** The distance between extremums of a PFO is calculated with above 80 cm being classified as a track.
- **Path Length Ratio:** Path length is calculated by stepping along a sliding linear fit. If the ratio of path length to straight line length is above 1.05 it is classified as a shower.

- **Track Width Ratio:** The maximum transverse displacement from a sliding linear fit is found. If the ratio between the width and straight line length is greater than 0.05 it is considered a shower.
- **Vertex Distance:** The vertex distance is calculated from the displacement between the PFO start and the neutrino vertex. If the ratio of this distance to the straight line length is above 0.5 it is considered a shower.
- **Shower Fit Width:** A sliding fit is performed to the transverse edges of the PFO. The width can the be calculated from the average transverse distance between the edge fits. If the ratio of the fit width to the straight line length is above 0.35 it is considered a shower.

However, these cuts have shortcomings in particular areas of phase space. To demonstrate this the distribution of true energy deposited by various particles, broken down as to how the particle was reconstructed and characterised, is shown in the left column of figures 5.9 and 5.10 for tracks and showers respectively. The fraction of particles not reconstructed is largest at lowest energies, as these smaller particles leave less energy deposits in the detector and are thus harder to pick out. There are two areas where showers are often misclassified as tracks: at low deposited energy, below 200 MeV, and at high energy, above 1000 MeV. The latter is an example of where a cut that can generally be useful in distinguishing between tracks and showers, the straight line length, can perform poorly for specific topologies, high energy showers. The classification performance for track-like particles is, generally, much better than for showers. The main area for improvement is the low energy tracks which are often misclassified as showers, particularly protons.



(a) Low energy shower (b) High energy shower

Figure 5.8: Event displays comparing low and high energy electrons, left and right respectively

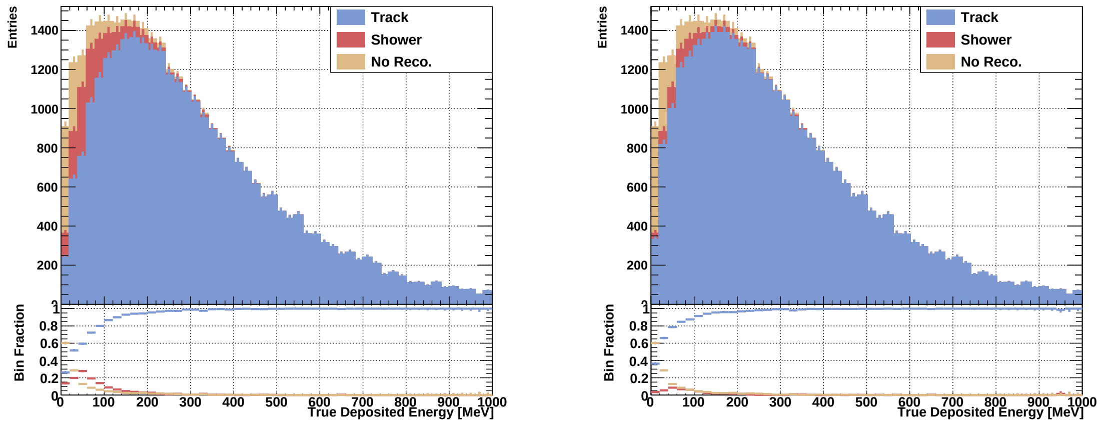
In order to overcome the issues with the cut based characterisation an MVA was introduced to give each PFO a score which determines whether it is considered as track-like or shower-like. This combines many of the variables used in the cut characterisation with several other predictors. The MVA was applied to only the final classification performed, as this defines the high level reconstruction pathway the particle will progress down.

Two of these variables use Principal Component Analyses (PCAs) to quantify the spread of the space points. A PCA first finds the principal axis that that has the largest spread of points along the axis[169]. This is a process which is repeated for subsequent axes, in this case for the secondary and tertiary axes, with each subsequent axis maximising the spread of points whilst remaining orthogonal to the previous axes. Distributions for each of the input variables can be found in Appendix B.

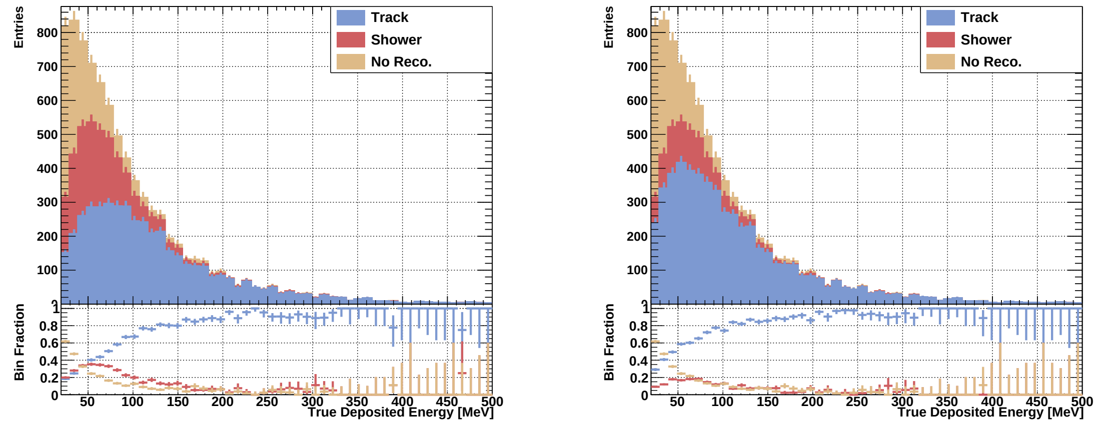
The features used are:

- **Length:** Straight Line Length used in the cut.
- **Fit Distance:** Mean transverse displacement from sliding linear fit.
- **Fit Gap Length:** Maximum gap between points in sliding fit.
- **Fit RMS:** RMS of transverse displacement from sliding linear fit.
- **Vertex Distance:** Vertex Distance used in the cut.
- **PCA Secondary Ratio:** PCA is performed and the ratio of primary to secondary eigenvalues is computed.
- **PCA Tertiary Ratio:** PCA is performed and the ratio of primary to tertiary eigenvalues is computed.
- **Open Angle Diff:** Difference between the opening angles of the first and second half of the PFO.
- **Charge Relative Spread:** Relative spread of charge (σ/μ).
- **Charge End Fraction:** Fraction of charge in the last 10% of the PFO.

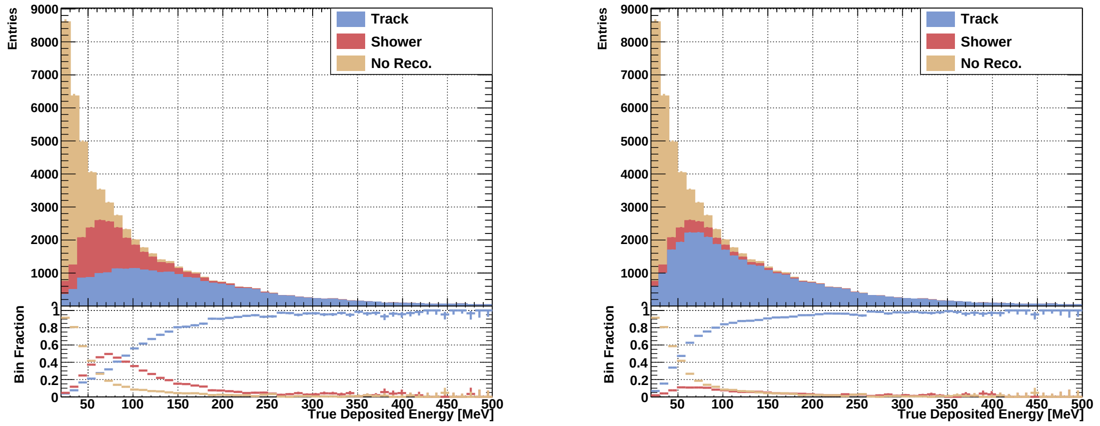
The performance of the MVA can be seen in the right column of figures 5.9 and 5.10 with the standard, cut based, results shown in the left column. These show the distribution of true deposited energy for different particle types broken down by how the PFO was reconstructed and characterised. The fraction of particles that are not reconstructed remains the same regardless of the characterisation used, as only the classification of particles is changing.



(a) Muons with cut characterisation (b) Muons with MVA characterisation

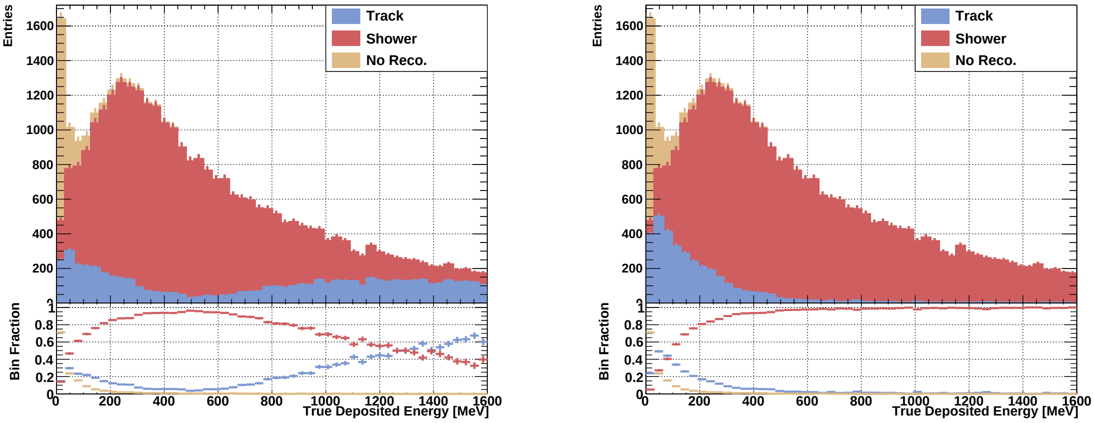


(c) Pions with cut characterisation (d) Pions with MVA characterisation

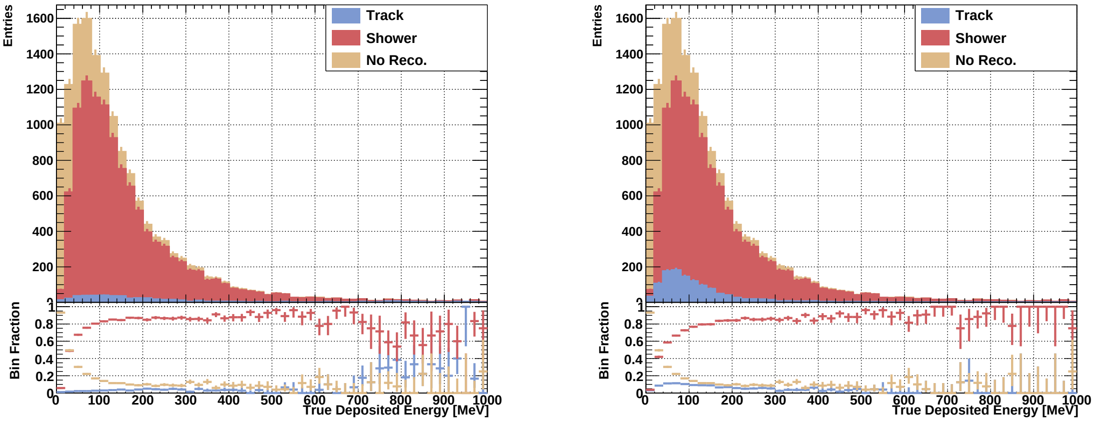


(e) Protons with cut characterisation (f) Protons with MVA characterisation

Figure 5.9: Stacked histograms showing how particles were reconstructed and classified for muons, pions and protons, top, middle and bottom respectively, with ratio plots showing the relative fraction of each bin. Comparing cut based and MVA characterisation, left and right respectively.



(a) Electrons with cut characterisation (b) Electrons with MVA characterisation



(c) Photons with cut characterisation (d) Photons with MVA characterisation

Figure 5.10: Stacked histograms showing how particles were reconstructed and classified for electrons and photons, top and bottom respectively, with ratio plots showing the relative fraction of each bin. Comparing cut based and MVA characterisation, left and right respectively.

A large improvement can be seen for high energy showers, with almost all being correctly classified when using the MVA where the cut based approach previously made many mistakes. However, this comes at the cost of a larger fraction of low energy showers being classified as tracks. Whether this is correct is somewhat unclear due to the previously discussed ambiguities regarding the true classification of such low energy electrons and photons. Indeed, the performance for low energy tracks is dramatically improved, particularly for protons, when using the MVA compared to the cut based approach.

To quantify the performance of the characterisations the fraction of reconstructed particles correctly classified is shown in Table 5.4. This shows that using the MVA to classify the particles yields large improvements, with 25% and 15% uplifts in correct

Particle	Electron	Photon	Muon	Pion	Proton
Cut Characterisation	83.8	95.7	94.6	69.4	64.4
MVA Characterisation	87.2	89.7	98.0	84.5	90.5

Table 5.4: Percentage of reconstructed particles that were correctly classified by each of the characterisation methods for different true particles.

classifications for pions and protons respectively. A more moderate 4% gain is found for muons but there is much less room for improvement here.

Whilst there is a small uplift, 4%, for electrons the performance for photons actually drops by 6% due to the degradation in performance at low energies. In order to avoid the ambiguities surrounding the true classification of low energy electrons and photons, minimum true deposited energy cuts were placed on each of the particles, as shown in Table 5.5. The efficiencies for particles passing this cut increases for all particle types when using the MVA relative to the cut based approach, with particularly large improvements for electrons, pions and protons of 14%, 14% and 25% respectively.

Particle	Electron	Photon	Muon	Pion	Proton
Energy Cut [MeV]	200	200	20	20	40
Cut Characterisation	81.3	93.6	94.9	70.1	66.0
MVA Characterisation	95.1	95.5	98.1	84.5	91.4

Table 5.5: Minimum true deposited energy cuts and percentage of reconstructed particles that were correctly classified by each of the characterisation methods for different true particles.

This increase has a significant impact on the ability to correctly characterise and ultimately select the particles and make physics measurements. In particular for the ν_e selection, the improved characterisation for electrons will correlate to an improved signal acceptance whilst the reduction in misclassified tracks will yield a decrease in the number of background candidates, as will be discussed in Chapter 7.

5.3.4 Neutrino Slice Identification

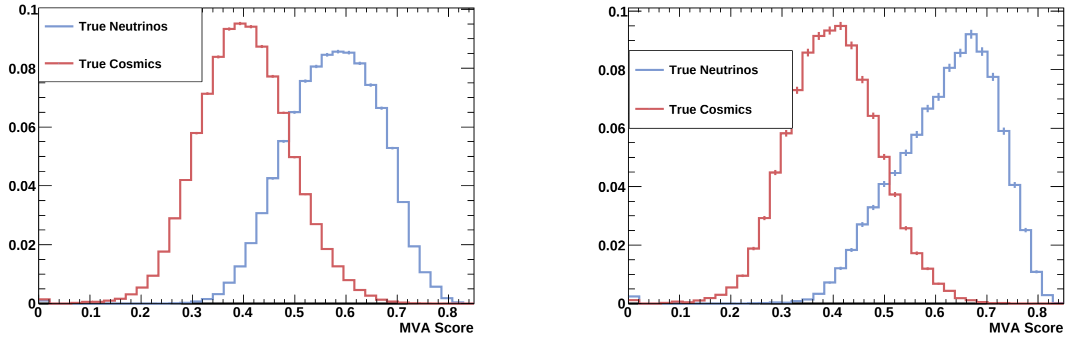
The cosmic removal that is discussed in Section 5.2 is aimed at removing unambiguous cosmic rays which leaves a significant number of ambiguous cosmic rays as neutrino candidates, alongside the true neutrino interactions. Each of these neutrino candidates are then run through the PandoraNu and reconstructed as if they were neutrino interactions. Once this pass has been performed another attempt can be made at identifying cosmic slices, utilising the additional information available.

This is performed using an MVA to assign a "Neutrino Score" to each neutrino candidate slice, with a low score indicating it is more likely a cosmic and a higher score a neutrino interaction. This score is based purely on topological arguments and does not use any calorimetric or subsystem information. Distributions for each of the input variables can be found in Appendix C. The variables are designed to analyse how much the slice is dominated by a single large track, like a cosmic, compared to how much multiple particles emerge from the vertex of the slice, like a neutrino interaction:

- **Final State PFOs:** The number of "primary" PFOs originating from the interaction vertex.
- **Total Num Hits:** The total number of hits in the slice.
- **Vertex Y:** The Y (vertical) position of the neutrino vertex.
- **Weighted Dir Z:** Average Z (beam) displacement of space points in the slice from the neutrino vertex.
- **Num. Space Points in Sphere:** Number of space points within a 10 cm sphere around the neutrino vertex.
- **Eigen Ratio in Sphere:** Ratio of secondary to primary eigenvalues of space points within 10 cm sphere, describing how much the points lie in a straight line.
- **Longest Track Y Direction:** The Y (vertical) projection of the direction of the longest track in the slice.
- **Longest Track Deflection:** How much the space points of the longest track lie along a straight line.
- **Longest Track Hit Fraction:** The fraction of hits in the slice that belong to the longest track.
- **Longest Track Num Hits:** The number of hits in the longest track.

As SBND has such a high neutrino flux, the number of events with several neutrino interactions, referred to as pile up, is non negligible at 4% of events. This means that simply choosing the most neutrino-like slice per event would not only reject neutrino interactions but also introduces the possibility of additional biases. As such, Pandora is configured to pass all of the neutrino candidates to analysers along with the "Neutrino Score" so that the decision can be made when the subsystem information is available, which will be discussed in Chapter 7.

The score distributions of the slice ID MVA can be seen in Figure 5.11, showing clear separation between cosmic and neutrino slices. Performance for ν_e is better than that for ν_μ which is expected as the electron showers produced in the former are more topologically distinct from cosmic rays than the muon tracks produced in the latter. The efficiencies and background rejections achievable by placing various cuts is also summarised in Table 5.6. A conservative cut, 0.4, removes only 1.8% (0.7%) of BNB (ν_e) signal respectively whilst reducing cosmics by over a third. More aggressive cuts are able to remove more cosmic background at the expense of neutrino signal. The optimal cut will depend on the requirements of the analysis so the score is persisted and the slice treated as a neutrino interaction, allowing for the decision to be made later in the workflow as will be discussed in Chapter 7.



(a) GENIE BNB

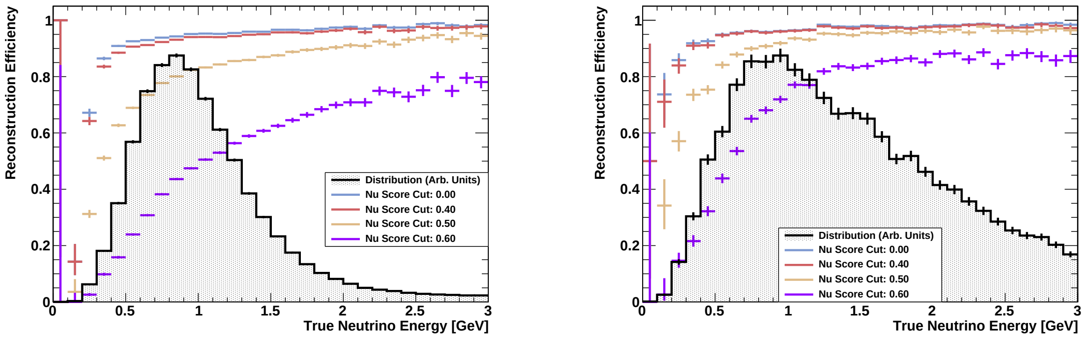
(b) GENIE intrinsic ν_e

Figure 5.11: SliceID MVA score distributions for cosmic rays and neutrino interactions.

Sample	Score Cut	Cosmics Accepted	Neutrinos Accepted
BNB + Overlay	0.40	65.16	98.16
	0.50	20.71	81.76
	0.60	2.49	43.90
NuE +Overlay	0.40	64.41	99.32
	0.50	19.68	90.93
	0.60	2.12	65.96

Table 5.6: Percentage of slices removed by placing various cuts on the SliceID MVA score for both cosmics and neutrinos.

It is important to check not only the amount of signal that is being removed but if any biases are introduced when cutting on the “Neutrino Score”. Figure 5.12 shows the efficiency of applying different cuts profiled across the true neutrino energy, with the underlying energy distribution also shown for both BNB and NuE. Cutting more harshly on the “Neutrino Score” removes low energy interactions disproportionately due to the lower vertex activity upon which the MVA relies. Thus, care should be taken when applying this cut not to introduce additional biases to the efficiency.



(a) GENIE BNB

(b) GENIE Intrinsic ν_e

Figure 5.12: Reconstruction efficiency comparison between different SliceID cuts.

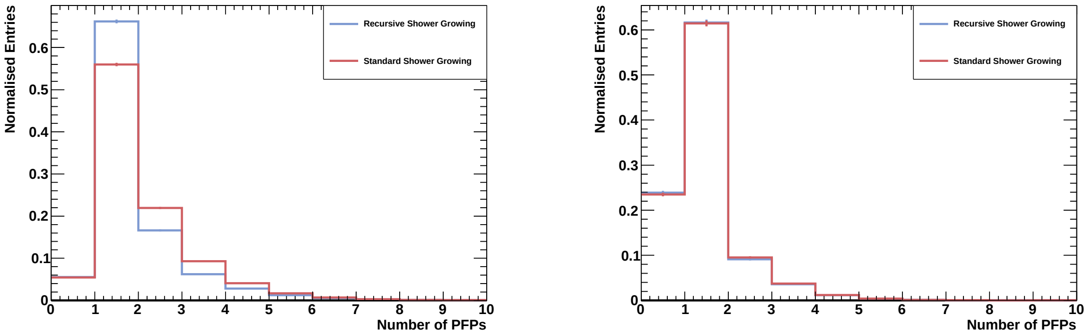
Overall, introducing the neutrino slice identification MVA complements the unambiguous cosmic removal already present in Pandora. It is able to reduce the remaining cosmic background by over a third with an almost negligible efficiency loss for neutrino interactions. More aggressive configurations allow for selecting much purer samples of neutrino interactions at the expense of signal efficiency. This MVA is another tool that can be utilised along with the Cosmic Ray Tagger (CRT) and Photon Detection System (PDS) subsystems to reject cosmic ray backgrounds, with the application of this score discussed in Chapter 7.

5.4 Recursive Shower Growing

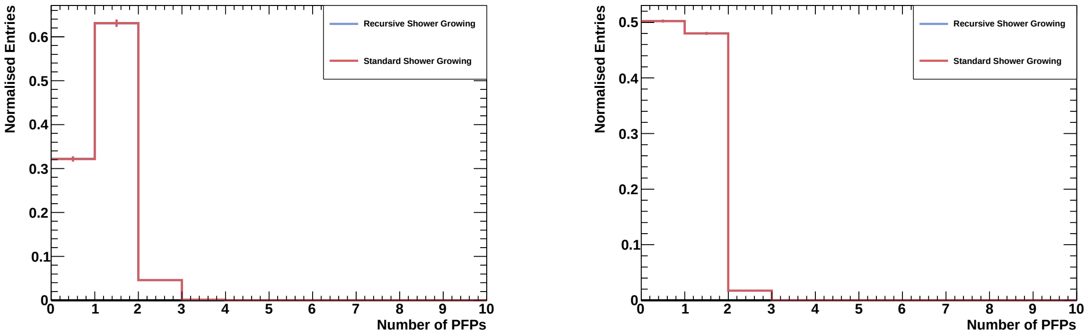
After applying the PFO characterisation BDT discussed in Section 5.3.3, the issue with high energy showers being misclassified as tracks was largely eliminated. Nevertheless, these PFOs had been considered tracks throughout the pattern recognition process which leads to many incomplete and fragmented showers, as can be seen in Figures 5.13. This shows the number of reconstructed particles that match back to a given true particle: ideally this would be one, with zero representing a particle that was not reconstructed at all and more than one a particle that was split into several PFOs. The electrons are the most severely segmented particle due to the previously discussed misclassification issues. In order to correct this, a new algorithm was developed to rerun the “Mop up” algorithms after the MVA classification has been performed, referred to as the “Recursive Shower Growing Algorithm”.

These “Mop up” algorithms are designed to merge together shower-like PFOs, or clusters, based on overlaps between cones fitted performed to the clusters, sliding fits, and vertex proximity, as was discussed in Section 5.1. The new algorithm recursively repeats the algorithms that were previously run until no more merges are made, or a maximum number of iterations is reached. The 3D space point creation is additionally rerun using the updated PFO characterisation to determine whether to use the algorithms targeted at tracks or showers. This results in more faithful recreations of previously misclassified showers in 3D.

The segmentation of true particles is demonstrated in Figure 5.13 which shows the number of reconstructed particles for each true particle. Electrons exhibit the worst segmentation with only 56% of particles having a single reconstructed cluster and some having as many as 7 using the standard configuration. Implementing the recursive shower growing reduces the severity of this problem with 67% of particles now having a single reconstructed PFO. The segmentation of photons is less effected as these are less often misclassified originally, as discussed in Section 5.3. Track-like particles remain largely unaffected by segmentation with no visible change when applying this new algorithm.

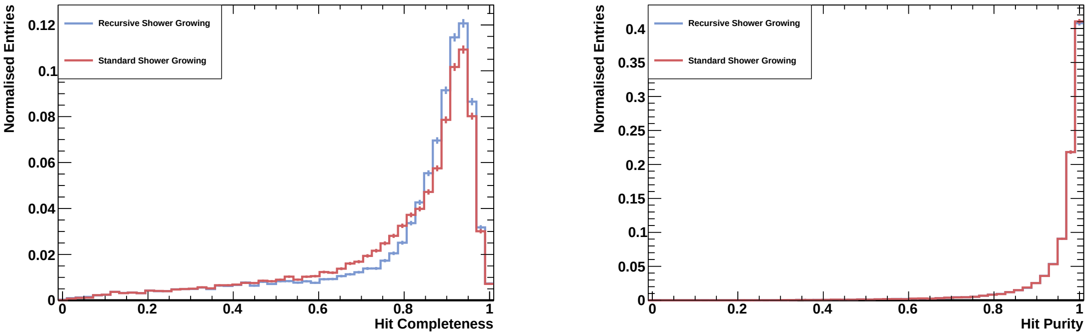


(a) Electrons (b) Photons

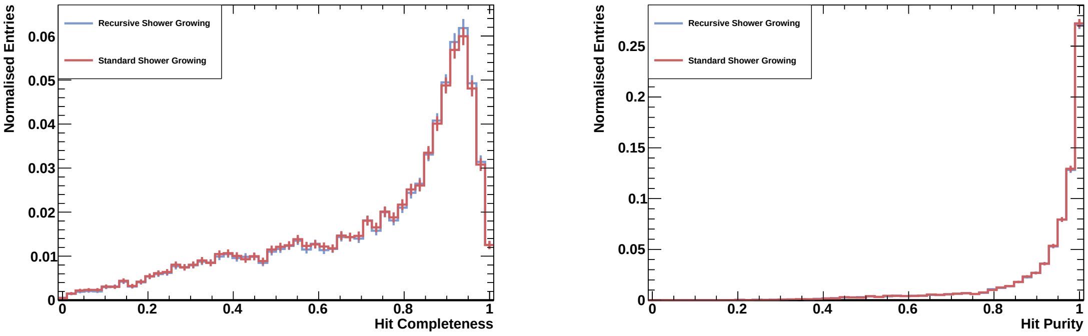


(c) Pions (d) Protons

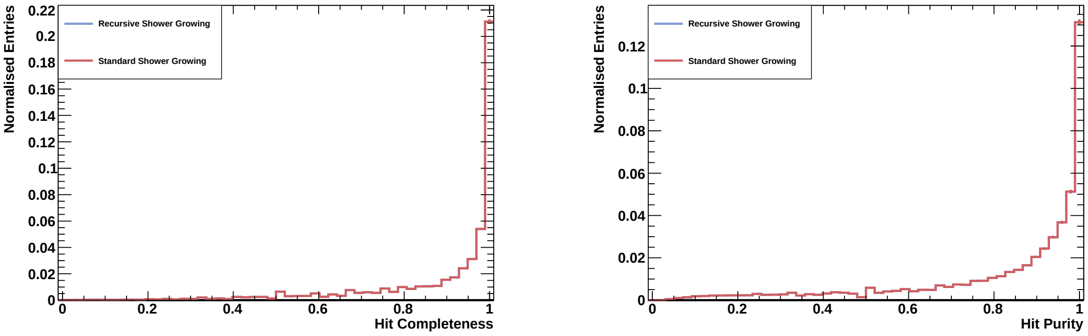
Figure 5.13: Comparison of the number of reconstructed PFParticles per true particle with/without recursive shower growing for various particles.



(a) Electron Completeness (b) Electron Purity



(c) Photon Completeness (d) Photon Purity



(e) Proton Completeness (f) Proton Purity

Figure 5.14: Comparison of the completeness (left) and purity (right) with/without recursive shower growing for different particles.

Merging these clusters together impacts both the completeness and purity of the PFOs that match back to a given true particle as shown in Figure 5.14. There is a significant increase in the completeness of electrons when re-running these mop-up algorithms, increasing the fraction of events with a completeness above 80% from 65% to 71%, with no significant change to purity. There is no appreciable change to either the photons or protons, showing that this algorithm is succeeding in targeting the misclassified high energy electrons.

The increase in completeness has a significant impact on the energy resolution achievable for electrons, this is discussed in Chapter 6. The impact can most visibly be seen for higher energy showers, as expected as these were the most commonly misclassified prior to the BDT. Additionally, the reduced shower segmentation leaves many fewer shower candidates per event which improves the event selection, as will be discussed in Chapter 7.

5.5 Concluding Remarks

The Pandora pattern recognition package has been described, outlining the overall process and some key algorithms. The unambiguous cosmic background rejection has been shown to be performing well, rejecting in excess of 90% of cosmic rays whilst maintaining over 90% of neutrino interactions. MVAs have been applied to guide decision making at key points in the pattern recognition process. This has resulted in improved vertexing, with a 40% reduction in the average displacement of reconstructed vertex for ν_e interactions. Similarly, the PFO characterisation MVA has improved the identifications of high energy electrons as shower-like and protons as track-like by 14% and 25% respectively. The neutrino slice identification MVA provides an additional rejection of the remaining cosmic rays, rejecting over a third with a sub-percent loss to neutrino interactions. Finally, the recursive shower growing algorithm has been developed to provide a 11% increase in the number electrons reconstructed as a single PFO, reducing the shower segmentation issues. This results in an additional 6% of electrons having a completeness in excess of 80% whilst maintaining equivalent purity, without degradation of performance for other particles. These performance improvements to the pattern recognition in turn correspond to improved reconstruction of high level quantities, particularly for showers, as will be discussed in the following chapter.

Chapter 6

High Level Reconstruction

After the Pattern recognition, discussed in Chapter 5, has been run the next stage in the reconstruction is to characterise the Particle Flow Objects (PFOs) produced, with separate pathways for track-like and shower-like PFOs. Specifically, the aim is to extract topological and calorimetric variables that can be used to identify particles and measure their kinematics. This is a vital stage for any physics measurement as it underpins any downstream Particle IDentification (PID) that may be used to select signal and reject backgrounds. Similarly, the ability to measure the properties of particles produced is crucial to measuring the kinematics of a neutrino interaction, important for cross section and oscillation measurements alike. The modules discussed here exist within the LArSoft framework shared across Liquid Argon Time Projection Chamber (LAr TPC) experiments[5].

Firstly in Section 6.1, an overview is given of the track fitting and calorimetry that is used in LAr TPCs. This including calculating the direction and energy loss profile for the track, formed from the dE/dx and residual range. An introduction to the key aims of the shower characterisation is presented in Section 6.2, explaining the parameters that are to be extracted. A new shower characterisation framework, PandoraModularShowerCreation, that was developed in the Short-Baseline Near Detector (SBND) is then discussed in Section 6.2, with the details on how each characteristic is extracted. Finally, the impact of the work presented in Chapter 5 on the shower characterisation is discussed in Section 6.3.

Overall, the improvements discussed in this chapter shows an increase in the quality of all reconstruction metrics, the impact of which will be discussed in subsequent chapters. In particular, the improved dE/dx calculations leads to increased separation power between electrons and photons whilst the improved energy reconstruction improves the prospect for differential cross section measurements as will be presented in Chapter 8.

6.1 Track Reconstruction

Once the pattern recognition, discussed in Chapter 5, has produced a track-like PFO and a corresponding series of space points the track fitting is run to reconstruct the trajectory. The resultant track that is produced consists of a series of smoothed trajectory points, each corresponding to a given space point with a direction and refined position. This is a vital part of the reconstruction as the track direction can be used to extrapolate tracks to the Cosmic Ray Tagger (CRT), measure the kinematics of neutrino interactions, and extract calorimetric information from the track.

This is undertaken by PandoraTrack which performs a sliding linear fit to the space points provided, as depicted in Figure 6.1[170]. The space points shown as the dots, are first ordered along an initial axis of the track, calculated via a Principal Component Analysis (PCA), depicted as the purple line[169]. For each space point a linear fit is performed using 10 points both upstream and downstream, for the point indicated by the orange star the points in red are used to fit the line shown in blue. This fit is used to extract the refined position and direction of each trajectory point.

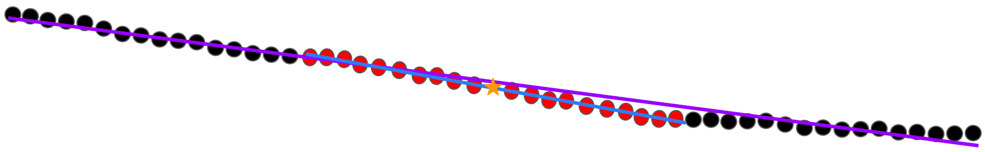


Figure 6.1: Diagram of the track fitting performed by PandoraTrack. The orange star indicates the point currently being fitted with the red and black dots representing the points that are included and excluded from the sliding linear fit respectively. The purple line depicts the initial axis of the shower, found via a PCA, and the blue line the result of the sliding linear fit.

Once the track has been reconstructed calorimetric reconstruction is run to calculate the energy loss per unit length, dE/dx, for each point in the track[171]. This information can then be used in both PID and to calculate the particle kinematics, covered later in Chapter 7.

Using the direction of a given trajectory point, the effective pitch (dx) can be calculated as the distance that a particle would need to travel in order to traverse a single wire in 3D. A 2D depiction of this is shown in Figure 6.2 where trajectory points are depicted as black points and the red lines between adjacent points represents the dx. For a particle travelling perpendicularly to the wire direction this corresponds simply to the wire pitch, the spacing between the vertical lines representing the wires, and when a particle is parallel it is infinite. The charge deposited in a hit is

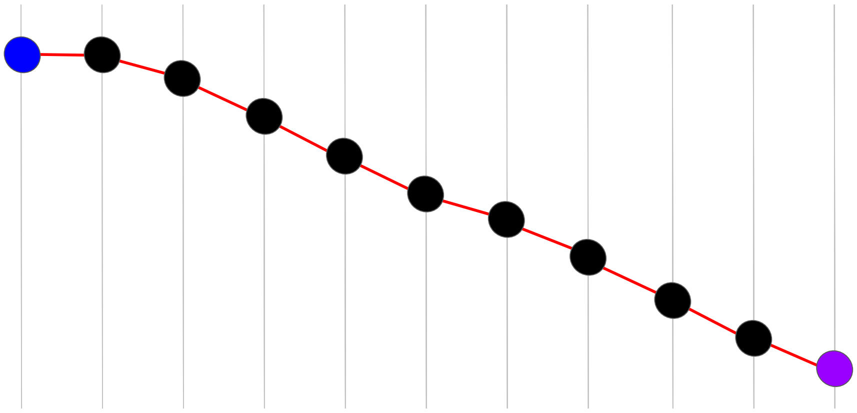


Figure 6.2: Diagram of the track pitch calculation between track trajectory points, depicted as black points. The vertical gray lines represent the wires and the red lines the dx between trajectory points. The residual range is defined as the path length, sum of the individual dx, between a given point and the end of the track.

calculated by converting the ADC to charge, the dQ, using calorimetry constants and correcting for attenuation from impurities, as discussed in Chapter 3. The dQ/dx can then be converted into the dE/dx, accounting for recombination. This is performed using the modified box model with parameters measured in the Argon Neutrino Teststand (ArgoNeuT), as discussed in Chapter 3[140].

The residual range of each trajectory point is also found; defined as the path length, the sum of the dx, between a given trajectory point to the end of the track, denoted by the purple point. This is vital for understanding the distribution of the dE/dx along a track, for example searching for Bragg peaks[172]. The total range of the track, the residual range at the start of the track, can also be used to estimate the energy of the track via a look-up table, assuming the particle stops via ionisation[173]. The relationship between range and energy is dependent on the particle type, so the energy is calculated under several hypotheses and the appropriate one chosen once PID has been performed.

Once this procedure has been performed for each trajectory point in the track the dE/dx can then be used to estimate the kinematics of the particle. Summing the dE/dx multiplied by the pitch for all points in a track yields the total calorimetric energy deposited by the particle in the detector. The energy can also be estimated from the Multiple Coulomb Scattering (MCS) of the track, most commonly used for particles that exit the detector where the range-based and calorimetry methods performs poorly[173]. Similarly to the range based method, the MCS fitting is performed under multiple particle type hypotheses, allowing analysers to determine the appropriate hypothesis once PID has been performed. In addition to measuring particle kinematics, these complimentary methods of calculating energy can be used to check for consistency in the PID, later covered in Chapter 7.

6.2 Shower Reconstruction

As an intrinsically stochastic process, electromagnetic shower development contains a high degree of randomness which makes reliably extracting metrics to describe the shower a particular challenge. Extraction of such characteristics from a shower is vital for the ability to identify and measure these showers nevertheless. In particular the key characteristics used for these purposes in a LAr TPC are itemised below and depicted in Figure 6.3:

- **Start Position:** The first energy deposition by the showering particle, vital for reconstructing the hierarchy of an event, represented in Figure 6.3 by the blue star. Identifying conversion gaps between the shower start and parent interaction vertex is a key identifier for photon showers.
- **Direction:** The direction of the underlying showering particle represented by the purple line, used for reconstructing event topologies including π^0 invariant mass reconstruction and differential cross section measurements.
- **Initial Track Hits:** The energy depositions of the showering particle before the development of the electromagnetic cascade, the blue points in Figure 6.3, used to infer information on the underlying particle.
- **dE/dx:** The charge deposited per unit length by the initial track of the shower which can be used to differentiate between electron and pair producing photons.
- **Energy:** The total energy deposited by the cascade which is representative of the energy of the true showering particle, provided the particle is contained.
- **Length and Opening Angle:** Parameters describing the dimensions of a cone, indicated by the shaded grey region, that encapsulates the shower.

In order to extract these characteristics in SBND, a new framework was developed in conjunction with Dominic Barker: PandoraModularShowerCreation[6][174]. This framework breaks down the calculation of each of these characteristics into individual, interchangeable stages calculated by Art tools. Isolating the algorithms in this way enables rapid analysis and development of tools as they are interchangeable via only a configuration change.

Many of these stages of shower characterisation are intrinsically linked; for example the start position may be used to find the initial track hits which, in turn, are used to calculate the dE/dx. This necessitates the ability to pass information between the various tools, enabled via the ShowerElementHolder within the new PandoraModularShower framework. This holder contains the characteristics required

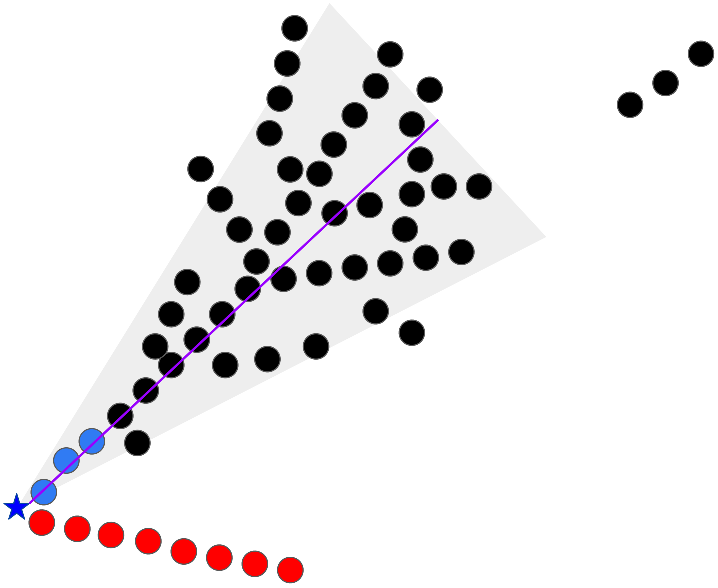


Figure 6.3: Diagram of the key characteristics reconstructed in a shower. An charged pion track, the red points, and an electron shower, in black, are simulated from a common vertex indicated by the blue star. The blue points represent the initial track hits before the shower begins to develop, these are used when calculating the dE/dx of the shower. The direction of the shower is represented by the purple line with the grey cone depicting the length and opening angle.

to create a shower and additionally enables each tool to both read and write additional information to be used by subsequent tools. This modularity allows for the interplay between the various characteristics to be analysed, for example changing only the tool that calculates the start position and evaluating the impact that this has on the calculation of the dE/dx.

In order to investigate the performance of the shower reconstruction a pair of vertex-like samples were created. These samples consisted of either an electron or a photon originating from a shared vertex with a charged pion as depicted in Figure 6.3. The pion was used to create a more neutrino-like scenario for the pattern recognition which improved performance, particularly in vertexing, compared to single particle samples. Both the electron and photon were generated using the energy distribution expected from BNB electrons to be able to compare reconstruction quantities without biases from differing underlying kinematic distributions.

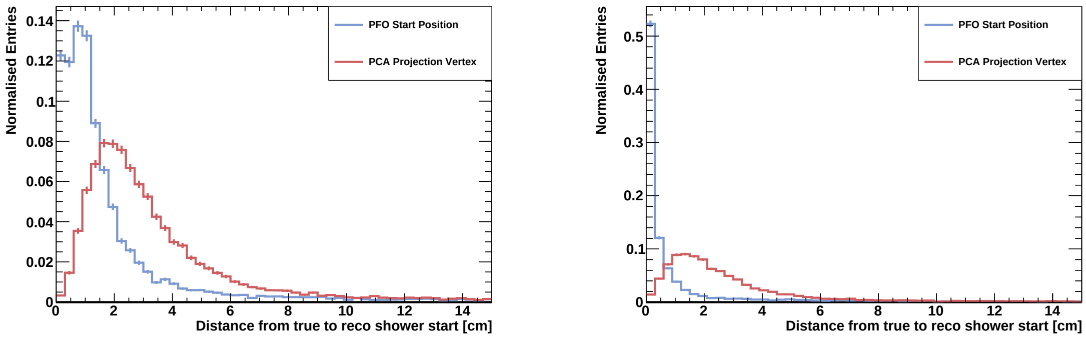
The following subsections present a comparison of the available tools for each of the characteristics previously presented using these samples. The most performant tool is identified in each of the subsections which is then used to underpin the calculations in the subsequent subsections.

6.2.1 Start Position

The start position is not only vital for identifying conversion gaps for photons but additionally underpins the initial track hit finding and subsequent dE/dx calculation. Two methods exist to find this start position within PandoraModularShower: using the Pandora vertex with the PFOStartPosition tool and projecting this vertex onto the shower axis with the PCAPropagationStartPosition tool.

The PFO start position method simply takes the vertex that is assigned to the PFO by Pandora during the pattern recognition, found by finding the space point that is closest to the parent particle. Primary showers, which have no parent particle, use the space point closest to the neutrino vertex, discussed in Chapter 5.3.2. The PCA propagation method relies on reconstructing the axis of the shower using a PCA, which is additionally used to calculate the shower direction as will be discussed in Section 6.2.2. The PFO vertex is then projected onto this axis to calculate the shower start position. This method theoretically reduces the dependence on correctly assigning hits near the vertex to the correct PFO. However, errors when calculating the shower axis can lead to worse performance with this method. This method can also lead to a start position displaced from any hits assigned to the shower causing issues when trying to find the initial track hits within some distance of the start, as will be discussed in Section 6.2.3.

A comparison of the performance of these methods is shown in Figure 6.4 which shows the distance between the true and reconstructed shower start positions. The fraction of showers with a reconstructed start within 1 cm of the true start is 42% (73%) for the PFO tool and 14% (29%) for the PCA tool, for electrons (photons). This superior performance for the PFO tool demonstrates that projecting the vertex



(a) Electron Showers (b) Photon Showers

Figure 6.4: Comparing distance between the true and reconstructed shower start position across different methods for electron showers, left, and photon showers, right.

onto the axis of the shower introduces additional offsets from the true shower start. The start position being reconstructed more accurately for photons than for electrons is expected as the photons travel some distance from the neutrino interaction vertex before showering. As such, there is less hadronic activity around the shower start and thus it is easier to assign hits to the correct cluster.

Overall the PFOStartPosition demonstrates the best performance and as such is used to calculate the start position for the remainder of the chapter, underpinning the calculation of the remaining metrics.

6.2.2 Direction

Correctly reconstructing the direction of showers is vital for accurately reconstructing the topology of neutrino interactions, from making differential cross section measurements against the leptonic angle to reconstructing π^0 mass peaks. There are two distinct approaches to calculating the direction of the shower employed by PandoraModularShower: calculating the direction from the entire shower using the PCADirection tool and calculating the direction from only the hits at the start of the shower using the TrackDirection tool.

The former method using the entire shower relies on the conservation of momentum as the shower develops, assuming the energy deposits are evenly deposited around the central axis representing the true direction. This method minimises the impact of misclustering and statistical fluctuations at the start of shower. Any missing energy or impurities can skew the direction of the shower however so incomplete or uncontained showers are likely to have poor results with this tool. Additionally, this does not account for any scattering of the initial particle before the shower starts to develop or uncontained energy depositions.

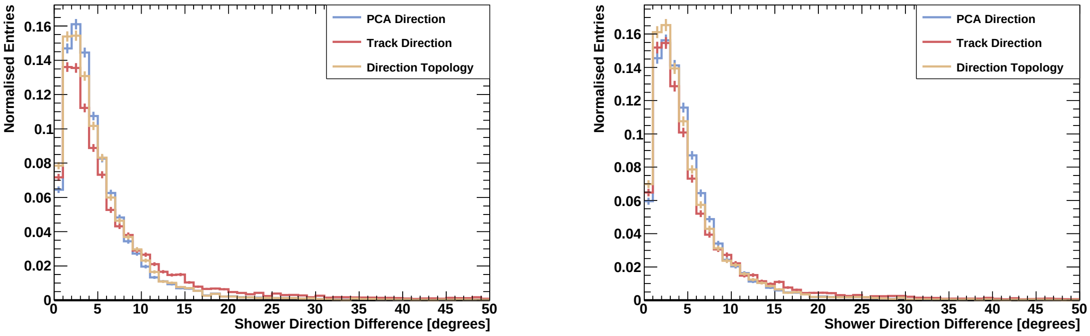
A PCA is used when calculating the direction of the shower using all of the hits, the details of the PCA calculation are discussed in Chapter 5.3.3[169]. The axis of the shower is defined as the primary axis of the PCA, the axis along which the spread of points is maximised. However, the PCA does not indicate whether the shower develops forwards or backwards along this axis. The direction of the shower is thus defined as the direction along the primary axis that points from the start to the centre of the shower.

In contrast, the track direction tool uses the hits from the start of the shower and attempts to reconstruct the primary particle before the shower begins to develop, attempting to reconstruct the true initial direction of the particle. This can reduce the impact from missing energy deposits from the shower from exiting photons, below threshold depositions or clustering inefficiencies. Additionally, using the direction at the start of the shower can reduce the impact of the electron scattering before the shower develops. However, this method heavily relies on accurate clustering around the vertex and any erroneously included, or excluded, hits can skew this measurement heavily.

In order to calculate the direction at the start of the shower, the hits at the start of the shower need to be identified, which will be discussed in Section 6.2.3. In order to identify these hits, the direction from the whole shower is initially used to create an approximate direction. Once these hits have been identified, a sliding linear fit is performed, using a method within Pandora used for track-like particles as discussed in Section 6.1. The direction of the shower can then be extracted from the fitted trajectory points of the track, in this case by taking the direction of the first point. This method is heavily reliant on the ability of the pattern recognition and initial track hit identification to correctly identify the hits at the start of the shower and as such the performance may improve as these evolve.

The two methods can also combined together using the DirectionTopologyDecision tool. This attempts to combine the accuracy of the PCA and the precision of the track direction by using the track direction when there is good agreement between the two methods, defined as being within 17° of each other, and otherwise using the PCA direction. This, in theory, enables a refinement of the direction when the track direction tool performs well but rejects egregious mistakes that occur due to poor clustering around the vertex.

A comparison of the three methods can be seen in Figure 6.5 which shows the difference between the reconstructed and true shower direction for both electrons and photons. This track direction method has a much larger spread, with 21% (18%) of electrons (photons) being more than 10° out from the true direction. In comparison, the PCA method is more consistent, with only 12% of both electrons and photons



(a) Electron Showers

(b) Photon Showers

Figure 6.5: Comparing the angle between the true and reconstructed shower start position across different methods for electron showers, left, and photon showers, right.

meeting the same criteria. Utilising the direction topology decision tool yields a boost to the number of events within 5° of the true direction, 62% (65%), relative to the PCA, 62% (62%), whilst maintaining the same fraction outside of 10°.

Overall, the DirectionTopologyDecision succeeds in its aim of combining the PCA and track direction methods to achieve the best performance. As such, this is the method used to calculate the direction for the remainder of the chapter.

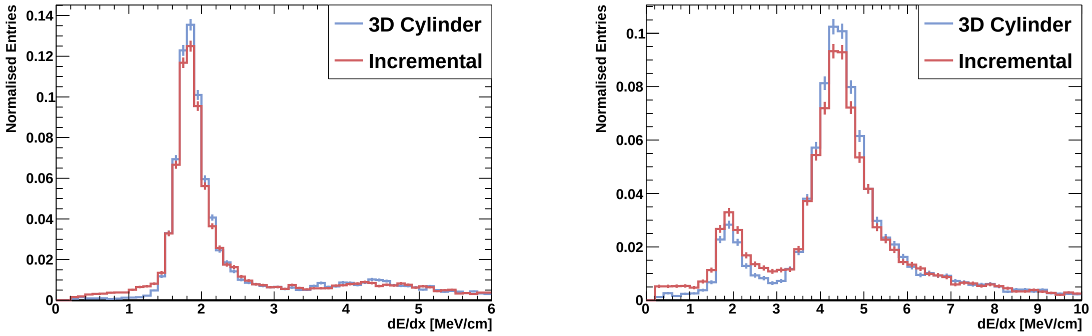
6.2.3 Initial Track Hits

The hits before the primary particle begins to shower contain information pertaining to the true underlying primary particle that is lost once the shower begins to cascade. These initial track hits are vital to being able to differentiate between electron and photon induced showers using the dE/dx, which will be presented in Section 6.2.4, in addition to being used for refining the direction as previously discussed.

The simplest way to define the initial track hits is to define a cylinder around the start of the shower, using the start position and direction previously calculated, using the 3DCylinderTrackHitFinder tool. This method relies heavily on the clustering around the start of the shower to have assigned the correct hits to the cluster, as all hits within the cylinder are considered.

A more advanced algorithm also exists in the form of the IncrementalTrackHitFinder. This algorithm initially selects a set of N seed hits closest to the vertex, performs a PCA and calculates the residual distances from the hits to the primary axis. If the average residual is above a given tolerance the hit with the highest residual is removed, this process is then repeated until the average residual drops below the tolerance. Hits further from the vertex are then tested to see if the average residual would increase or decrease if they were added, hits that satisfy the latter are added. This process is repeated until three successive hits are rejected. This process can help to remove spurious hits from the initial track and prevent the track being extended into the bulk of the shower.

The truth definition for the initial track hits is somewhat ill-defined as it is extremely challenging to robustly identify the point at which to classify the shower as developed. As such, the metric used to assess the performance of the initial track hit finding is the dE/dx, the calculation of which is discussed in the next section and shown in Figure 6.6. This shows the expected single Minimum Ionising Particle (MIP) peak at, 1.8 MeV/cm, for electrons and both the single and double MIP peaks for photons[175]. The 3D cylinder provides a more defined shape to the dE/dx compared to the incremental hit finder for electrons, with 71% having a dE/dx below 3 MeV/cm for both tools. Similarly, the 3D cylinder correctly calculates 87% of



(a) Electron Showers

(b) Photon Showers

Figure 6.6: Comparing dE/dx for different track hit finding methods for electron showers, left, and photon showers, right.

photons having a dE/dx above the same 3 MeV/cm cut compared to 81% for the incremental tool.

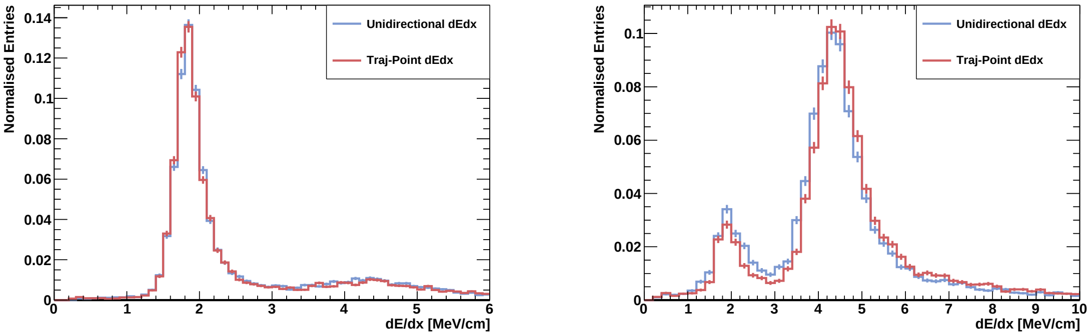
As such, the 3D cylinder provides the most separation power between electrons and photons and was thus the tool used for the following section. The incremental hit finder was tuned to maximise the separation between electrons and photons but upstream reconstruction changes have removed the performance advantage originally had over the 3D cylinder, demonstrating the robustness of the simpler tool.

6.2.4 dE/dx

Analysing the dE/dx of the track stub at the beginning of a shower allows for powerful discrimination between electrons and pair producing photons. The dE/dx of an electron will correspond to that of a single MIP whereas a photon that pair produces an e^+e^- pair will correspond to two MIPs, as discussed in Section 3. Once the initial track hits have been found via the methods discussed in the previous section, here using the 3DCylinderTrackHitFinder, the dE/dx can then be calculated.

This calculation largely follows the method used for finding the dE/dx of tracks discussed in Section 6.1: first calculating the effective pitch (dx) and then converting the charge (dQ/dx) to energy (dE/dx). Although this base method is used in each of the tools there are some subtleties regarding the details of this calculation. The UnidirectionaldEdx tool simply uses the direction of the shower as a whole to calculate the effective pitch, as defined by the algorithms discussed in Section 6.2.2. In contrast, the TrajPointdEdx tool relies on the sliding fit performed to the initial track hits to provide a direction for each hit, allowing more precise calculations of the pitch, and hence dE/dx.

These methods are compared in Figure 6.7 which shows the dE/dx distribution for both electrons and photons. Overall, the TrajPointdE tool produces more defined distributions with both electrons and photons having sharper, better defined peaks. By placing a cut at 3 MeV/cm, the Unidirectional tool accepts 69% of electrons whilst rejecting 83% of photons compared to 71% and 87% respectively for the TrajPoint tool. This increased separation power will directly impact the selection efficiency and background rejection discussed in Chapter 7 where the TrajPointdEdx is used.



(a) Electron Showers

(b) Photon Showers

Figure 6.7: Comparing dE/dx for different calculation methods for electron showers, left, and photon showers, right.

6.2.5 Energy

Due to the sprawling nature of showers, it is not feasible to calculate the pitch for every hit of the shower. Therefore, the energy cannot be calculated by summing the dE/dx as was done for tracks, discussed in section 6.1. Instead, a direct conversion from charge to energy is used, assuming a linear relationship.

First, the charge collected is converted to a number of electrons collected on the wires, using the calorimetry constants previously discussed. Subsequently, the number of electrons collected at the wire is related to the number of electrons deposited by correcting for electron lifetime and recombination using a constant factor of 0.64 derived from Monte Carlo (MC) studies. Finally, this number of electrons is converted to the deposited energy by multiplying by W_{ion} = 23.6 eV: the energy required to ionise an argon atom[140][141].

The hit finding plays a vital role as any energy not included in a hit is not passed to Pandora and is thus excluded from the shower energy calculation. The performance of the hit finding is quantified using the hit energy completeness, defined as:

$$\textbf{Energy Completeness} = \frac{\textbf{Energy in Reco. Hits}}{\textbf{Total Deposited Energy}} \tag{6.1}$$

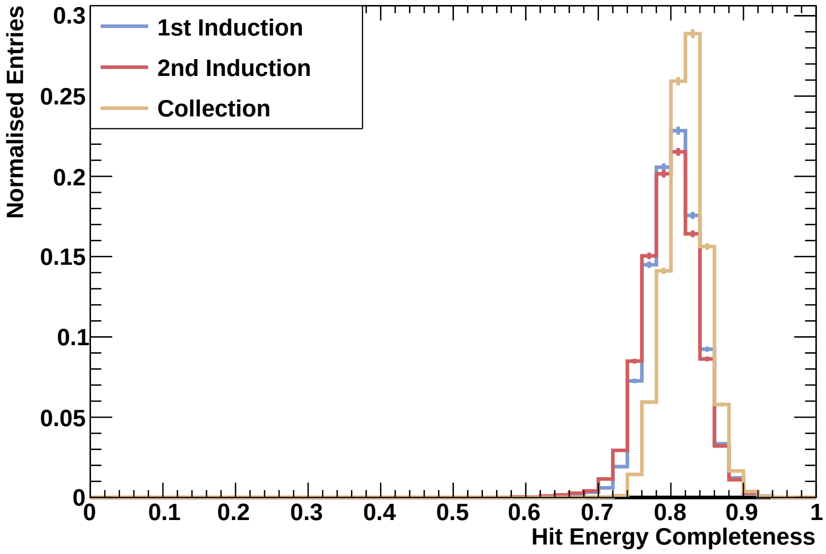


Figure 6.8: Energy completeness of all of the reconstructed hits for simulated electrons shown for each plane.

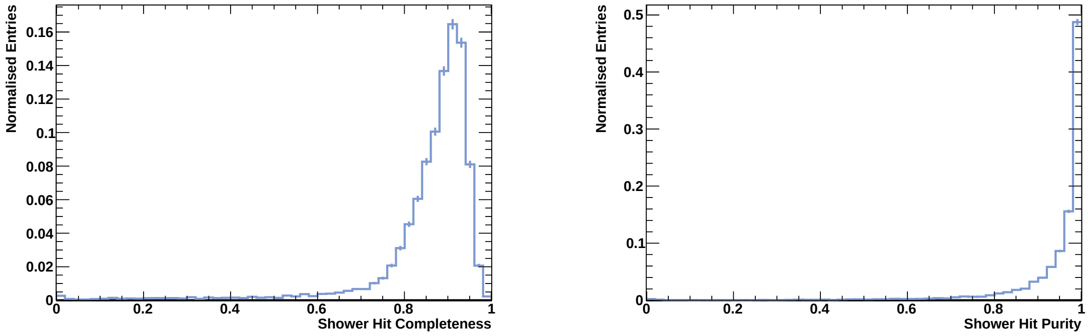
This hit energy completeness for this sample of electrons is shown in Figure 6.8, separated into planes. This shows that the performance is worse in the induction planes relative to the collection plane; this is expected due to the intrinsically worse signal-to-noise ratio in these planes. Overall, the hit finding encapsulates 80% (83%) of the deposited energy by the true simulated particle in the induction (collection) planes. The majority of this missing energy is due to small, isolated energy depositions that are not large enough to pass the hit-finding thresholds. Nevertheless, the broadness of the distribution limits the extent to which this inefficiency can be calibrated out without introducing additional uncertainties.

The pattern recognition described in Chapter 5 then defines which hits are clustered into the shower to be used in the energy calculation. Figure 6.9 shows the purity and completeness for electron showers both in terms of hits, respectively defined as:

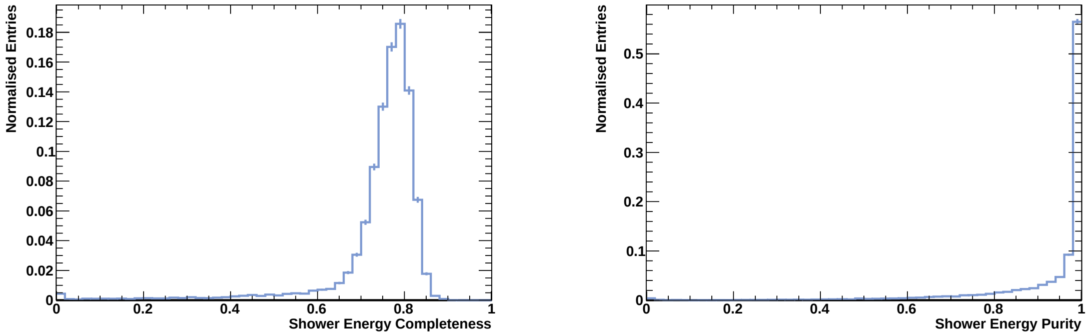
$$\textbf{Hit Purity} = \frac{\textbf{\# True Hits in Reco. Shower}}{\textbf{\# Hits in Reco. Shower}} \tag{6.2}$$

$$\textbf{Hit Completeness} = \frac{\textbf{\# True Hits in Reco. Shower}}{\textbf{\# True Hits}} \tag{6.3}$$

These hit metrics show how correctly the pattern recognition assigns hits to showers. Generally, it was found that the showers are very pure with a hit purity of above 90% in 83% of electron showers. Completeness is, generally, not as high with only 43% of electron showers having a hit completeness above 90%.



(a) Hit Completeness (b) Hit Purity



(c) Energy Completeness (d) Energy Purity

Figure 6.9: Evaluating the performance of the pattern recognition with hit, top, and energy, bottom, metrics. Hit completeness is assessed on all planes and the energy completeness in the best plane only.

The energy metrics shown in Figure 6.9 additionally incorporate the effect of the hit finding, as shown in Figure 6.8, in addition to the hit purity and completeness. This effect can be observed that the energy completeness is less that the hit completeness, as even with 100% hit completeness the upper limit of efficiency is determined by the hit energy completeness. Incorperating both the hit finding and clustering inefficiencies shows that the showers are typically very pure but somewhat incomplete.

The shower energy ratio, E_{Reco}/E_{True}, calculated from these showers is shown in Figure 6.10, with the shape closely resembling that of the distributions shown in Figure 6.9. The distribution is peaked at a shower energy ratio of 0.875, with the low tail and high tails from incomplete and impure showers respectively. There is a slight offset from the peak of the shower energy completeness of 0.775, indicating a slight offset in the calibration. Therefore, a downstream correction is needed to account for the inefficiencies in hit finding and clustering. Profiling the energy ratio against the true shower energy, the energy is overestimated at low energy, where the

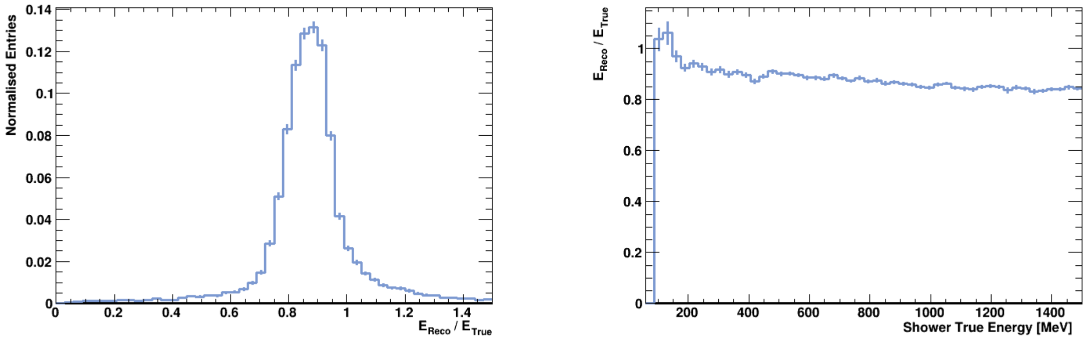


Figure 6.10: Comparing true and reconstructed shower energy for electron showers integrated, left, and profiled, right, over true shower energy.

effect of impurities is largest, but the distribution levels off and becomes mostly flat above 200 MeV.

Overall, it was shown that the energy reconstruction is highly dependent on both the hit finding efficiency and clustering performance. The energy calculation accurately extracts the energy from the provided hits thus it was concluded that the energy resolution is limited by the hit finding and clustering.

6.2.6 Length and Opening Angle

The final quantities calculated by the shower reconstruction are the length and opening angle of the shower. These variables define the topological shape of the shower and can be used to differentiate between true showers and tracks that were misclassified as showers, discussed further in Chapter 7.

The length and opening angle variables are intrinsically linked as the opening angle of a shower is defined as:

$$\theta_{Open} = \mathbf{tan}^{-1}\left(\frac{\textbf{Shower Width}}{\textbf{Shower Length}}\right) \tag{6.4}$$

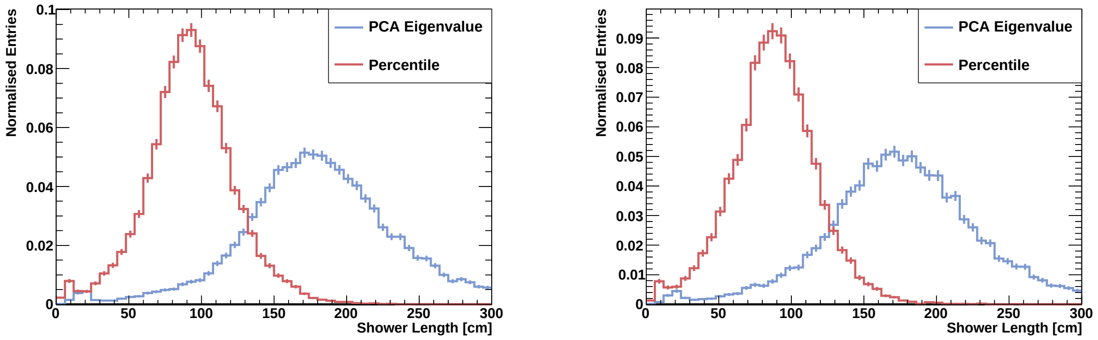
The width and length of the shower are calculated via the same methods using the longitudinal and transverse spread of points around the centre of the shower.

There is some ambiguity surrounding the definition of the length, and width, of a shower however due to the lack of a well defined end point. The length can be calculated by either using the furthest point from the shower as the end or by defining a characteristic length that encapsulates the majority of the shower.

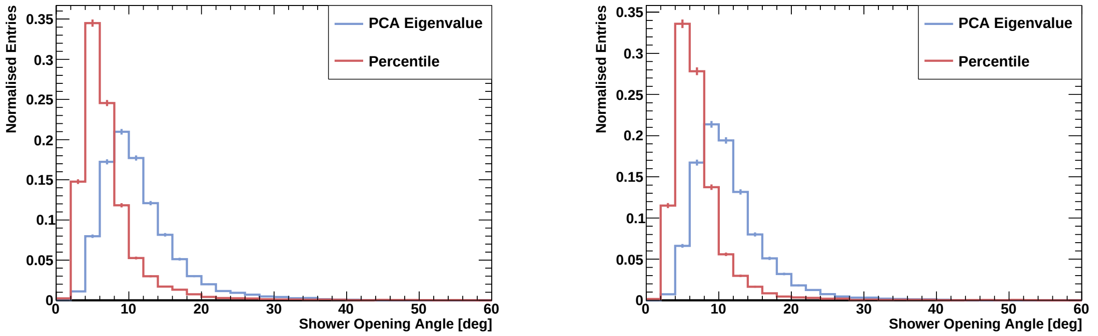
There are several methods of calculating these quantities: calculating the eigenvalues of the PCA using the PCAEigenvalueLength tool or calculating the percentiles of the distribution of the spread using the LengthPercentile tool. The first method uses the PCA calculated in Section 6.2.2 to find the direction of the shower. Alongside the eigenvectors representing the axis of the shower, the eigenvalues computed describe the variance of points around these axes and hence the square root of these describes the standard deviation. In particular, the standard deviation along the primary and secondary axes are used to calculate the length and width of the shower respectively. This method can be susceptible to being skewed by extremely detached energy deposits however, as these points will largely inflate the average distance from the shower centre.

In comparison, the percentile method defines the length and width to which a given percentile, by default the 90th, of all hits lie within from the centre of the shower. This method aims to be robust against outlying energy depositions far from the centre of the shower that may skew methods relying on the average distance.

Due to the previously stated ambiguities, there is no definition of true shower length or opening angle, so the reconstructed quantities cannot be compared to the true as has been done for previous sections. Nevertheless, a comparison of the methods can be seen in Figure 6.11 which shows the distributions of length and



(a) Electron Length (b) Photon Length



(c) Electron Opening Angle (d) Photon Opening Angle

Figure 6.11: Comparing the shower length, top, and opening angle, bottom, between electrons, left, and photons, right, for different calculation methods.

opening angle the two methods. Generally, the two methods exhibit the same shaped distribution with the PCA eigenvalue method being more stretched, as expected due to outlying depositions skewing the mean above the median. The percentile method was ultimately used as it proved more useful in estimating shower containment, as will be presented in the proceeding Chapter.

6.3 Impact of Pattern Recognition

As the shower characterisation builds directly upon the output of the pattern recognition it is sensitive to any upstream changes, such as those implemented in Chapter 5. This section includes a comparison of the performance of the shower reconstruction metrics presented in the preceding section for both the standard Pandora configuration and the latest SBND version including the improvements discussed in Chapter 5. Specifically, this includes the introduction of vertexing and PFO characterisation Multi Variate Analyses (MVAs) and the recursive mop up algorithms. The comparison is performed on the same vertex-like sample of electrons previously introduced and uses the most performant tool for each stage of the characterisation.

Prior to the introduction of the PFO characterisation MVA many high energy electrons were erroneously classified as tracks by Pandora. As the figures presented in this section only consider reconstructed showers, and not tracks, if such a misclassification occurs the PFO that best matches to the true particle will not be considered. Instead, a smaller PFO that was correctly classified would often be considered the best match or sometimes no match would be found. This lead to having reconstructed showers that were detached and incomplete relative to the underlying showering particle. The introduction of the PFO characterisation MVA largely eliminated this misclassification and thus the fraction of detached and incomplete particles was dramatically reduced.

The impact on the topological shower variables, namely the start distance and direction difference, is shown in Figure 6.12. This shows the difference between the true and reconstructed quantities when using both the standard Pandora and SBND tuned Pandora. Overall, both of these metrics show large improvements when using the SBND tuned Pandora. The fraction of showers with a reconstructed start position within 1 cm of the true shower start almost doubles from 23% to 42% with the improvements to pattern recognition, largely due to the impact of the vertex finding MVA. Similar increases are seen in the direction calculation with the fraction of showers with a direction reconstructed within 5° of the true direction increasing from 45% to 61%. The direction calculation particularly benefits from the increased

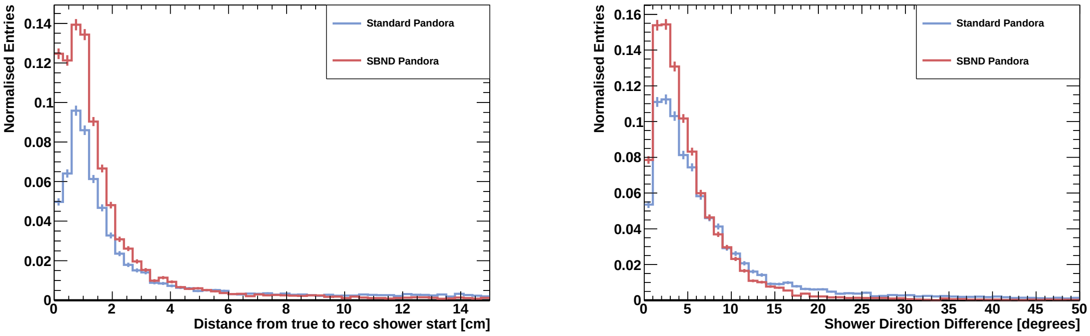


Figure 6.12: Comparing the difference between the true and reconstructed start position, left, and direction, right, for electrons with the standard Pandora and SBND tuned Pandora.

completeness and improved 3D hit creation resultant from the recursive shower mop up algorithms.

Figure 6.13 similarly shows the relative performance of the calorimetric variables with and without the pattern recognition improvements. The dE/dx distribution shown on the left demonstrates a sharper and more defined shape with the improvements. Applying a cut at 3 MeV/cm the improvements increase the fraction of electrons selected from 64% to 70% whilst simultaneously raising the fraction of photons rejected from 74% to 86%. This metric is improved particularly from the improved vertexing allowing an improved selection of initial track hits. Improvements to the shower completeness, both as a result of the improved vertexing and recursive mop up algorithms, result in a dramatically improved energy resolution, as shown on the right. The effect of the PFO characterisation MVA has the largest impact on the energy reconstruction due to the previously discussed shower selection. Nevertheless,

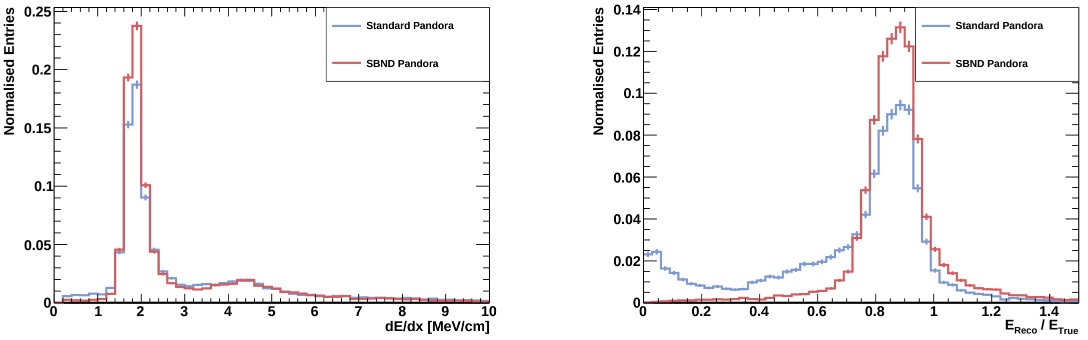


Figure 6.13: Comparing the shower dE/dx, left, and energy ratio, right, for electrons with the standard Pandora and SBND tuned Pandora.

the improved completeness from the recursive mop up algorithm also dramatically improves the energy resolution for these showers.

The performance of the pattern recognition underpins the performance of the downstream characterisation. The improvements to the downstream pattern recognition translate to dramatically improved shower reconstruction across all areas, enabling better selections and measurements of ν_e interactions as discussed in Chapters 7 and 8.

6.4 Concluding Remarks

An overview of the high level reconstruction chain in SBND has been presented, beginning with an overview of the track reconstruction in Section 6.1. The new PandoraModularShowerCreation framework that was developed in SBND was then presented, demonstrating the various tools available for extracting key characteristics of showers. Finally, the impact of the pattern recognition improvements introduced in Chapter 5 are demonstrated to provide a significant uplift in performance across all of the high level shower reconstruction metrics. The presented reconstruction underpins the selection that will be presented in the subsequent chapter, for both PID and kinematic measurements alike.

Chapter 7

Selecting ν_e and $\overline{\nu}_e$ CC Interactions

In order to make any physics measurements in the Short-Baseline Near Detector (SBND), a selection must first be performed to reject backgrounds and ensure the quality of the selected signal. The work presented here aims to select ν_e and $\overline{\nu}_e$ Charged Current (CC) interactions, building upon the reconstruction improvements discussed in Chapters 5 and 6, towards the ultimate goal of a ν_e CC inclusive cross section measurement. This includes reconstructing an electron candidate so the interaction kinematics can be measured and rejecting both cosmic and neutrino induced backgrounds. Cuts were typically optimised to maximise the product of efficiency and purity of the sample, with the exceptions of cuts placed to ensure signal quality rather than to reject backgrounds.

First, the aims of the selection are outlined in Section 7.1 which defines the signal considered in the selection presented here. An overview of the reconstruction is additionally presented which analyses both efficiency losses and the number of potential backgrounds. Next, Section 7.2 contains an overview of the preselection that is applied and a review of the remaining signal and backgrounds. Cosmic removal is then presented in Section 7.3, utilising each of the Time Projection Chamber (TPC), Photon Detection System (PDS) and Cosmic Ray Tagger (CRT) subsystems. The identification of muons, indicative of ν_μ CC backgrounds, is subsequently discussed in Section 7.4. The selection of the electron showers produced in the signal interactions is presented in Section 7.5. Finally, the combined results of all of the cuts is presented in section 7.6 alongside an analysis of the remaining backgrounds.

Overall, the selection presented is able to reject the vast majority of backgrounds whilst selecting many of the signal ν_e and $\overline{\nu}_e$ CC interactions. The selected sample is then used as an input into the cross section measurement discussed in Chapter 8.

7.1 Signal Definition

A selection first begins with defining the signal to be selected, namely ν_e and $\overline{\nu}_e$ CC interactions that occur within the Fiducial Volume (FV) of the detector, as will be defined in Section 7.2.2. For brevity, the combined ν_e and $\overline{\nu}_e$ will be simply referred to as ν_e. Additionally, it is required that the primary electron produced has an energy, E_e^{True}, above 200 MeV to ensure that it is likely to be well reconstructed, motivated by studies discussed in Chapter 5. Any interactions that fail the FV cuts are labelled “Other ν” alongside ν_e CC interactions that fail the minimum lepton energy and which are considered a background. No requirements are placed on the hadrons produced in the interactions as this is a inclusive selection.

This signal definition yields a total of 20,039 signal interactions, compared to a background of 2.91 and 1.14 million ν_μ CC and Neutral Current (NC) interactions respectively, assuming an exposure of 6.6×10^{20} Protons On Target (POT). An additional 9.63 million dirt interactions, neutrinos that interact outside of the FV, are considered as backgrounds, although many of these will deposit no energy in the detector.

The selection relies on the “events” defined by the triggering of the detector, as discussed in Chapter 3, in addition to the “slices” that are created by Pandora, as discussed in Chapter 5. Each slice attempts to encapsulate all of the energy depositions in the TPC from a single origin: either a neutrino interaction or a primary cosmic ray. Some of the cuts presented reject entire events as the information is not associated with a single slice, e.g. cutting on the CRT information which is not connected to a specific slice. The remaining slices are then accepted or rejected based on a series of cuts placed on the reconstructed information within the slice or by matching the slice to activity in one of the detector subsystems, e.g. the PDS.

Due to reconstruction inefficiencies some neutrino interactions have no corresponding slice, this is most prevalent for dirt and NC interactions which commonly deposit little energy in the detector. Similarly, some neutrino interactions will be split into multiple slices. To ensure that these split neutrino interactions are not double counted, any slice with a completeness below 50% is labelled as an “Other ν” background, which further places a requirement for the slice to be well reconstructed.

Therefore, there are a total of 18,860 reconstructed slices corresponding to the signal ν_e CC interactions, an efficiency loss of 5.9%. Similarly, ν_μ CC interactions are reduced by 5.9% to 2.74 million. However, the dirt and NC interactions are reduced by 49.9% and 66.5% respectively, yielding 3.2 million and 570,072 slices respectively. An additional 350,019 slices fail the completeness requirement and thus are classified as split neutrinos. Finally, there are 62 million cosmic induced slices, vastly outnumbering all of the neutrino interactions, with contributions from both out-of-time cosmics which overlay the neutrino interactions and in-time cosmics which triggered the detector. This low intrinsic signal to background ratio resulting from both the low ν_e fraction of the Booster Neutrino Beam (BNB) and the high rate of cosmic rays associated with a surface level detector present a particular challenge for the reconstruction and selection.

7.2 Preselection

The selection begins with a preselection, designed to remove any candidates for which the later selection and analysis stages are not possible, namely slices without a suitable candidate shower. An electron candidate is required to enable the measurement the lepton kinematics of the interaction by taking the largest, highest energy, shower in the slice. Removing these slices not only ensures the suitability of the surviving slices but also rejects a large number of potential backgrounds, allowing the later stages of the selection to be more targeted at removing the remaining backgrounds.

The preselection begins with rejecting slices that fail Pandora's unambiguous cosmic removal; as these slices are reconstructed as cosmic rays and not neutrino interactions they are missing the required data products, such as a reconstructed interaction vertex. Requiring that the reconstructed vertex lies within a FV rejects both cosmic and dirt backgrounds whilst also maintaining reconstruction quality for showers. Finally, it is required that there is a suitable electron candidate shower: defined as a reconstructed shower with an energy, E_e^{Reco}, above 200 MeV, the same cut applied to E_e^{True} in the signal definition, that is fully contained within the detector. Without such a shower it is failing to meet the signal definition and thus is impossible to measure the kinematics of the ν_e CC interactions.

7.2.1 Pandora Unambiguous Cosmic Removal

Being a surface detector, the vast majority of the reconstructed slices are due to cosmic rays; with a total of 62 million cosmic slices expected for an exposure of 6.6×10^{20} POT, compared to 6 million neutrino interactions. In order to initially reduce this all slices that are labelled as "unambiguously cosmic" by Pandora are rejected, this tagging procedure is outlined in Chapter 5. This removal targets slices which are topologically impossible to be neutrinos, such as through going tracks which both start and end at the detector walls or slices that lie outside of the main drift window and thus cannot have entered the TPC during the beam spill window.

This cut was found to successfully remove 83% of the cosmic background slices whilst removing only 0.6% of the ν_e CC signal, with similar reductions the background neutrino interactions. This additionally ensures that all remaining slices are reconstructed as neutrinos and not cosmic rays, guaranteeing that each slice will have a reconstructed vertex and that the dedicated shower growing algorithms have been run as mandated by the upcoming cuts.

7.2.2 Fiducial Volume Cut

A Fiducial Volume (FV) cut is subsequently applied to remove any slices where the interaction vertex lies within a given distance of the detector walls. This cut is primarily designed to remove cosmic rays and dirt neutrino interactions that enter the detector through one of the walls. Additionally, this cut ensures that the electron showers produced in the signal interactions deposit the majority of its energy in the TPC, as will be further discussed in Section 7.2.3. This is vital as the only, current, method of calculating the electron energy is to sum all of the energy depositions within the detector and thus any uncontained showers will have poor energy resolution, as will be presented in the subsequent section.

The distribution of reconstructed neutrino vertices of the slices remaining after the unambiguous cosmic removal is shown in Figure 7.1, with the edges of the detector indicated by the dashed lines. Due to the preponderance of cosmic backgrounds at

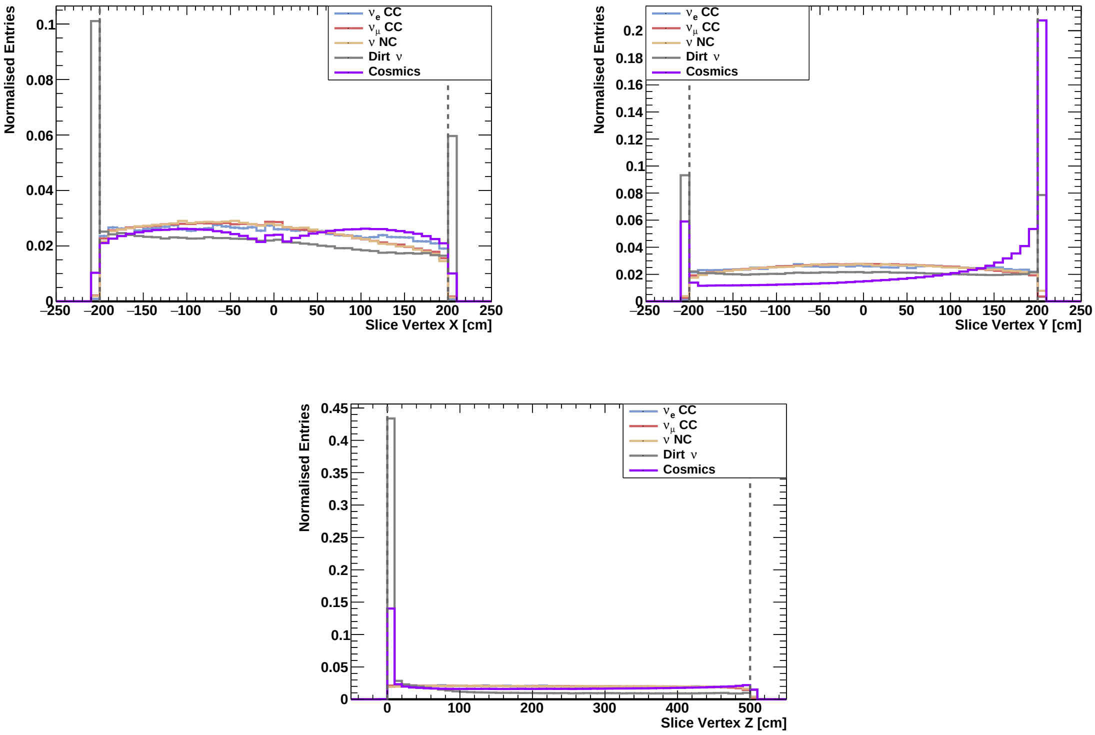


Figure 7.1: Reconstructed neutrino vertex position for fiducial volume cut, the edges of the detector are shown by the dashed lines.

this early stage of the selection, each of the distributions has been normalised so the shape of each distribution can be clearly seen and compared. Cosmic ray slices predominantly enter through the top face of the detector, high Y, with a significant number having a reconstructed vertex at either the bottom of the detector, low Y, or the upstream face of the detector, low Z. The vertices at the upstream and bottom faces of the detector is largely due to residual through-going cosmic rays where Pandora has placed the vertex at the upstream end of the slice. Similarly, the dirt interactions typically enter the TPC via the faces of the detector, particularly the upstream face, as expected.

The FV was chosen by finding the overlap of the normalised dirt background distribution and the ν_e CC signal distribution and is shown in Table 7.1. In the beam direction, Z, a 20 cm cut is applied to the upstream face to remove dirt backgrounds and a 30 cm cut at the downstream face to ensure containment of low energy hadronic activity. Additional 20 cm cuts are placed from the top, bottom, and sides of the detector to remove the remaining cosmic and dirt backgrounds. Finally, a 5 cm buffer is placed around the cathode, located at X = 0, to ensure that the reconstruction around the vertex, i.e. the 3 cm initial track used to calculate the shower dE/dx, is unhindered by the dead region.

Coordinate	Minimum	Maximum
|X|	5 cm	180 cm
Y	-180 cm	180 cm
Z	20 cm	470 cm

Table 7.1: Definition of the Fiducial Volume (FV). The absolute X co-ordinate is used to avoid the cathode at X=0 between the TPCs.

Applying this FV cut removes 43.9% of the cosmic ray interactions that survive the unambiguous cosmic removal. Similarly, 90.3% of dirt interactions are removed, with split neutrinos additionally reduced by 38.4%. The FV interactions are only reduced by 2% for CC and 6.6% for NC, demonstrating the excellent vertex reconstruction achieved.

7.2.3 Reconstructed Shower Requirement

In order to measure the kinematics of ν_e CC interactions a slice must have a reconstructed shower to consider as an electron candidate. In particular, as the selection presented here focuses on interactions where $E_e^{True} > 200$ MeV it is required that a candidate shower must have a reconstructed energy $E_e^{Reco} > 200$ MeV. The reconstructed energy distribution of the remaining showers can be seen in Figure 7.2a.

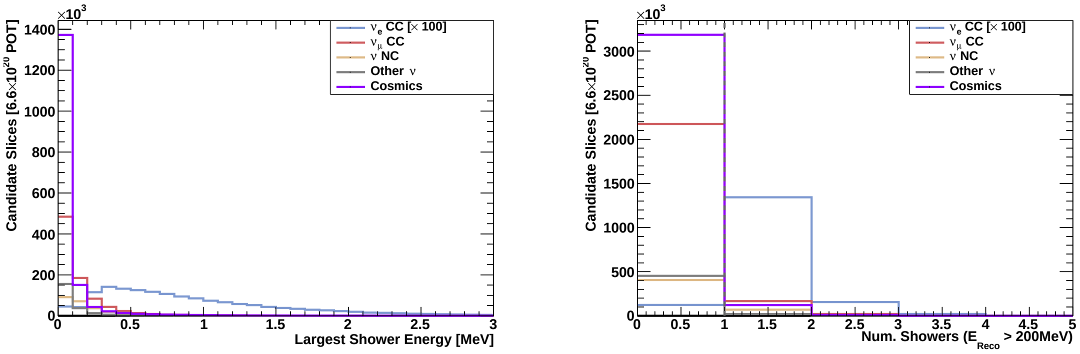


Figure 7.2: The distribution of leading shower candidate reconstructed energies, left, and the multiplicity of candidates per slice.

The backgrounds dominate the signal at this early stage of the selection, particularly at low energy, so the signal has been scaled up by a factor of 100 for visibility.

Requiring at least one shower above 200 MeV removes 91.4% (82%) of background ν_μ CC (NC) interactions and 95.8% of cosmic slices remaining after the FV cut. Other neutrino interactions are removed by 94.7%, including a 65.0% reduction in ν_e CC interactions which fail the true electron energy cut. However, 9.3% of ν_e CC interactions are removed, largely due to the shower segmentation issues discussed in Chapter 5 degrading the energy resolution.

Cuts to remove events with multiple showers, e.g. from interactions producing a π^0, currently perform poorly due to these shower segmentation issues. This can be seen in Figure 7.2b which shows the number of candidate showers per slice. A large fraction of the ν_e CC signal has multiple candidate showers per slice and thus no cuts are placed on this metric to avoid large efficiency losses. Future improvements to pattern recognition should reduce the segmentation issue and allow such a cut to be applied with less signal loss.

Current energy reconstruction methods for showers rely on summing the energy deposited within the detector, as discussed in Chapter 6.2, and thus perform poorly for uncontained showers. As such, additional requirements can be placed on the start and end of the reconstructed showers, with the end is defined as the start plus the direction multiplied by the length. The energy containment of these showers describes the fraction of a particles energy that is deposited in the detector, as given in Equation 7.1:

$$\text{Energy Containment} = \frac{\text{True Energy Deposited in TPC}}{\text{True Initial Energy}} \tag{7.1}$$

This energy containment of ν_e CC interactions is profiled across the start and end of the shower in Figure 7.3, with the edges of the detector indicated by the dashed lines. Each line represents a different combination of cuts that are placed on the start and end of the shower. As shown in the left column, the energy containment drops off as the shower start approaches the edges of the detector, in particular there is a large decline towards the downstream face of the detector, high Z. The right hand column demonstrates a similar drop off as the shower end position approaches the edges of the detector, although the drop off is significantly sharper.

The FV cut placed on the shower start position was the same as that applied to the neutrino interaction vertex, given in Table 7.1. However, the cut placed on

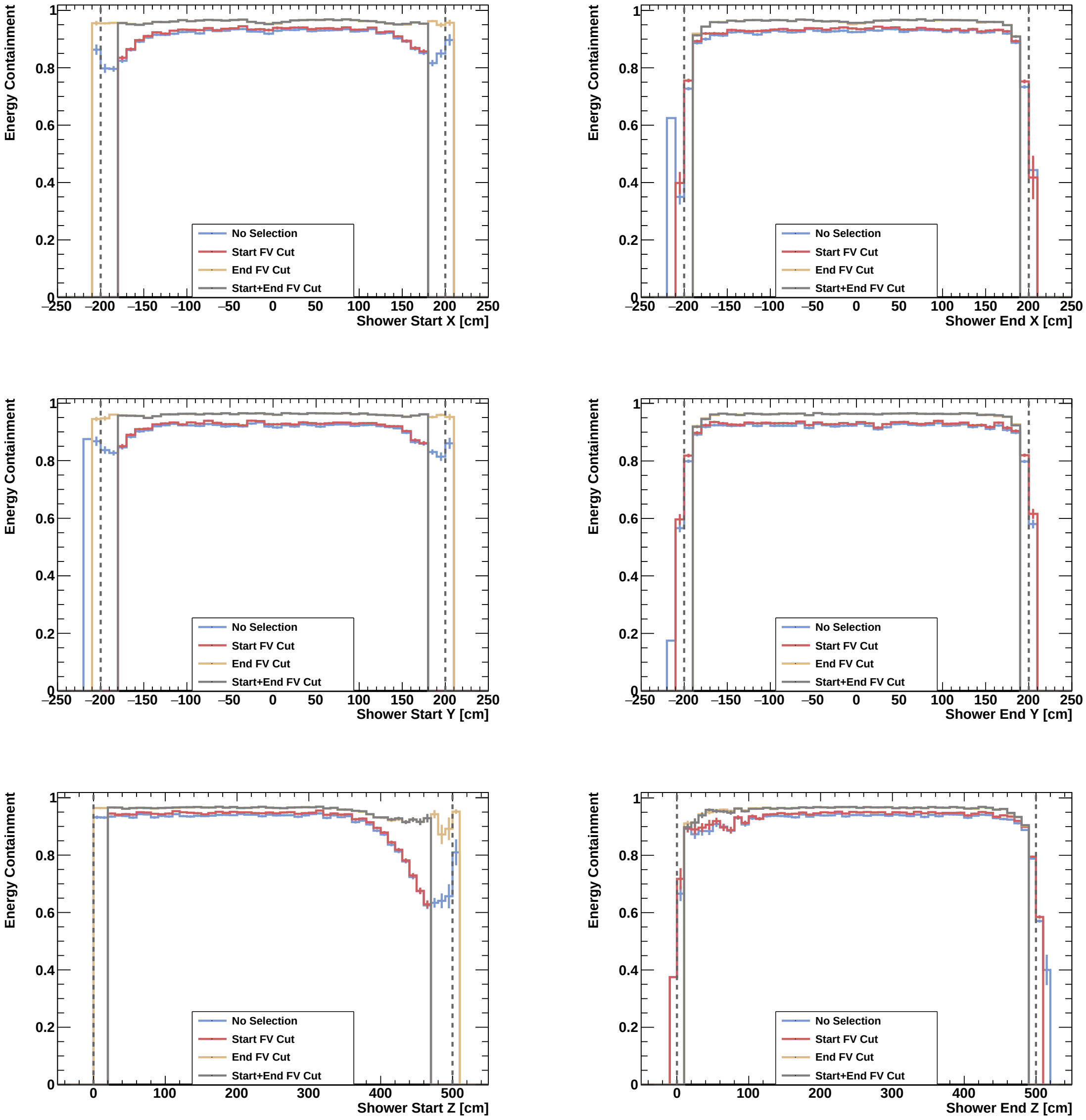


Figure 7.3: Reconstructed electron candidate energy containment profiled over the shower start, left, and end, right, positions. The edges of the detector are shown by the dashed lines.

the shower end position requires a distance of only 5 cm from the closest wall as the distributions from off more sharply. Comparing the cuts placed on the showers, it is clear that the cut placed on the end of the shower provides the largest uplift to energy containment, in particular reducing the drop off at downstream shower start positions. Placing a requirement on both the start and the end positions further increases the energy completeness and flattens the distributions.

The integrated distribution of energy containment is shown in Figure 7.4, with effect of the same cuts shown by the different lines. Again, this demonstrates that the largest improvement to the energy containment comes from applying the FV cut to the shower end position, with the start adding a marginal increase. By applying both of these cuts, over 86% of showers have an containment above 95%, a significant increase over the 68% without these cuts in place.

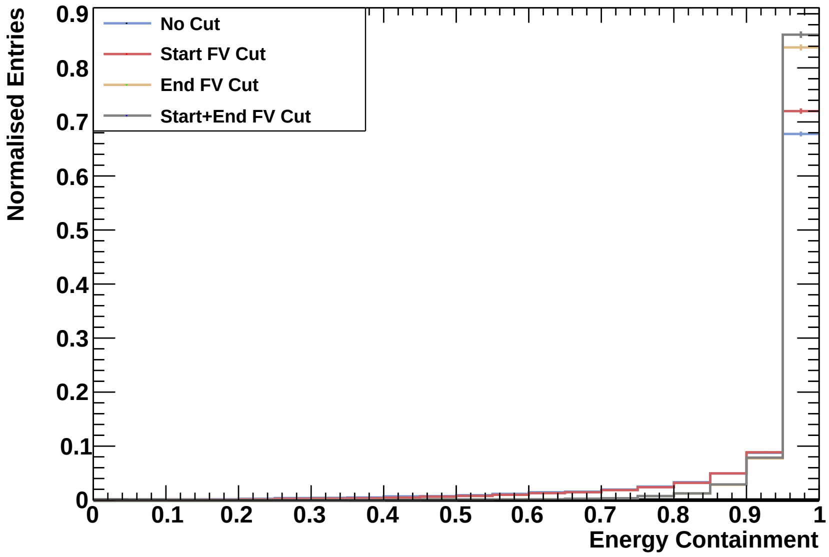


Figure 7.4: Comparion of the energy containment for combinations of fiducial volume cuts placed on the start and end of showers.

Applying these shower containment cuts comes at the cost of 17.9% of the ν_e CC signal. Similarly, 12.5% and 9.8% of the ν_μ CC and ν NC interactions are removed by this cut. Cosmic and dirt backgrounds are removed at slightly higher rates of 19.3% and 33.3% respectively. Despite the efficiency loss, this shower containment cut is crucial for ensuring that the energy of the selected events can be reliably estimated, vital for the measurements that will be discussed in Chapter 8.

7.2.4 Initial Purity and Completeness

The preselection was found to reduce the ν_μ CC and NC backgrounds by 92.8% and 85.2% respectively, largely due to the absence of showers present in these interactions. Additionally, the cosmic background has been reduced by 99.8%, with each of the cuts in the preselection rejecting a portion of this background. The other ν interactions have been reduced by 99.4%, consisting of a 99.7% reduction in dirt interactions, 97.3% reduction in split neutrinos, and a 69.0% reduction in ν_e CC interactions that fail the lepton energy requirement. However, the number of ν_e CC signal slices is also reduced by 28%, largely due to reconstruction inefficiencies in misclassifying and segmenting showers degrading the energy resolution.

Despite the large reduction in backgrounds, the sample remains background dominated, with the ν_e CC signal constituting only 3.2% of the remaining slices. The dominant background is the ν_μ CC interactions making up 44.7% of the sample with NC contributing 20.2%. Cosmic induced slices constitute 26.1% of the remaining sample, still posing a significant background after the unambiguous cosmic removal and FV cuts. The remaining 5.8% is made up of other neutrino interactions, dominated by dirt and split neutrino interactions in equal measure.

The distribution of reconstructed energies E_e^{Reco} can be seen in Figure 7.5, broken down by the true interaction that is matched back to the slice. Due to the low purity of the sample, the signal ν_e CC has been scaled up by a factor of 10 for visibility. The backgrounds dominate at low energy but drop off sharply towards higher shower

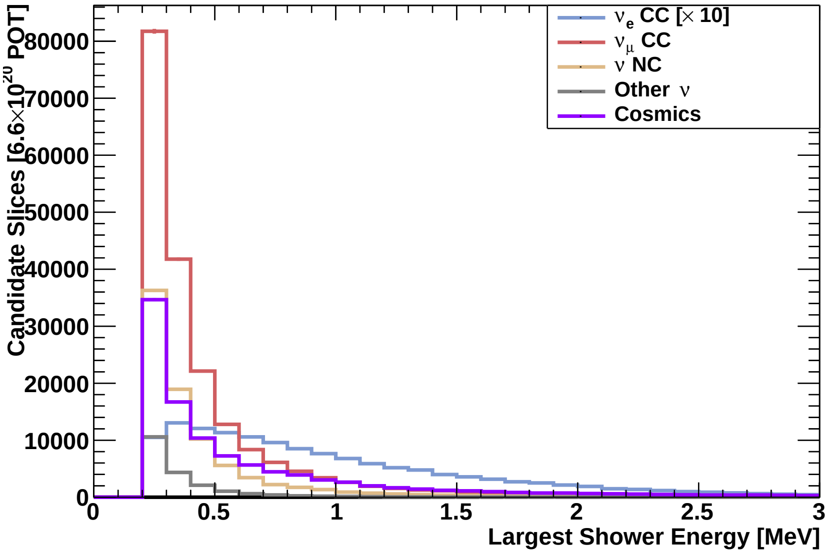


Figure 7.5: The distribution of leading shower candidate reconstructed energies after the preselection, broken down by the true interaction the slice is matched back to.

energies. As such, additional backgrounds could be rejected by raising the electron energy cut but this eliminates an important and interesting region of phase space for the ν_e CC interactions to be measured.

7.3 Cosmic Background Removal

The first selection cuts placed on the candidate slices are aimed at removing the dominant cosmic ray background, as such many of these cuts are common between ν_e and ν_μ selections[9]. These cuts can be separated into two categories: topological and timing cuts. The former is focused on the TPC activity to determine whether it is topologically neutrino-like or cosmic-like using the slice ID Multi Variate Analysis (MVA) discussed in Chapter 5. The latter utilises the timing resolution of the detector subsystems to differentiate between activity that occurred coincident with the beam spill, referred to as in-time, and activity outside this time window, referred to as out-of-time. Any out-of-time activity is assumed to be of cosmic origin and is thus rejected.

Liquid Argon Time Projection Chambers (LAr TPCs) generally have poor timing resolution due to the time required to drift electrons across the detector volume, in SBND 1.25 ms. Additionally, the degeneracy between activity at an early time and a large drift distance and activity at a later time closer to the anode makes the TPC timing alone somewhat unreliable. Certain topologies, i.e. cathode crossing tracks, can break the degeneracy but this is not possible for all tracks so the TPC can not reliably time tag all activity. When a slice can be tagged the time resolution is dominated by the readout frequency of the cold electronics of 0.5 μs[1]. This is comparable to the beam spill width of 1.6 μs so the TPC timing is not used to differentiate between in-time and out-of-time slices.

In contrast, the PDS and CRTs subsystems are capable of achieving timing resolution on the order of ns[1]. As such, by matching a TPC slice to the corresponding subsystem activity cosmic backgrounds that enter the TPC outside of the beam window can be rejected. This matching also breaks any degeneracies between the timing and drift position of any TPC activity, this is shown in Figure 7.6.

This diagram shows four interactions, coloured points, occurring at different times, x axis, and drift positions, y axis, in the detector. The electrons produced travel to the anode along the dashed line which indicates their drift trajectory and the intercepts represent the apparent position in the detector assuming they occurred at the trigger time. The black dashed line represents the start of the beam spill and trigger time, labelled T0, with the grey dashed line representing the end of the beam spill. The second dashed line represents the end of the main drift window, the time

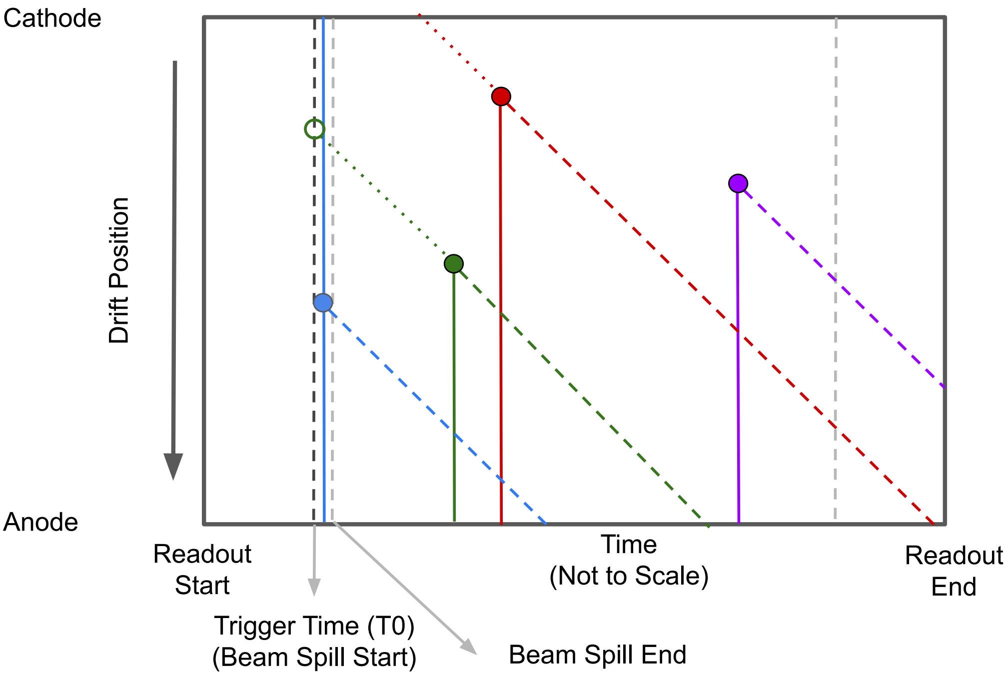


Figure 7.6: Diagram demonstrating the degeneracy between drift co-ordinate and interaction time in a LAr TPC. For an interaction, filled point, that happens at a given time, x coordinate, and drift position, y coordinate, the electrons drift to the anode along the dashed lines. The apparent position that the TPC sees the charge is shown by the intersection with the x axis or black dashed line denoting the trigger time. Four interactions are shown: an in-time interaction in blue, an out-of-time interactions in green and red and an interaction that is not read out in purple.

it would take for electrons produced at the cathode at the trigger time to drift to the anode.

The blue interaction is defined as in-time as it occurs during the beam spill time, marked by the vertical dashed lines. Under the assumption this interaction occurs at the trigger time one obtains the correct drift position for this interaction. Conversely, the green interaction occurs outside of this time window and so is considered out-of-time. By assuming that it occurred at the trigger time the apparent position, the hollow green circle, is offset from the true drift position, the solid circle. The TPC alone is unable to determine that this is out-of-time however as it is contained within the main drift window, this can only be done by matching the interaction to a detector subsystem with better timing resolution.

In contrast, the red interaction happened too late and close to the cathode to be read out in the main drift window, instead being read out in the “back porch”. The TPC is thus able to determine that this is out-of-time as it would not be contained in the TPC if it occurred at the trigger time. Finally, the purple interaction occurs too late to be read out by the TPC, so only a flash will be present for this activity and no charge. It is similarly possible for interactions to occur before the readout start but the electrons arrive during the readout window, such cases would have charge in the TPC but no flash. It is assumed that all activity occurs at the beam spill time, configured to be 0 in SBND, for any FV cuts and calorimetric corrections, e.g. electron lifetime attenuation.

7.3.1 TPC Cosmic Removal

Cosmic rays and neutrino interactions can be distinguished by their differing topologies: cosmic rays typically have a single long track, with accompanying delta rays, and neutrino interactions have multiple particles originating from a single vertex. To augment the unambiguous cosmic ray removal within Pandora an MVA is run to calculate a “Neutrino Score” for each slice. This score characterises how neutrino-like a slice is based on topological variables described in detail in Chapter 5.3.4. Running this MVA further reduces the number of cosmic rays using only the TPC without relying on the other detector subsystems.

The distribution of scores is shown in Figure 7.7, with a higher score indicating a more neutrino-like interaction. There is clear separation between cosmic slices and neutrino slices in the distributions, with ν_e CC generally having slightly higher scores than ν_μ CC and NC interactions as discussed in Chapter 5.3.4. Placing a conservative cut at 0.4 reduces the number of cosmic slices by 16.1% whilst removing

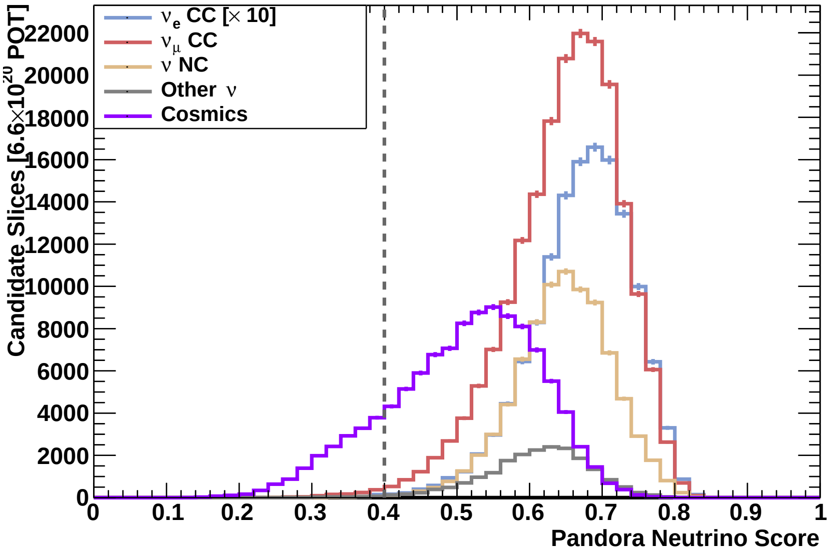


Figure 7.7: Distribution of Pandora“Neutrino Score” measuring how neutrino-like a given slice is. Slices shown are after the preselection previously discussed and break down each slice by truth matching information.

less than 1% of all neutrino interactions. As such, this cut proves as a useful addition to the unambiguous cosmic removal with little degradation to signal efficiency.

7.3.2 PDS Cosmic Removal

In addition to the ionised electrons, the scintillation light created when a charged particle traverses a LAr TPC can be collected. SBND has a comprehensive Photon Detection System (PDS) to collect this light, as described in Chapter 4. This PDS has ns level timing resolution which allows for identification of activity outside of the beam spill time[1].

By comparing the charge and light centers, the TPC and PDS activity can be matched together and in turn reject out-of-time slices, referred to as flash matching. As the PDS lies in the Y-Z plane behind the anode, the Y and Z coordinates of the flash can be calculated by taking the PE weighted average of the individual photon detector positions.

There are currently two ways to calculate the drift, X, position of the flash: using the flash spread and the coated/uncoated ratio. The further from the anode an interaction occurs the further the light has to spread out, thus the position in X can be estimated by the total spread of the light collected by the PDS. SBND has both TetraPhenyl Butadiene (TPB) coated Photo-Multiplier Tubes (PMTs) which collect the Vacuum Ultra-Violet (VUV) scintillation light directly and uncoated PMTs that collect the reflected light from the cathode mounted foils, as discussed in Chapter 4. The relative proportions of direct and reflected light detected is dependent on the X position of the interaction, with interactions closer to the cathode creating a higher proportion reflected light than interactions at the anode. Thus, the X position of the interaction can be estimated by the relative ratio between direct and reflected light, taking into account the number of each type of PMT. Ultimately these two methods are combined to give a refined estimate of the X position.

A comparison of the charge and light centers is performed to estimate the match quality using the following formula:

$$Score = \frac{|\Delta_x|}{\sigma_x} + \frac{|\Delta_y|}{\sigma_y} + \frac{|\Delta_z|}{\sigma_z} \tag{7.2}$$

Here Δ represents the difference between the flash and charge coordinates and σ the expected spread of each metric, derived from simulation, to account for the varying accuracy in each coordinate. A lower score represents a better match and a higher score a worse match. Unphysical matches, for example matches between light and charge in opposing TPCs, are assigned a score of -5 and are rejected by this analysis. This procedure is only capable of differentiating between in-time and out-of-time activity and thus in-time cosmic rays cannot be rejected as they are consistent with the light produced during the beam spill window.

The distribution of the score given by Equation 7.2 is shown in Figure 7.8, containing both in-time and out of time cosmics. The unphysical peak at -5 is dominated by cosmic slices contained in a TPC volume with no flashes during the beam spill window. Scores for neutrino slices peak at the expected value of 3, showing that the expected spreads of each score are being adequately estimated. The tail of high scores above 6 is dominated by out-of-time cosmics with bad matches to the in-time light. Requiring a score between 0 and 6 was found to reject 89.7% of all cosmics whilst removing only 6.6% of the ν_e CC signal. The ν_μ CC and ν NC backgrounds are additionally reduced by 15.0% and 6.6%. Other ν backgrounds are reduced by 26%, with a 16% reduction for dirt interactions and 39% for split neutrinos. This is the most powerful of the cosmic rejection tools presented and demonstrates the importance of the timing information afforded by the PDS.

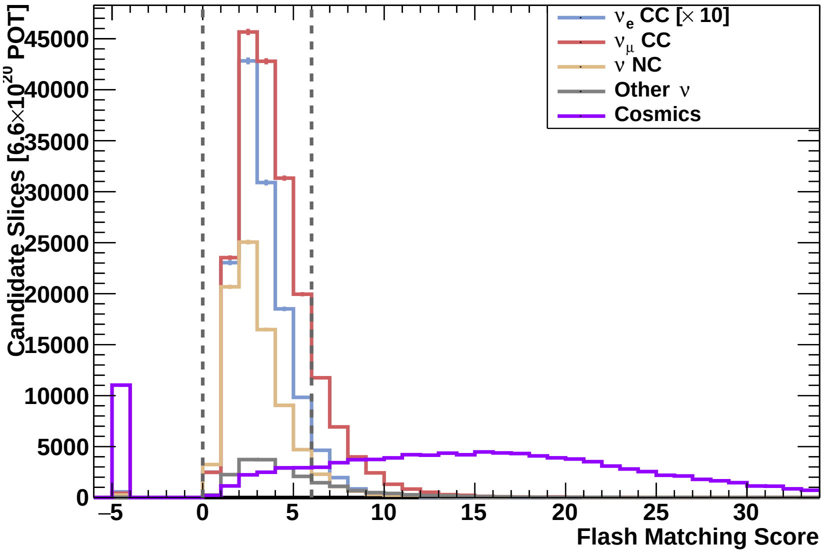


Figure 7.8: Distribution of flash matching score between in-time light and TPC slices. Lower scores represent better agreement between the light and charge whilst a score of -5 is assigned to incompatible matches, e.g. light and charge in different TPCs.

7.3.3 CRT Cosmic Removal

The ns level timing resolution afforded by the CRT system allows triggers caused by in-time cosmic rays, referred to as in-time cosmic events, to be rejected. Matching between the subsystems can be performed by extrapolating TPC tracks to the CRTs, accounting for any necessary timing offsets and finding the Distance of Closest

Approach (DCA). However, this approach has limited utility in this selection as all tracks are required to be contained within the detector, as will be discussed in Section 7.4.

Instead, a veto was developed to remove events with hits in the CRTs during the beam window can dramatically reduce the in-time cosmic events where no neutrino interaction occurs. Specifically, vetoing events which have a CRT hit during the beam spill window with at least 100 Photo-Electrons (PE), tuned to reject Minimum Ionising Particles (MIPs) whilst ignoring neutrons. This is effective for a ν_e selection due to the electron showers rarely inducing a CRT hit, in contrast to the muon tracks which create hits a significant fraction of the time. The distribution of CRT hit times for in-time cosmic triggered events is shown in Figure 7.9. A large number of hits occur at the trigger time, defined to be 0, which are successfully removed by the veto alongside a flat rate of random cosmic rays.

This veto was found to reject 80% of the in-time cosmic triggered events, significantly reducing a background which is hard to remove with other methods which are targeted at out-of-time activity. Due to the ns level timing resolution and low beam occupancy, the random occurrences of cosmic rays causing CRT hits in neutrino triggered events was found to be only 3%. Therefore, it was found that only 5% of ν_e CC interactions create a CRT hit compared to 23% (6%) of ν_μ CC (NC) interactions. Dirt interactions are similarly more likely to induce a CRT hit and are thus reduced by 31%.

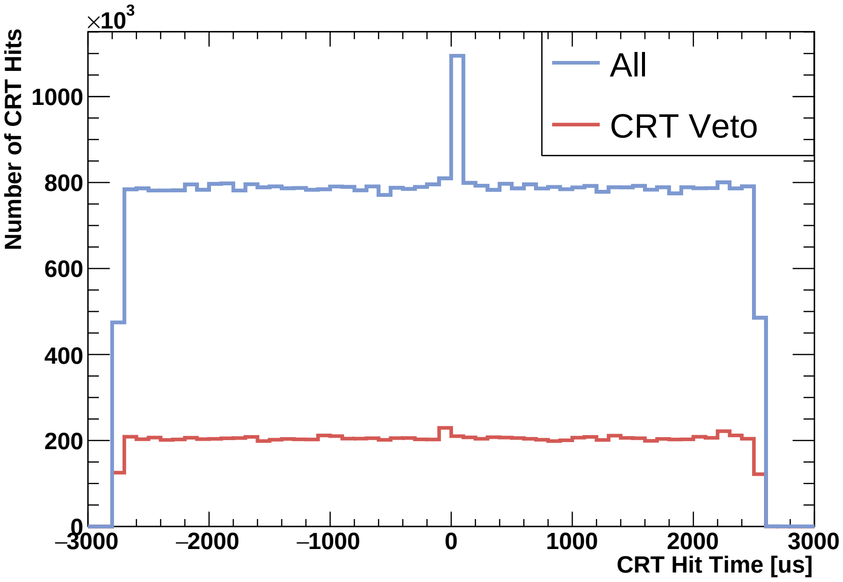


Figure 7.9: Distribution of CRT hit times for in-time cosmic triggered events, with the expected peak at the trigger time of 0 being reduced by the CRT veto.

7.3.4 Combined Cosmic Removal

Combining the cosmic rejection from the TPC, PDS and CRT discussed in this section was found to remove 93.7% of the remaining cosmic ray slices not removed by the preselection, reducing from 111,629 to only 6,993. When including the rejection provided by the preselection the total cosmic rejection rises to 99.99%. The remaining cosmics are now outnumbered by the ν_e CC signal, with only 10.9% removed by these cuts leaving 12,161 candidates. The ν_μ CC, ν NC and other ν backgrounds are reduced by 38.6%, 15.6% and 46.9%, slightly more than the ν_e CC signal primarily due to the CRT veto. The distribution of electron candidate shower energies for the remaining slices is shown in Figure 7.10 which shows that the dominant backgrounds are the ν_μ CC and NC backgrounds with 120,159 and 68,121 candidates respectively.

The efficacy of each of the cosmic removal cuts, including the combined cut, is shown in Table 7.2 which shows the fraction of events surviving each cut, after the unambiguous cosmic removal. Overall, the flash matching was the most powerful tool to reject backgrounds, whilst the CRT veto is particularly effective at removing the in-time cosmic events. Further improvements to the flash matching are under development to further increase the cosmic rejection power of the PDS by using more advanced matching techniques. The Pandora “Neutrino Score” removes a significant number of backgrounds with a negligible impact on the signal efficiency. Further improvements could also be made to this MVA by including additional variables, especially extending to use calorimetric variables, rather than just topological.

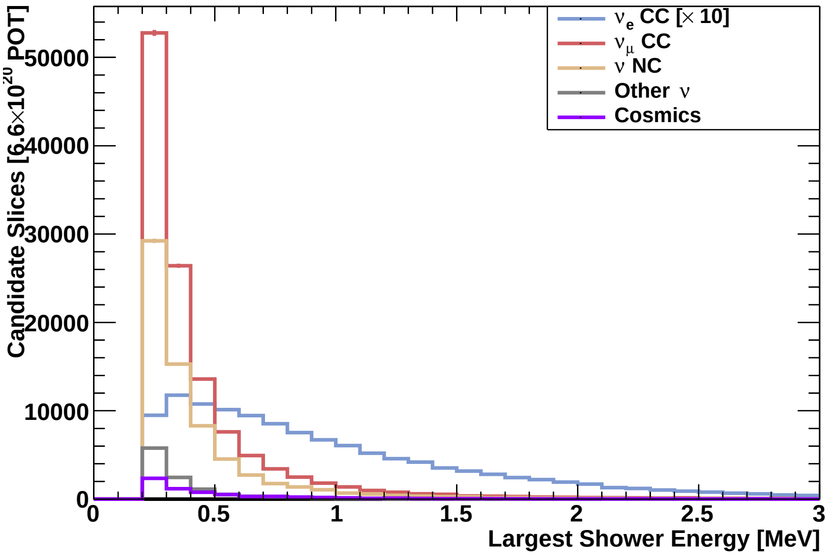


Figure 7.10: Distribution of electron candidate energies after all cosmic removal plotted by the true slice interaction.

True Type:	ν_e CC	ν_μ CC	ν NC	Other ν	Cosmics
CRT Selection	95.5	71.8	86.0	72.2	69.3
TPC Selection	99.7	99.4	99.8	99.0	83.9
PDS Selection	93.5	85.0	93.5	73.9	10.6
Combined Selection	89.1	61.7	80.4	54.1	6.3

Table 7.2: Selection efficiency for each cosmic removal cut after the unambiguous cosmic removal.

7.4 Muon Track Removal

After the cosmic background removal, the largest remaining background for the ν_e CC selection is ν_μ CC interactions, making up 55% of the remaining slices. These interactions are identifiable via the muon track which is not produced in either NC or ν_e CC interactions. Moreover, such track cuts are orthogonal to the cuts placed when trying to select the electron showers so provide an alternative pathway to reject this particular background.

The cuts developed to identify muons in the ν_μ CC selection can be utilised to identify these interactions as backgrounds for the ν_e CC selection[9]. Traditionally, this has involved placing cuts on topological and calorimetric variables to identify a track as a muon, pion, or proton, a process referred to as Particle IDentification (PID). An overview of these cuts is presented in Section 7.4.2. A Boosted Decision Tree (BDT) was developed to combine together the variables used in these traditional cuts to enhance the performance by utilising correlations between variables and including additional weak predictors. The design, implementation and performance of this BDT will be presented in Section 7.4.3.

7.4.1 Track Containment

Many of the PID cuts that are performed rely on having contained tracks, for example calculation the track length or identifying the Bragg peak produced when a particle stops within the detector. Therefore, requiring that all of the tracks within a slice are contained is a prerequisite for many future PID cuts. Moreover, as muons deposit less energy per unit distance travelled, they have a lower dE/dx, than protons they are more likely to exit the detector than protons, for a given energy and start position within the FV. Thus a requirement for track containment additionally removes ν_μ CC backgrounds.

A track is deemed to be exiting if the end point lies within 5 cm of a detector wall, the same requirement placed on showers. The number of exiting tracks in the remaining candidate slices after the cosmic removal is shown in Figure 7.11, broken down by the true interaction type the slice was matched to. Only 7.0%

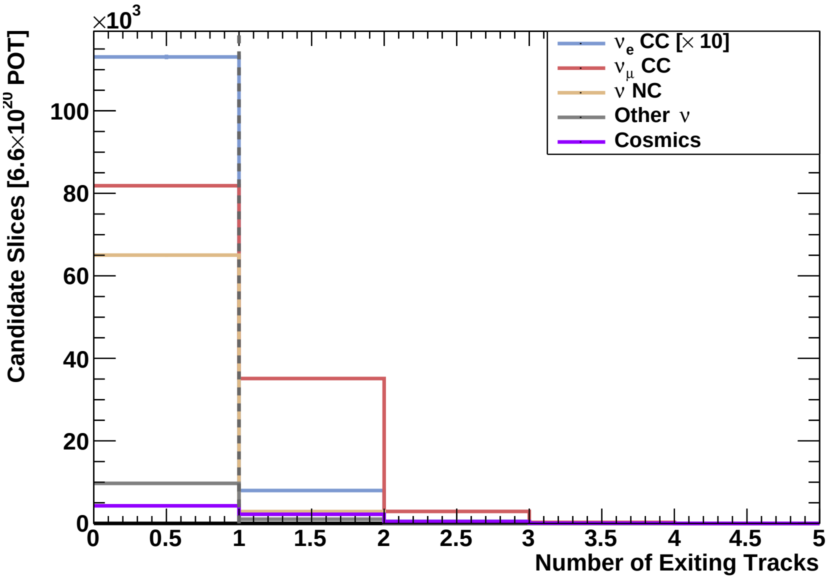


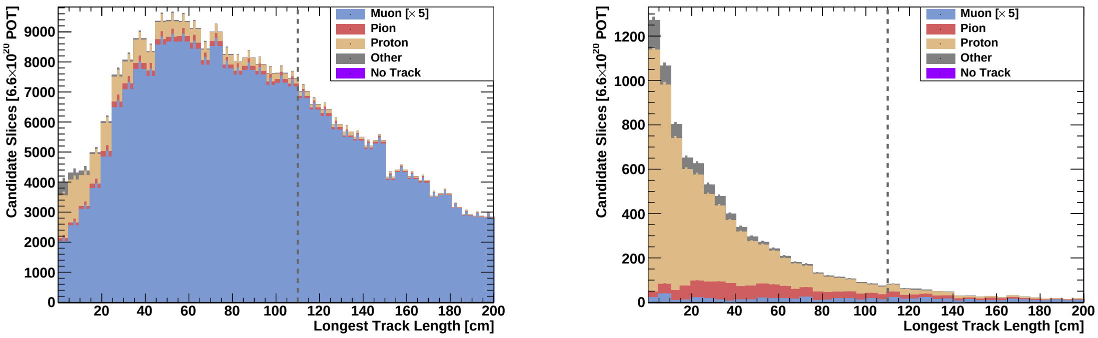
Figure 7.11: Distribution of the number of exiting tracks per candidate slice broken down by true interaction type. Exiting tracks are defined as tracks that end within 5 cm of any detector wall.

(4.5%) of fiducial ν_e CC (ν NC) slices contain any number of exiting tracks, so any containment cuts would have little effect here as expected. In contrast, 31.9% of ν_μ CC interactions and 38.9% of the remaining cosmic slices have at least one exiting track. Therefore, a cut was introduced to remove any slice that contains any number of exiting tracks, which are now removed from the remainder of the chapter.

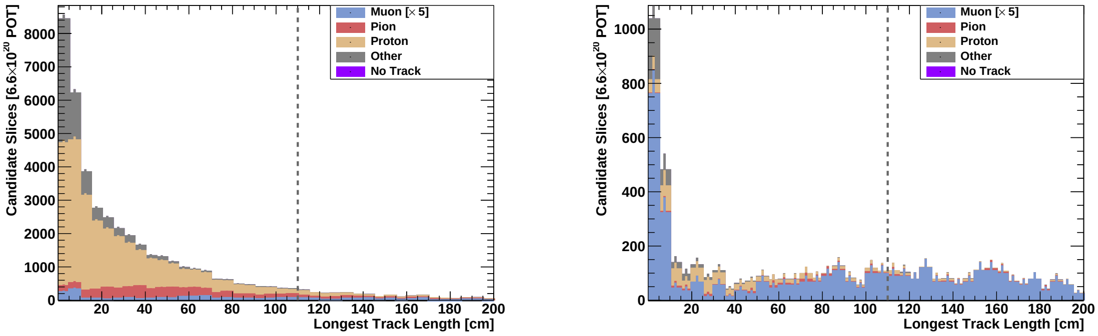
Many of the remaining slices contain no reconstructed tracks and are thus not affected by any of the cuts discussed in this section. Only a relatively small fraction, 9.7%, of ν_μ CC slices contain no tracks. However, much larger fractions of ν_e CC, NC, and cosmics have no tracks at 25.3%, 38.1% and 50.0% respectively. In many of the ν_μ CC and cosmic trackless cases the muon has been reconstructed but misclassified as a shower, this will be discussed in more detail in Section 7.5.

7.4.2 Muon ID Cuts

Now that the remaining sample consists entirely of contained tracks, cuts can be placed on these tracks to identify muons. Similarly to how the largest shower in a slice was defined as being the electron candidate, the longest track in a slice is defined as the muon candidate. This assumption can be validated by looking at the true simulated particle that matches back to this longest track in truth. This is shown in Figure 7.12 which shows the length of the longest track broken down by the true particle the track was matched to for each interaction type. This shows that in



(a) ν_μ CC (b) ν_e CC



(c) NC (d) Cosmic

Figure 7.12: Distribution of the length of the longest track for different interaction modes after requirement for track containment. Each histogram is stacked and broken down by the true particle that matches back to the track.

72.3% of the ν_μ CC slices, ignoring those with no tracks, the longest track matches back to a muon. The remainder is dominated by protons, 19.6%, and pions, 5.7%, which typically occur at low track lengths. In contrast, ν_e CC slices with tracks are dominated by protons, 75.6%, and pions, 15.3%, with the remainder being made up of misclassified showers. Neutral current slices are similarly made up of 64.6% protons and 14.4% pions with a relatively larger fraction of misclassified showers. Muons make up the longest track in 40.6% of cosmic slices with the remainder made up of protons, 32.1%, and misclassified showers, 19.5%.

Due to the combination of low dE/dx and higher typical initial energy, muons tend to produce longer tracks than protons, and to a lesser extent pions. As such, placing a cut on the length of the muon candidate in each slice is a first step towards identifying and rejecting muons as demonstrated by the distribution of track lengths shown in Figure 7.12. The maximum track length for each slice is shown in Figure 7.13, broken down by interaction type.

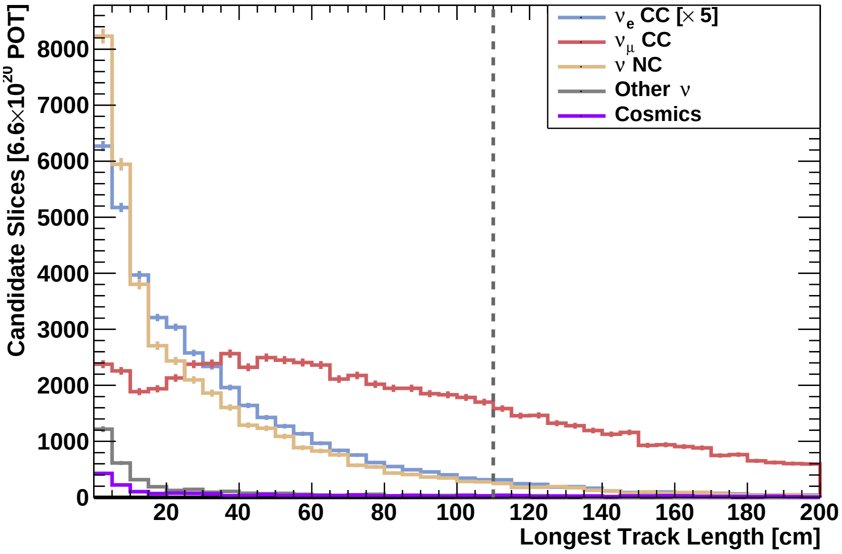
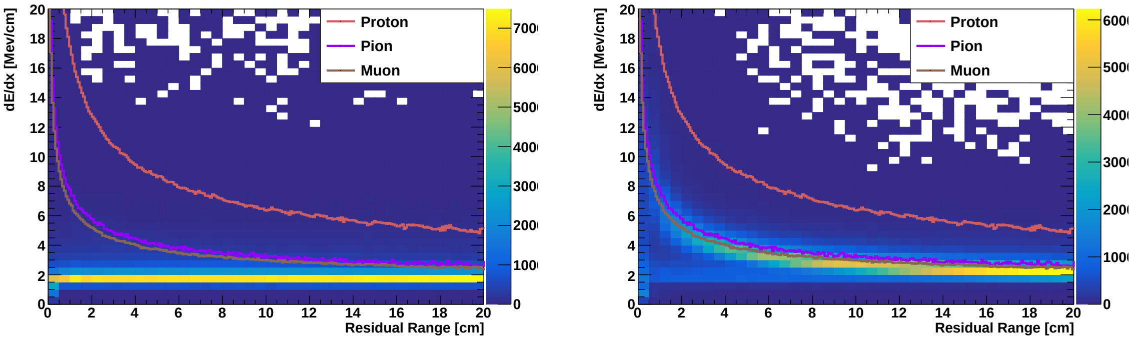


Figure 7.13: Distribution of the length of the longest track broken down by the true interaction type of each slice after track containment requirement.

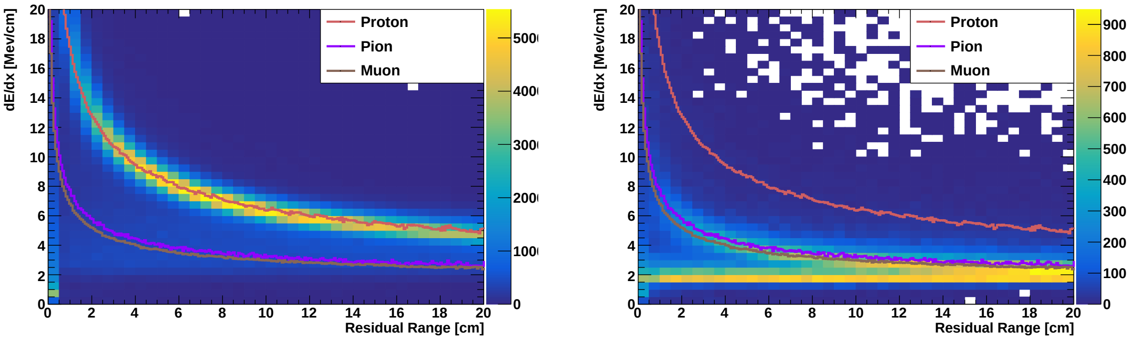
By rejecting slices where the muon candidate is longer than 110 cm, indicated by the dashed line, 31.8% of ν_μ interactions are rejected whilst 95.5% and 96.6% of ν_e CC and NC interactions are kept respectively. As can be seen in Figure 7.12, this mainly removes slices where the longest track matches back to a muon. Specifically, across all interaction modes this cut rejects 55.7% of the slices where the longest track is a muon compared to 17.8% and 5.7% for pions and protons respectively. Lowering this cut further would remove additional muons but additionally begins to remove many more slices where the longest track is a proton, including ν_e CC signal.

Rather than just lowering the track length cut, an additional PID check on the particle calorimetry can be applied to ensure the track is consistent with the expectations for a muon track, and not a proton track. This is performed using the χ^2 PID which compares the energy loss profile, dE/dx vs residual range, to a series of templates for different particle hypotheses, in this case protons and muons. A sample of ν_μ CC events was generated in the FV previously discussed in Section 7.2.2 in order to investigate the energy loss profile of different particle types and topologies. This energy loss profile is shown in Figure 7.14 for the last 20cm of exiting muons, contained muons, contained protons, and contained pions. The, simulation derived, expected behaviour for stopping particle of each type is overlaid for comparison.

Exiting muons shown on the top left show a flat distribution of dE/dx across the residual range, yielding poor agreement with all hypotheses. As previously eluded to, the decision to remove any uncontained tracks was, in part, motivated by the poor performance of this PID method. In contrast, stopping muon, pions, and protons



(a) Exiting Muons (b) Contained Muons



(c) Contained Protons (d) Contained Pions

Figure 7.14: Track energy loss profiles, dE/dx vs residual range for various particle types and topologies. The expected profiles for protons, pions and muons used to calculate the χ^2 PID are overlaid.

exhibit a Bragg peak, indicated by the rise in dE/dx at low residual range, and thus demonstrate good agreement with their respective hypotheses.

These contained muons clearly show a Bragg peak consistent with the expected distribution shown in brown, although it is also similar to the pion hypothesis in purple. Nevertheless, this sample is clearly distinct from the proton hypothesis shown in red. Similarly, the contained protons shown in the bottom left show good agreement with the proton hypothesis and clear separation from both the muon and pion hypotheses. As such, this method of looking at the energy loss profile can be used to differentiate MIPs, muons and pions, vs Highly Ionising Particles (HIPs), protons, but performs poorly at distinguishing between muons and pions.

However, the contained pion sample in the bottom right show two topologies: one that matches expectation and one flat distribution similar to that of the exiting muons. The flat distribution was found to be caused by pions that hadronically interact before losing all of their energy via ionisation and thus do not have a Bragg peak. The fraction of each particle that stops via ionisation is shown in Figure 7.15,

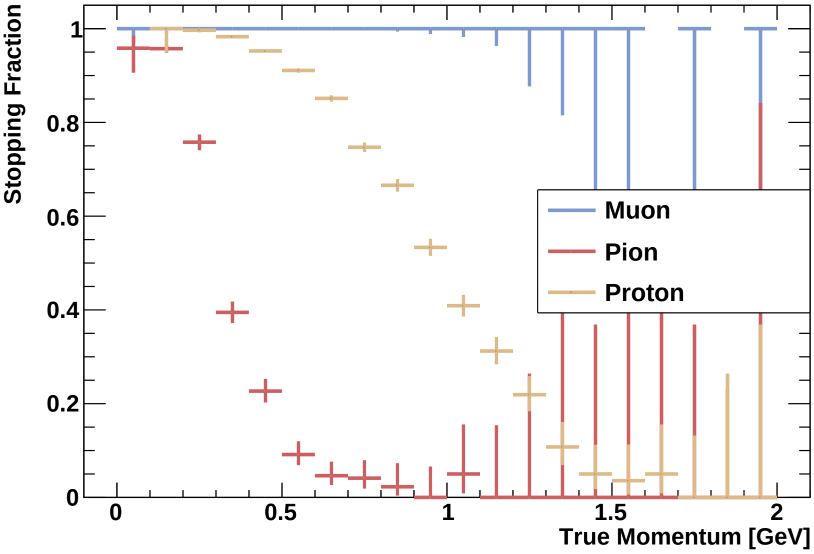


Figure 7.15: Fractional distribution of the contained particles that stop via ionisation as a function of true initial momentum for muons, pions and protons.

for the same sample of contained muons, pions and protons. This shows that almost all contained muons stop via ionisation, regardless of their energy. Some protons interact hadronically before losing all of their energy via ionisation, but this only becomes dominant above 1 GeV and the majority of the protons produced in the BNB are below this energy. In contrast, hadronic interactions for pions are much more common in SBND and become dominant above 300 MeV. These interacting particles can therefore be identified by the absence of a Bragg peak, something that is used in the PID that will be discussed in section 7.4.3.

The reduced χ^2 is calculated for each track using the muon and proton hypotheses to quantify the level of agreement, the particle is typically classified as the hypothesis with the lowest χ^2[140]. This test can also be performed against a kaon hypothesis but this is not performed as the BNB is at sufficiently low energy that the production of kaons is negligible. The distribution of the reduced χ^2 for the longest track for both the muon and proton hypotheses, χ^2_μ and χ^2_p respectively, is shown in Figure 7.16 for all remaining slices after the track containment cut. This shows that muons and pions generally have low χ^2_μ and high χ^2_p indicating good agreement with the MIP hypothesis. Conversely, protons show the opposite behaviour and are consistent with the HIP hypothesis.

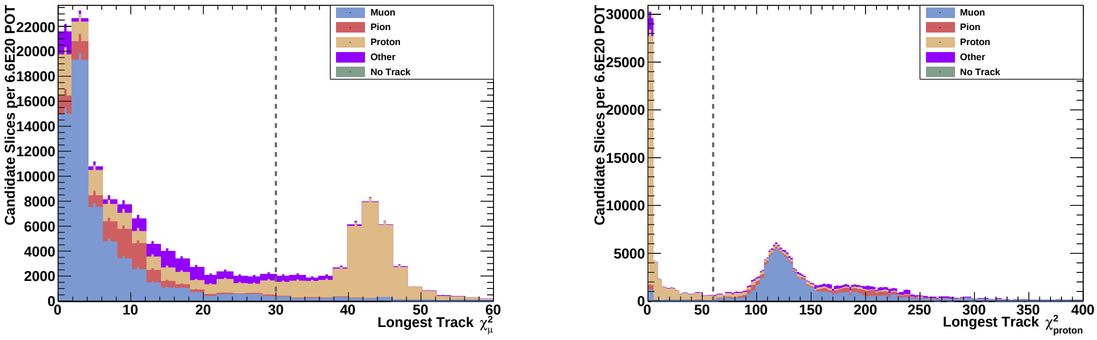
Placing a requirement on both χ^2_μ and χ^2_p largely separates the muon and pion tracks from the proton tracks. As such the cut on track length can be refined to only remove slices where the longest track is consistent the muon hypothesis, defined as satisfying all of the following criteria:

- **Track Length:** Require that the length of the track is at least 80 cm
- **MIP Hypothesis:** Require the track is consistent with the MIP hypothesis ($\chi^2_\mu < 30$)
- **HIP Hypothesis:** Require the track is inconsistent with the HIP hypothesis ($\chi^2_p > 60$)

By adding these extra checks and lowering the length threshold it was found that an additional 27.5% of the remaining muon tracks are removed compared to the simple track length cut. Pion tracks are similarly rejected 14.7% more often due to the lack of separation between muons and pions. However, 2.1% fewer proton tracks are removed by this cut despite the lower length threshold, demonstrating the powerful separation between muons and protons. This corresponds to an additional 14.8% reduction in the remaining ν_μ CC slices compared to the simple track length cut with only 0.6% and 0.3% losses to ν_e CC and NC interactions respectively.

The distribution of remaining track lengths is shown in Figure 7.17, with the cut at 80 cm indicated by the dashed line. This shows that at 80 cm there is a large drop in the number of remaining ν_μ CC interactions compared to ν_e CC and NC interactions which continue their gradual decline shown in Figure 7.13. This demonstrates that the χ^2 PID is protecting the proton tracks whilst removing the muon tracks as intended.

However, this also demonstrates the inability of the χ^2 PID to sufficiently differentiate between muons and pions. Without separation between muons and pions, the aggressiveness of such cuts is limited before the signal efficiency is reduced by removing ν_e slices with pion tracks. This loss of efficiency additionally risks introducing biases and model dependencies into an inclusive selection, as the selection



(a) χ^2_μ (b) χ^2_p

Figure 7.16: Distributions of χ^2 calculated for the energy loss profile of longest tracks compared with the muon, left, and proton, right, hypotheses, stacked by track truth matching for all remaining slices after track containment cut.

efficiency will be highly dependent on the level of pion production. Therefore, to further reduce the ν_μ CC background improved track PID is requires, particularly for identifying pions.

7.4.3 Dazzle Track PID BDT

In order to improve the track PID, a multi-classification BDT was developed to replace the traditional cuts, referred to as Dazzle, and implemented using TMVA[176]. This BDT not only includes additional variables but can additionally exploit correlations between variables, improving the performance over traditional cuts. In particular, BDTs can be extremely effective at including variables that perform well in a limited area of phase space.

This multi-classification approach consists of a BDT for each of the desired output classifications, in this case four: muon, pion, proton, and other. Each of these BDTs produces a score of how consistent an input track is with their hypothesis; i.e. the proton BDT gives a score of how proton-like a given track is. The BDTs are trained in parallel and normalised such that the sum of scores across all hypotheses sums to unity, thus each score can be treated as a probability of that hypothesis.

These output scores can be used in two ways: the particle can be classified as the hypothesis with the highest score or a simple cut can be placed on the scores. For example, to classify a track as a muon either a requirement that the score from the muon BDT exceeds that of all other BDTs or just that the score is above a given

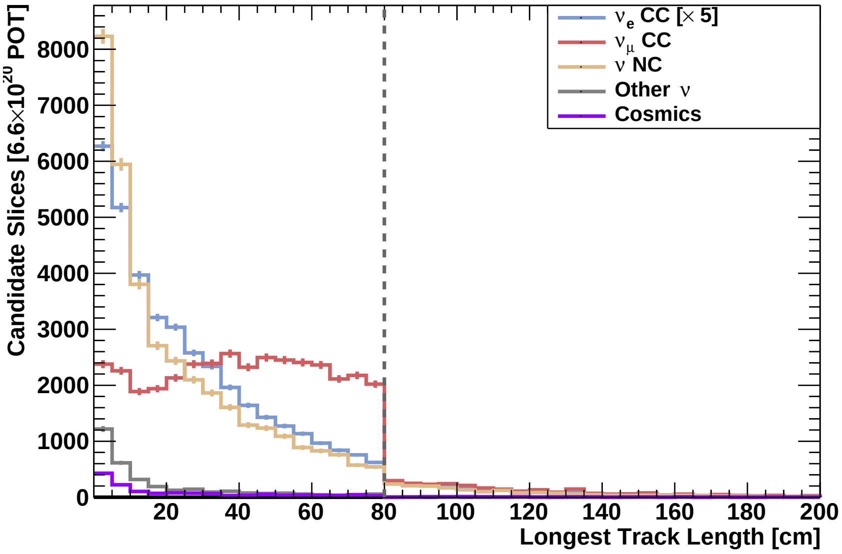


Figure 7.17: Distribution of the length of the longest track broken down by the true interaction type of each slice after track containment and muon candidate removal.

threshold can be placed. Classifying the track as the hypothesis with the highest score ensures that every track will be classified in, exactly, one output. This avoids any ambiguities with double counting or neglecting tracks but leaves little flexibility for tuning cuts for specific analyses. In contrast, optimisation of a single cut can be performed for each individual analysis and affords great flexibility. This introduces the risk of a track falling into multiple, or no, categories however which is not always desirable.

The input variables used to construct these BDTs are listed below and are predominately based on previous studies performed[9], with distributions for each variable shown in Appendix D:

- **Track Length:** The length of the track can be used not only to differentiate muons and protons, as previously discussed, but additionally to estimate the quality of other metrics, whose calculation is often worse for shorter tracks.

- **Stopping χ^2 Ratio:** In order to identify the presence of a Bragg peak two fits are performed to the energy loss profile of the last 20 cm of the track, the same range as used in the χ^2 PID shown in Figure 7.14. An exponential and 0-degree polynomial are both fitted to model the stopping and non-stopping hypotheses respectively. The ratio $\chi^2_{Pol0}/\chi^2_{exp}$ compares the goodness of fit between the two hypotheses with a low ratio suggesting a non-stopping track and a high ratio the presence of a Bragg peak. This absence of a Bragg peak is indicative of a hadronic interaction, and can thus help in distinguishing between muons and pions.

- **Fitted dE/dx:** The best fit from the 0-degree polynomial fit discussed above can be used to estimate the dE/dx of the track. For non-stopping tracks this can be used to differentiate between MIPs and HIPs.

- **χ^2_μ:** The stopping χ^2 under a muon hypothesis previously discussed

- **χ^2_p:** The stopping χ^2 under a proton hypothesis previously discussed

- **$\chi^2_\mu - \chi^2_\pi$:** The difference in stopping χ^2 between muon and pion hypotheses. This is used rather than just the χ^2_π to reduce correlations with χ^2_μ

- **Scattering Distance of Closest Approach (DCA):** The average transverse displacement of the track trajectory points from the axis of the track, defined by interpolating the start and end positions. Muons typically undergo the most Multiple Coulomb Scattering (MCS) so have the largest displacement whilst protons typically have the least.

- **MCS Scattering Angles:** The average scattering angle between consecutive segments of the track, using the MCS calculation discussed in Chapter 6.1[173]. This is a complementary way of measuring the scattering which can reduce the impact of a single large scatter.

- **MCS Scattering Ratio:** The ratio between the largest scatter in the MCS calculation and the mean discussed above. This can be used to either identify a single large hadronic scatter or the presence of a stopping particle, as the angles increase as the particle loses energy.

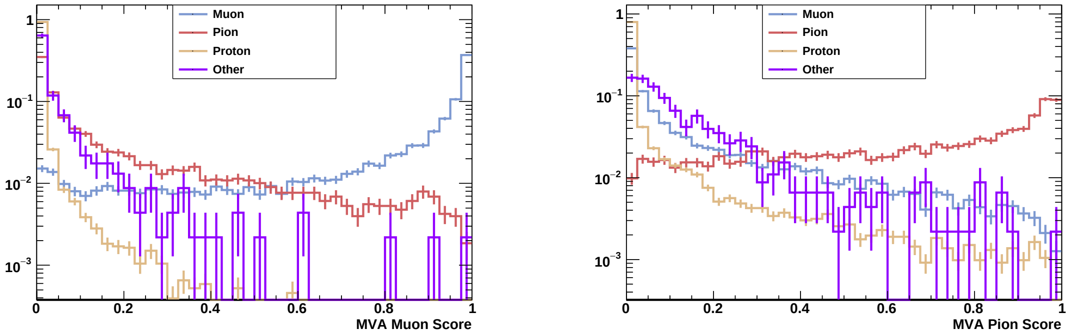
- **Number of Daughters:** The number of daughter particles defined in the particle flow hierarchy by Pandora. This can be used to identify both stopping muons with Michel electrons or hadrons that interact inelastically, aiding in the differentiating between muons and pions.

- **Max Daughter Hits:** The number of hits in the largest daughter particle. This can be used to differentiate between Michel electrons vs hadronic interactions.

- **Momentum Agreement:** The relative agreement between the MCS and range based momentum estimators, both assuming the muon hypothesis, discussed in Chapter 6.1. For contained particles, there should be good agreement for muons and poor agreement for protons and pions, particularly for those that interact hadronically.

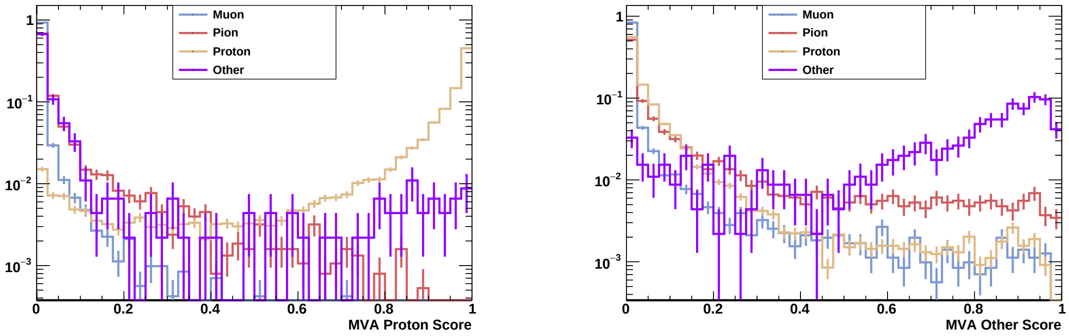
These BDTs were trained on a sample of half ν_μ and half ν_e events using the nominal BNB simulation. This training follows the same procedure discussed in Chapter 5.3.1 of splitting the sample into separate train and test samples. The BDT was then trained on the training sample and validated on an unseen test sample. Kolmogorov-Smirnov Tests (KS tests) checks are then performed to check performance is similar across the training and test datasets to prevent overtraining.

The performance of this BDT can be seen in Figures 7.18 and 7.19. The former of these shows the BDT scores under the muon, pion, proton and other hypotheses with each line representing a type of true particle. The protons achieve the largest separation from the other particles, indicated by the high scores under the proton hypothesis and low scores under all other hypotheses. This is expected as these are the most calorimetrically and topologically distinct, as was seen for the χ^2 PID. Muons and pions are somewhat separated but have the most overlap of each of the particle types as is expected, particularly for stopping pions which are almost indistinguishable from muons. Other tracks, mostly misclassified electrons and photons, achieve decent separation from the other hypotheses.



(a) Muon

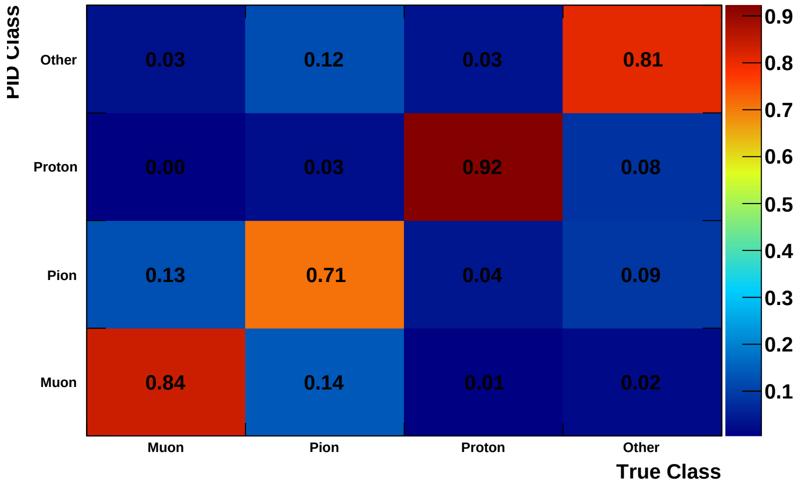
(b) Pion



(c) Proton

(d) Other

Figure 7.18: Dazzle Track PID BDT scores under muon, top left, pion, top right, proton, bottom left, and other, bottom right, hypotheses. Each line represents tracks matching back to the specified true particle and a higher score a higher probability of that hypothesis.



PID Class	Muon	Pion	Proton	Other
Other	0.03	0.12	0.03	0.81
Proton	0.00	0.03	0.92	0.08
Pion	0.13	0.71	0.04	0.09
Muon	0.84	0.14	0.01	0.02

Figure 7.19: Confusion matrix comparing true track types to the classification from Dazzle, column normalised to unity.

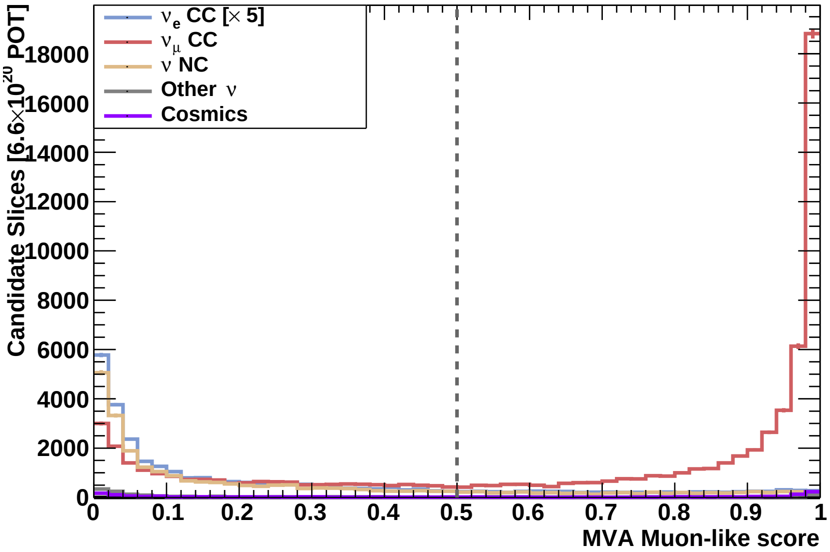


Figure 7.20: Dazzle Track PID BDT scores under muon hypotheses broken down by the true interaction type of each slice after track containment requirement.

The confusion matrix shown in Figure 7.19 takes the hypothesis with the highest score as the PID classification of the particle, a perfect PID algorithm would produce a diagonal matrix. Again, the performance is best for protons which are correctly identified 92% of the time with only the other tracks being misidentified as protons at a significant rate. Muons are correctly identified 84% of the time and are most often misclassified as pions, 13% of the time, as expected due to the similarities between stopping muons and pions. Pions are only correctly identified 71% of the time with misclassifications split between muon and other hypotheses fairly evenly. Finally, the other hypothesis is classified correctly 81% of the time and is misclassified as either a proton or pion in almost equal proportions. Overall, the majority of classifications lie on the diagonal demonstrating good performance across all particle types.

In the ν_e CC selection, this BDT is run over the muon candidate for each slice. The distribution of scores for the longest track in each slice is shown in Figure 7.20. Rejecting any slices where the muon candidate has a muon-like score above 0.5 removes 59.3% of ν_μ CC interactions and 10.3% of ν_e CC, compared to 32.8% (4.5%) respectively for the track length cut and 42.8% (5.1%) for the χ^2_μ (track length) cuts. Additionally, NC events are removed at 7.6% compared to 4.4% and 4.6% for the track length cuts respectively.

The distribution of track lengths for the longest track in each of the remaining slices is shown in Figure 7.21. This demonstrates that the BDT does not place a harsh cut on the track length and how many tracks with lengths exceeding the original cut at 110 cm remain.

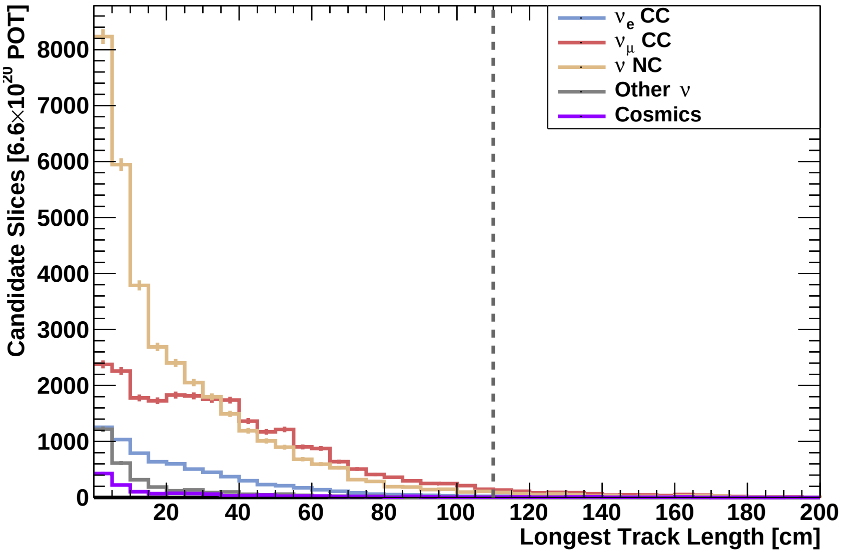


Figure 7.21: Distribution of the length of the longest track broken down by the true interaction type of each slice after track containment and Dazzle track PID BDT muon candidate removal.

Across all interaction types, this cut reduces the number where the longest track matches back to a muon by 72.8% compared to 28.1% and 8.6% for pions and protons respectively. This demonstrates that the BDT succeeds in identifying muons the majority of the time, with pions more commonly misidentified than protons as expected. Prior to this cut, muons were the longest track in 75.2% of the ν_μ CC slices, where 7.3% have no reconstructed track. Comparatively, this is reduced to 27.3% after the BDT has been run, with protons being the longest track 34.9% of the time and 22.5% of the remaining ν_μ CC slices having no reconstructed track. Any further attempts to reduce this background via cuts on tracks will have to revisit the assumption that muons comprise the longest track in the slice and address upstream reconstruction inefficiencies, particularly in misclassifying muons as showers.

7.4.4 Summary

Overall, the muon track cuts successfully remove a large portion of the ν_μ CC background, with the efficacy of the various cuts are compared in Table 7.3. This table calculates the efficiency after the track containment cut has been applied, reducing the ν_μ CC and cosmic backgrounds by 31.9% and 39.9% respectively at the expense of 7% of the ν_e CC interactions. This demonstrates the importance of the PID to identify the muon track, indicated by the performance increase when including the χ^2 PID in the muon track selection. The Dazzle selection further improved upon this performance by including additional variables to identify muon tracks, achieving the best overall performance.

True Type:	ν_e CC	ν_μ CC	ν NC	Other ν	Cosmics
Track Length Selection	95.5	68.2	96.6	97.4	86.2
Muon Track Selection	94.9	58.2	96.2	96.5	84.8
Dazzle Selection	89.7	40.7	92.5	94.9	82.6

Table 7.3: Selection efficiency for each of the muon track cuts after the track containment selection.

The number of slices where the longest track is a muon has been reduced by 82.0% compared to 34.5% and 13.1% for pions and protons respectively, demonstrating the performance of the PID. As such, only 8.6% of remaining slices have a muon as the longest track, including 27.3% of ν_μ CC interactions. In contrast, 38.2% of the remaining slices have no reconstructed tracks, including 22.5% of ν_μ CC interactions. The majority, 60.1%, of the remaining cosmics have no remaining tracks with only 10.1% where the longest track is a muon, reduced from 43% before the track cuts. Overall, this demonstrates the cuts are performing well at removing the events with muon tracks but are limited by the current reconstruction and paradigm taking the longest track as the muon candidate.

After being reduced by 62.3%, the ν_μ CC interactions now make up only 28% of the remaining slices, reduced from 55% after the cosmic removal. Cosmic induced slices have been reduced by 49.7%, primarily from the containment cut, and only constitute 3.0% of the remaining slices, a slight reduction from 3.2% previously. The dominant background is now the ν NC interactions, making up 52% of the remaining slices having being reduced by only 11.7%. The ν_e CC signal is reduced by only 16.6% and now has a 9.0% purity, up from 5.6% before the track cuts were applied.

7.5 Electron Shower Selection

Identifying an electron is necessary for not only selecting ν_e CC interactions but also measuring the interaction kinematics. This is a powerful tool to reduce remaining backgrounds, especially NC interactions with π^0 induced photon showers. The preselection ensures that all remaining slices contain an electron candidate shower upon which a series of cuts can be placed. The traditional cuts that are used to distinguish between electron and photon showers are first applied individually and later combined using a BDT.

7.5.1 Electron ID Cuts

Differentiating between electrons and photons is a key ability of LAr TPCs to efficiently select ν_e CC events[175]. Nevertheless, this identification is challenging due to the stochastic nature of showering particles like electrons and photons. The electron candidate in the remaining background slices after the muon track removal are dominated by photon showers, 76.9%, with misclassified muons and protons making up 9.2% and 10.3% respectively. As such, the upcoming cuts are focused on differentiating between photons and electrons.

Neutral particles, such as photons, do not deposit any energy via ionisation in a LAr TPC and are only detectable by the charged particles produced when they interact. As such, there is a gap between the start of the photon, typically at the neutrino interaction vertex, and where the visible energy is deposited in the detector.

As the distribution for photons is expected to be an exponential decay, the most common place for a photon to interact is in the first bin. Therefore, even with perfect reconstruction those photons that interact within a single wire of being produced, corresponding to 3 mm in SBND, cannot be rejected.

This conversion gap can be detected in reconstruction, manifesting as the distance between the neutrino interaction vertex identified by Pandora and the shower start position. The distribution of this metric for the electron candidates in the remaining slices after the Dazzle cut is shown in Figure 7.22. Electron showers display no

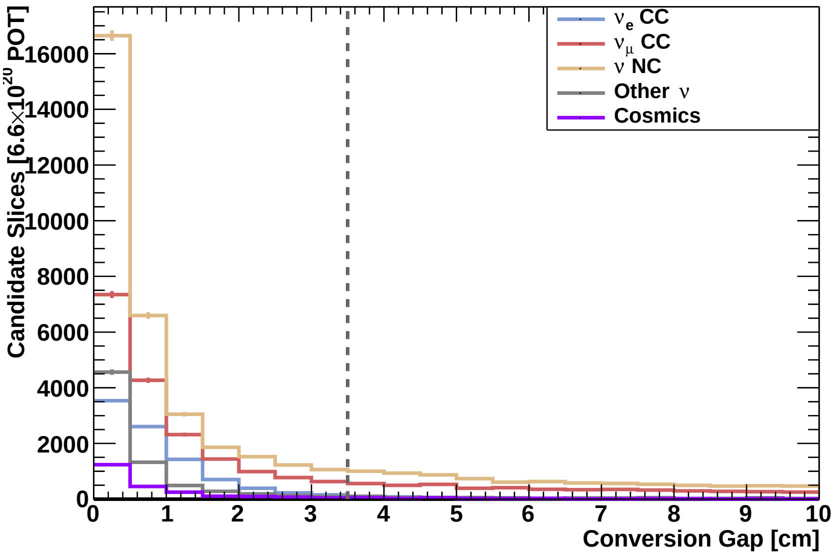


Figure 7.22: Conversion gap between neutrino interaction vertex and shower start position broken down by the true interaction type of each slice after Dazzle muon candidate removal. Dashed line represents the cut applied at 3.5 cm.

such gap and begin to shower immediately, although imperfections in reconstruction and limited detector resolution smear this. Therefore, placing a cut on the distance between the neutrino interaction vertex and shower start position can differentiate between electron and photon showers.

The ν_e CC interactions are concentrated at low conversion distances whilst the backgrounds have a longer tail as expected. However, a significant number of backgrounds are cases where the neutrino interaction vertex has been misplaced and instead the reconstructed vertex is placed at the start of a detached photon, hence the large peak at low conversion gaps . Placing a cut at 3.5 cm removes 11.6% of the ν_e CC signal whilst rejecting 48.6% and 48.7% of ν_μ CC and NC interactions respectively. This corresponds to a reduction in the number of true photon showers of 53.7% whilst reducing the number of misclassified tracks by 21.9%, the number for electrons is the same as for ν_e CC. The aggressiveness of this cut is limited by the resolution with which the shower start position and neutrino vertex are reconstructed.

For a single Minimum Ionising Particle (MIP), in this instance an electron, the energy loss per unit length, the dE/dx is expected to be 1.8 MeV/cm in a LAr TPC. However, if a photon pair produces and creates an e^+e^- pair they will be indistinguishable and the dE/dx of the resultant shower is expected to be double the single MIP expectation. Note that photons which Compton scatter will produce an electron shower with the single MIP dE/dx, as discussed in Chapter 3. The reconstruction of this quantity is discussed in detail in Chapter 6.2.4.

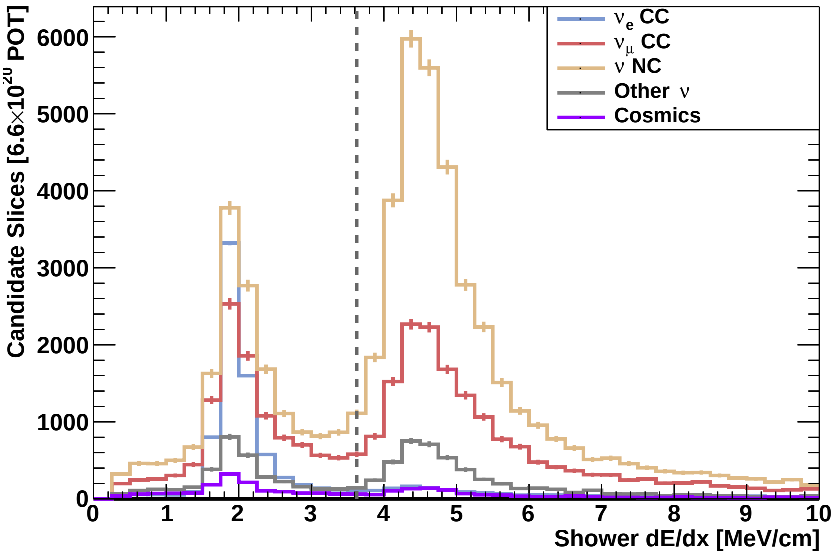


Figure 7.23: Shower initial track dE/dx broken down by the true interaction type of each slice after Dazzle muon candidate removal. Dashed line represents the cut applied at 3.625 MeV/cm.

The distribution of dE/dx is shown in Figure 7.23 which clearly shows both single and double MIP peaks. The ν_e CC interactions are dominated by the single MIP peak whilst the backgrounds show double peaked structures, with NC interactions dominated by the double MIP peak. Placing a cut just before the double MIP peak at 3.625 MeV/cm reduces ν_μ CC and NC backgrounds by 63.4% and 71.1% respectively. This corresponds to a 72.3% reduction in the number of remaining showers that match to true photons and only 49.2% for the misclassified tracks. The lower reduction from tracks is expected as misclassified muon or pion tracks would have the same dE/dx distribution as electron showers, although protons would be rejected. However, 25.8% of remaining ν_e CC signal events are removed largely due to the long tail of the Landau distribution.

Combining both the dE/dx and conversion gap cuts rejects 76.2% and 82.7% of the ν_μ CC and NC events respectively whilst removing 31.6% of the ν_e CC signal. The energy distribution of the remaining candidates is shown in Figure 7.24, demonstrating the dominance of backgrounds at low energy. Although this cut removes a large number of backgrounds they still dominate the remaining sample making up 76.2% of the remaining slices, albeit reduced from 91.0%. Photon induced showers are reduced by 84.6% whilst misclassified showers are reduced by only 63.3%. As such, photons now only make up 57.6%, reduced from 76.9%, of the remaining background showers and misclassified tracks now play a more significant role.

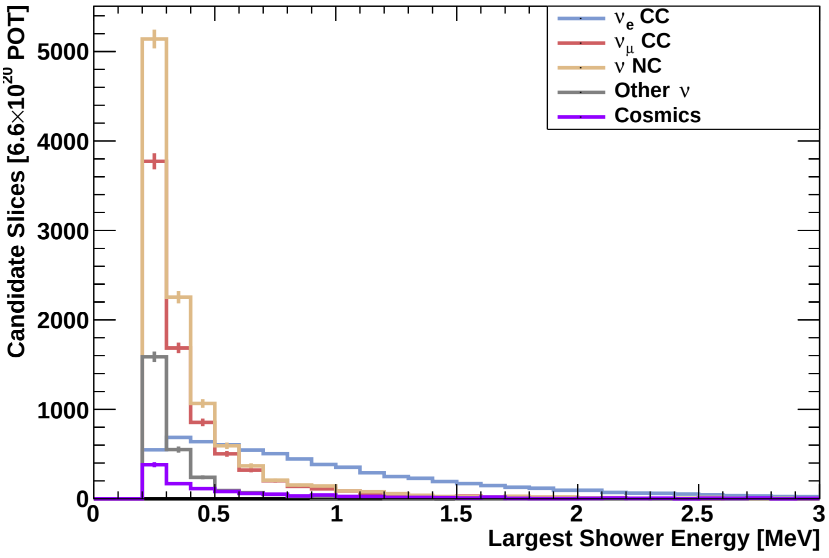


Figure 7.24: Shower reconstructed energy distribution broken down by the true interaction type of each slice after Dazzle muon candidate removal and dE/dx and conversion gap cuts. Dashed line represents the cut applied at 3.625 MeV/cm.

7.5.2 Razzle Shower PID BDT

In order to improve the efficiency and background rejection of the cuts placed on the electron shower candidates another multi-classification Boosted Decision Tree (BDT) was developed as a replacement. This BDT is designed using the same design philosophy as the track PID BDT presented in Section 7.4.3 but classifies each shower as: electron, photon or other.

As with Dazzle, the inputs to Razzle are a combination of the variables previously used to identify showers combined with additional weak predictors, largely aimed at identifying misclassified tracks. Care was taken to avoid, or modify, input variables that have a strong dependence on the shower energy to minimise any biases. Specifically, the chosen inputs for the BDT are listed below, with distributions for each variable shown in Appendix E:

- **Conversion Gap:** The gap between the neutrino interaction vertex and shower start position. Electron showers should have a no gap and photons will travel some distance before interacting and will have a gap as previously discussed.

- **dE/dx:** Differentiate between single and double MIP showers from electrons and pair producing photons as previously discussed.

- **Opening Angle:** The opening angle of the shower, defined as: $\theta_{Open} = \tan^{-1}\left(\frac{\textbf{Shower Width}}{\textbf{Shower Length}}\right)$, can help identify track-like particles that were misclassified as showers.

- **Modified Hit Density:** Calculate the number of hits per wire to identify misclassified track-like particles. This is calculated by dividing the length of the shower by the effective pitch, dx, of the shower, details of each of these are discussed in Chapter 6.2.

- **Energy Density:** Similarly to the hit density but can additionally differentiate between electrons and photons. This density is calculated using the formula $\textbf{Density} = (\frac{\textbf{Shower Energy}}{\textbf{Shower Length}^2})$ to account for correlations between energy and length.

Similarly to Dazzle, either a cut on any of the scores produced can be placed or hypothesis with the highest score can be selected. The distribution of scores is shown in Figure 7.25 with each line representing a true particle type, where a good level of separation is achieved for each particle. This can be also seen in Figure 7.26 which shows a confusion matrix, where the hypothesis with the highest score is used

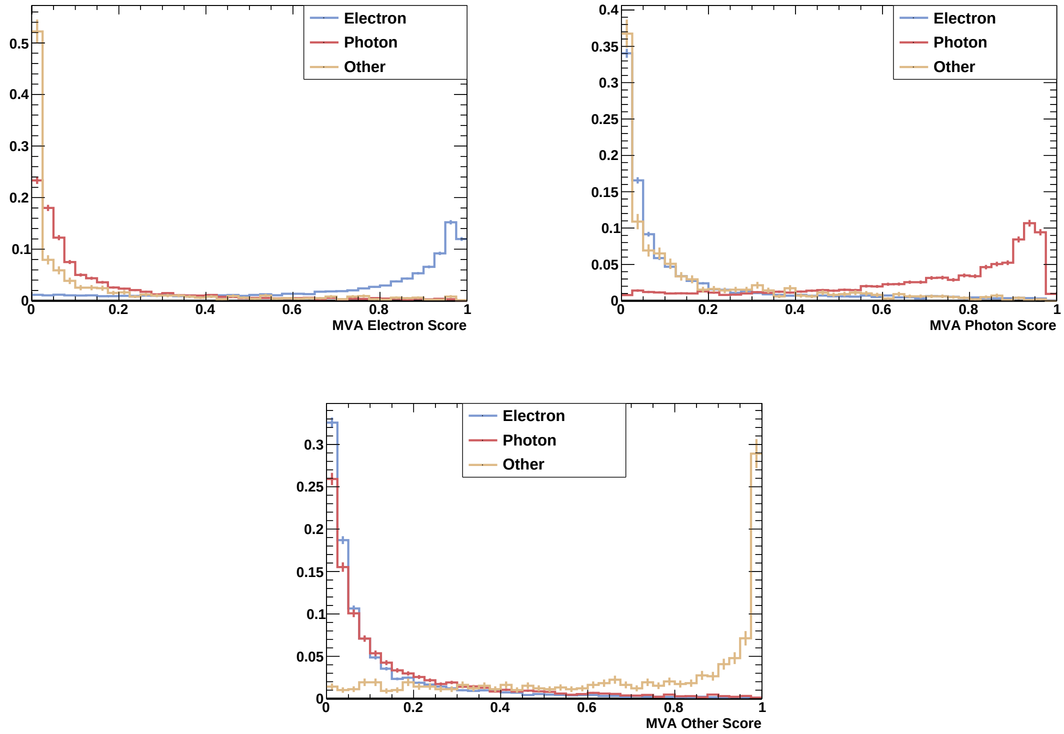
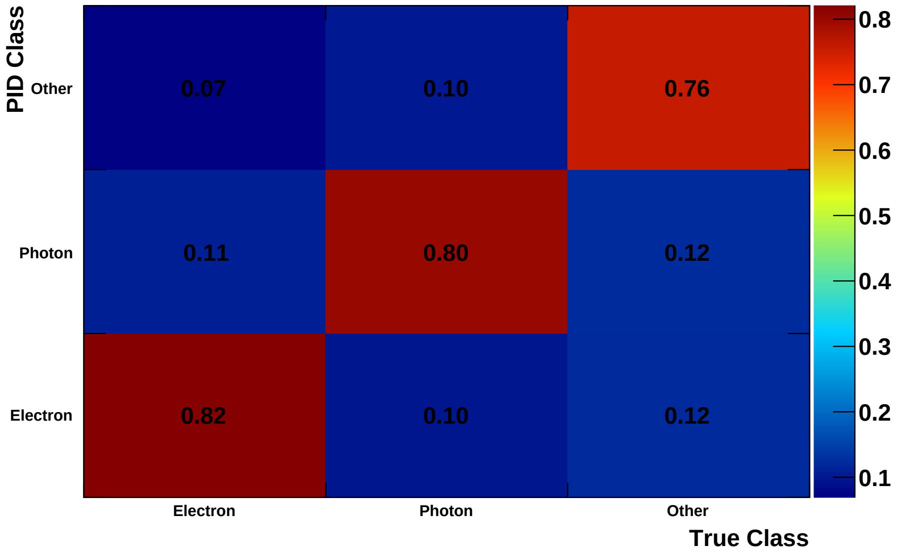


Figure 7.25: Razzle Shower PID BDT scores under electron, top left, photon, top right, and other, bottom , hypotheses. Each line represents showers matching back to the specified true particle and a higher score a higher probability of that hypothesis.



PID Class	Electron	Photon	Other
Other	0.07	0.10	0.76
Photon	0.11	0.80	0.12
Electron	0.82	0.10	0.12

Figure 7.26: Confusion matrix comparing true shower types to the classification from the Razzle shower PID MVA, column normalised to unity.

as the PID classification. Electrons are most often identified correctly, 82% of the time, where photons and other fall slightly lower at 80% and 76% respectively.

In the ν_e CC selection, this BDT is run over the electron candidate for each slice. The distribution of scores for the remaining slices after the muon track cuts is shown in Figure 7.27. Rejecting any slices where the electron candidate has an electron-like score below 0.75 rejects 93.3% of ν_μ CC and 93.9% of NC interactions compared to 76.2% and 82.7% respectively for the combined cut. That is, the remaining backgrounds are reduced by 71.8% and 65.0% for ν_μ CC and NC. This cut keeps 58.7% of the ν_e CC signal, a 14.2% loss over the combined cut. Relative to the combined cuts, only 15.0% more electron showers are rejected compared to 61.8% of photon showers and 75.9% of misclassified tracks.

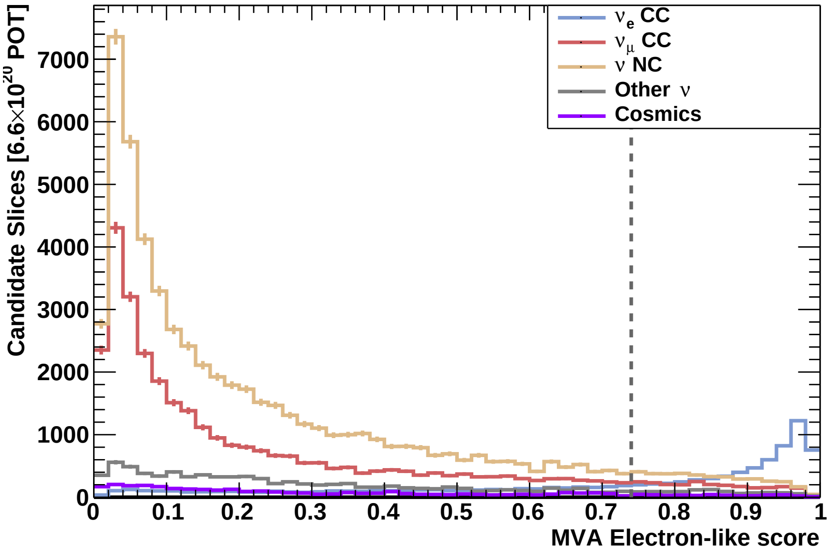


Figure 7.27: Razzle shower PID BDT scores under electron hypotheses broken down by the true interaction type of each slice after Dazzle muon cuts.

7.5.3 Summary

Overall, the cuts to identify electron showers perform well at rejecting the backgrounds whilst maintaining the majority of the ν_e CC signal, the performance of these cuts is shown in Table 7.4. This demonstrates that the performance of the Razzle BDT exceeds that from combining the dE/dx and conversion gap cuts. Due to the prevalence of backgrounds after the muon track removal, harsh cuts are required to remove the backgrounds which dominate over the signal. These cuts are able to reduce 92.3% and 92.9% of the ν_μ CC and ν NC backgrounds respectively. Similarly, cosmic ray and other neutrino backgrounds are reduced by 89.3%. However, these harsh cuts combined with the stochastic nature of showers result in an efficiency loss of the ν_e signal of 41.3%. Nevertheless, the purity is increased to 45.1% from an initial 9.0% after the muon track removal.

The number of slices where the electron candidate was a photon was reduced by 94.1% compared to 90.9% for misclassified tracks. In comparison, the loss of true electron showers was only 43.7% which now constitute 44.4% of the remaining slices, up from 8.6% after the muon track cut. Photon showers now make up 38.0% of the electron candidates with the remaining 17.5% being made up of misclassified tracks, down from 70.2% and 21.1% respectively previously.

True Type:	ν_e CC	ν_μ CC	ν NC	Other ν	Cosmics
dE/dx Selection	74.2	37.7	29.8	40.6	55.3
Conversion Gap Selection	88.4	52.5	52.3	76.2	63.5
Combined Selection	68.4	23.8	17.3	28.9	33.0
Razzle-Dazzle Selection	58.7	6.7	6.1	10.7	10.7

Table 7.4: Selection efficiency for each of the electron shower cuts after the muon track rejection.

7.6 Final Selection Results

The number of selected slices at each stage of the selection is shown in Table 7.5. In total there are 5,960 signal ν_e CC interactions selected compared to 2,240 ν_μ CC and 3,646 NC interactions. There are an additional 376 cosmic induced slices and 988 other ν slices selected, dominated by 496 dirt backgrounds and 438 split neutrino interactions. This yields an overall purity of 45.1%, a dramatic increase compared to the 0.03% before any cuts.

The efficiency for the ν_e CC signal for each stage of the selection can be seen in Table 7.6. There is a 27.1% efficiency loss in the preselection largely due to reconstruction inefficiencies in reconstructing a suitable electron shower candidate,

True Type:	ν_e CC	ν_μ CC	ν NC	Other ν	Cosmics
No Selection	18,860	2,738,666	571,163	3,580,692	62,127,049
Preselection	13,651	194,911	84,708	20,184	111,629
Cosmic Selection	12,161	120,159	68,121	10,911	6,993
Dazzle Selection	10,146	33,319	60,148	9,212	3,524
Razzle-Dazzle Selection	5,960	2,240	3,646	988	376

Table 7.5: Number of candidate slices selected for each stage of the selection, corresponding to an exposure of 6.6×10^{20} POT.

True Type:	ν_e CC	ν_μ CC	ν NC	Other ν	Cosmics
Reconstruction	94.04	94.08	50.07	33.53	-
Preselection	65.83	6.70	7.43	0.11	0.18
Cosmic Selection	58.65	4.13	5.97	0.06	0.01
Dazzle Selection	48.88	1.14	5.27	0.05	0.01
Razzle-Dazzle Selection	28.11	0.08	0.32	0.01	0.00

Table 7.6: Selection efficiency for each stage of the selection.

although this rejects a large number of all backgrounds. The containment requirement within the preselection additionally sacrifices some of the signal efficiency to ensure the reconstruction quality of these showers. The cosmic removal is extremely effective, reducing remaining the cosmic background by 93.7% with only an 10.9% reduction in the remaining signal. Using the track containment cut and Dazzle BDT to identify muon tracks reduced the remaining ν_μ CC backgrounds by 72.4% with only a 16.7% loss to the ν_e CC signal. Finally, the Razzle BDT reduces the residual ν_μ CC and ν NC backgrounds by 92.3% and 92.9% respectively whilst removing 41.3% of the ν_e CC signal. This efficiency loss for the signal is necessitated to reduce the preponderance of backgrounds that dominate even after the Dazzle selection. In particular, the NC backgrounds can only be removed by cuts placed on the showeres observed and outnumber the signal by a factor of 6 before this final cut. This yields an overall efficiency of 28.1%, whilst reducing ν_μ CC backgrounds by 99.9%, ν NC by 99.7%, other ν by 99.99% and cosmic rays by over 99.999%.

The distribution of reconstructed shower energies is shown in Figure 7.28, broken down by the interaction type the slice matches back to in truth. The final purity of the selected sample is 45.1%, a significant improvement over the 0.03% at the start of the selection. Nevertheless, the backgrounds still dominate at low energies which outlines the challenge associated a ν_e selection in SBND. This demonstrates the need for further development to the pattern recognition, reconstruction, and selection tools within SBND. In particular, unlocking this low energy phase space is key for differentiating between interaction models as will be discussed in the proceeding Chapter.

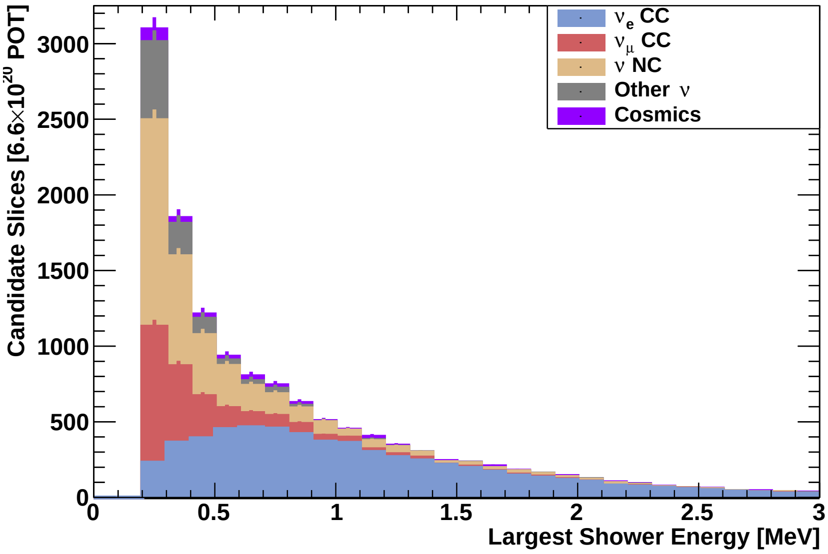


Figure 7.28: Shower reconstructed energy distribution broken down by the true interaction type of each slice after Razzle-Dazzle cuts.

This efficiency is profiled over the true lepton energy in Figure 7.29. The efficiency is lowest at low lepton energies with a gradual rise to a plateau above 1 GeV at an efficiency of 40%. The Razzle BDT disproportionately removes low energy showers as these are most commonly mistaken with backgrounds. However, the other cuts provide a largely flat reduction in signal across the lepton energy.

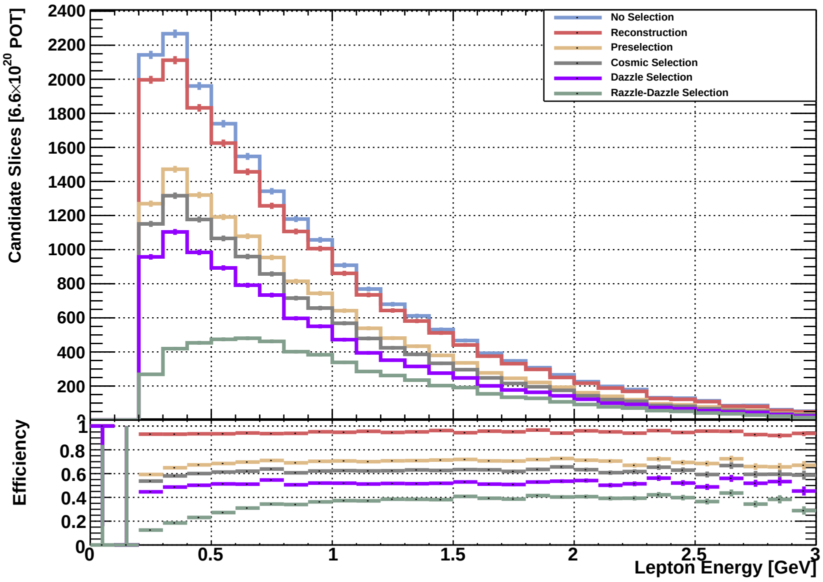


Figure 7.29: Selection efficiency of ν_e CC interactions for each stage of the selection as a function of true lepton energy.

7.7 Concluding Remarks

Despite the ν_e CC signal making up only 0.5% of the BNB, the final selected sample achieves a purity of 45.1%, with almost all of the cosmic induced backgrounds being rejected. The ν_μ CC and NC backgrounds are reduced by 99.9% and 99.7% respectively whilst maintaining a signal efficiency of 28.1%. The selected events can now be used to perform a measurement of the ν_e CC inclusive cross section, as will be presented in the forthcoming chapter.

Chapter 8

$\nu_e + \overline{\nu}_e$ Charged Current Inclusive Cross Section Measurement

The ultimate goal of this thesis is to assess the ability of the Short-Baseline Near Detector (SBND) to make a measurement of the $\nu_e + \overline{\nu}_e$ Charged Current (CC) inclusive cross section. In particular, to quantify the ability to differentiate between theoretical cross section model predictions. As such, this chapter represents the culmination of work presented in the preceding chapters to reconstruct and select these interactions.

First, a brief outline of the analysis procedure will be given in Section 8.1, with an overview of the, simulated, input data used in the analysis. Next, the formalism for the cross section calculation will be presented in Section 8.2, outlining the individual steps undertaken. A discussion of the systematic uncertainties is subsequently given in Section 8.3, outlining the techniques used, formalism, and sources of uncertainty included in the final measurement. The measurement of the cross section is then presented in Section 8.4, stepping through each stage of the process. A comparison of this measurement to neutrino interaction models is subsequently presented in Section 8.5 which quantifies the ability of SBND to differentiate between the models. Finally, the obtained results and future improvements that could improve both the realism and power of this measurement are discussed in Section 8.6.

8.1 Introduction

Due to the absence of any real data from SBND, simulation has been used to create “fake data” for which a cross section can be measured and uncertainties assessed. This simulated sample consisted of three individual samples:

- **BNB:** This simulates all neutrino interactions associated with the Booster Neutrino Beam (BNB), including dirt interactions, overlaid with CORSIKA simulated cosmics. This sample contains 3,560,488 events corresponding to 3.08×10^{20} Protons On Target (POT).

- **Intrinsic ν_e:** This sample runs the BNB simulation, and cosmic overlay, as above except only keeps the intrinsic ν_e and $\overline{\nu}_e$ from the beam, discarding ν_μ. This sample contains 174,600 events corresponding to 2.47×10^{21} POT.
- **In-time Cosmics:** CORSIKA cosmics were simulated in the absence of neutrino interactions to mimic events that would be triggered by an in-time cosmic interaction. This sample contains 417,245 events corresponding to 37,475,000 beam spills, or 1.87×10^{20} POT assuming a nominal intensity of 5×10^{12} POT per beam spill.

These samples were all scaled to the expected exposure of 6.6×10^{20} Protons On Target (POT), corresponding to three years of running for SBND, with the statistical errors scaled accordingly. The intrinsic ν_e sample was used to model the reconstruction and selection of the signal, used in calculating the forward folding matrix and efficiency. The BNB sample, with the signal removed, was combined with the in-time cosmics to assess the background contribution to the selected sample.

The true cross section presented by a theoretical model, $\sigma(E)$ depends on the energy of the incoming neutrino. Therefore, the rate of neutrino interactions observed in the detector, r, depends on the flux of incoming neutrinos ϕ and the cross section:

$$r(E) = \phi(E) \cdot \sigma(E) \cdot n_t \tag{8.1}$$

The rate of interaction scales directly with the number of nuclear targets in the detector, n_t. If a detector was able to faithfully reconstruct the neutrino energy the measured cross section could be presented in terms of neutrino energy and the flux dependence on the observed rate could be removed. However, due to nuclear effects, Final State Interactions (FSI), and imperfect detector performance it is not possible to accurately reconstruct the incoming neutrino energy from the observed particles in the detector without introducing model dependent assumptions. The introduction of any such assumptions could lead to biases and thus degrade the ability of a measurement to differentiate between models, any such model dependencies should be minimised. Instead, the cross section that is measured, whether integrated or differential in a variable other than neutrino energy, must be calculated from the energy-integrated number of observed interactions, which scales directly with the exposure of Protons On Target (POT).:

$$N = n_t \cdot POT \int [\phi(E) \cdot \sigma(E)] dE \tag{8.2}$$

This integration conflates the neutrino flux and cross section models and thus a cross section measured from this rate is dependent on the neutrino flux, referred to as being "flux-integrated". This flux-integrated cross section, σ^{FI}, is thus calculated from this rate correcting for the integrated flux, Φ, as described in Equation 8.3:

$$\begin{aligned}\sigma^{FI} &= \frac{N}{\Phi \cdot n_t} = \frac{1}{\Phi}\int[\phi(E)\cdot\sigma(E)]dE,\\ \Phi &= POT\int\phi(E)dE\end{aligned} \tag{8.3}$$

Therefore, to compare the measured cross section to a theoretical model a prediction must be made which convolves the true cross section with the flux. This method provides the most model independent measurement of a cross section, but adds complexity to external comparisons. In particular, measurements performed by experiments that are exposed to different fluxes do not produce directly comparable results.

8.2 Cross Section Extraction Procedure

When taking into account imperfect reconstruction and selection, the observed event rate N is:

$$N = \sigma \cdot \epsilon \cdot n_t \cdot \Phi + B \tag{8.4}$$

Where σ represents the flux-integrated cross section, ϵ the efficiency, n_t the number of target nucleons, Φ the integrated flux, and B the number of selected background events. The efficiency ϵ is defined for only the signal events, as the backgrounds are handled separately, and is given in Equation 8.5:

$$\epsilon = \frac{N^{Signal}_{Selected}}{N^{Signal}_{Total}} \tag{8.5}$$

In order to measure the cross section from these events Equation 8.4 can be rearranged to obtain Equation 8.6:

$$\sigma = \frac{N - B}{\epsilon \cdot n_t \cdot \Phi} \tag{8.6}$$

This integrated cross section can be measured but provides only limited information to differentiate between interaction models. Therefore, differential cross sections are typically preferred as they provide additional handles to improve the discriminating power between models. For a differential cross section in some arbitrary observed parameter X the number of observed events in bin i, N_i, is measured and the cross section for the corresponding bin is calculated in Equation 8.7:

$$\left(\frac{d\sigma}{dX}\right)_i = \frac{N_i - B_i}{\epsilon_i \cdot n_t \cdot \Phi \cdot (\Delta X)_i} \tag{8.7}$$

The number of backgrounds, B_i, and efficiency, ϵ_i, are the same as previously defined but are now evaluated in each bin. There is an additional term to incorporate the bin width $(\Delta X)_i$ which allows non-uniform binning to be used, such that a target statistical error in each bin can be achieved.

As only the observed number of interactions, N, is measured by the detector the remainder of the parameters must be estimated, typically from Monte Carlo (MC) simulations. There will be a systematic uncertainty associated with these estimations due to modelling of the detector, flux, and neutrino interactions. The measurement of side-bands, for example π^0 interactions that pose a major background in this analysis, can be used to provide an in-situ constraint and reduce these uncertainties, although that is not considered in this thesis[68].

In order to assess the uncertainties, a series of universes are created each of which contain some systematic variation, most commonly via reweighting which will be discussed in Section 8.3. The extracted cross section can then be calculated in each universe, denoted by ^{n}, as described in Equation 8.8.

$$\left(\frac{d\sigma}{dX}\right)_i^n = \frac{N_i - B_i^n}{\epsilon_i^n \cdot n_t \cdot \Phi^n \cdot (\Delta X)_i} \tag{8.8}$$

The parameters denoted by ^{n} indicate those that are effected by the systematic variations and thus assessed in each universe, whilst those without remain constant across the universes. Once this cross section is calculated in every universe, the information can be distilled into a covariance matrix which encodes the uncertainty in the measurement, including correlations between bins, as will be discussed in Section 8.3.1.

Once the cross section measurement has been performed comparison to different interaction model predictions can be made. The uncertainties encoded in the covariance matrix allows for a quantitative assessment of the level of agreement between the measurement and the predictions. This basic prescription of cross section extraction will be further expanded upon in the forthcoming sections, detailing the estimation of each of the parameters and discussing the sources and formalism of systematic uncertainties.

8.2.1 Background Subtraction and Purity Correction

In order to recover the true amount of signal, backgrounds must be removed from the observed number of interactions. These backgrounds can be separated into two distinct categories: on-beam and off-beam. The former represent backgrounds that are associated with the neutrino beam activity including ν_μ Charged Current (CC), ν Neutral Current (NC), and dirt interactions. The latter represent backgrounds that are not associated with the beam, namely cosmogenic backgrounds.

Off-beam backgrounds can be estimated by taking data in anti-coincidence with the beam. This in-situ constraint removes any systematic errors associated with the modelling of the background source, or indeed the detector itself. As such, the off-beam backgrounds have only a statistical uncertainty associated with them.

In contrast, the on-beam backgrounds cannot be separated out from the signal in the same way and thus MC simulations or side-bands must be used to estimate the expected rate, although only the former is explored here. Therefore, there will be associated systematic uncertainties corresponding to the modelling of these backgrounds.

In addition to the Poissonian uncertainty on the number of background interactions selected, there is an additional uncertainty due to finite statistics. That is, the knowledge of the true number of backgrounds is limited by the finite number of background interactions modelled, either simulation or anti-coincidence data for on-beam and off-beam respectively. This uncertainty is typically negated by generating significantly more backgrounds than are expected from the data, resulting in this error diminishing when added in quadrature with the Poissonian uncertainty.

Two ways to remove backgrounds from the data sample have been considered in this work: either via a background subtraction, as in Equation 8.7, or by a purity correction. The former predicts the absolute number of backgrounds that are expected that can thus be subtracted from the data rate. This background subtraction is a more model insensitive technique as it includes no correlations between how the systematic uncertainties will effect the signal and background. However, in the presence of large normalisation uncertainties and impure selections the impact of these uncertainties can be extremely large.

Rather than predicting the number of backgrounds directly, the purity of a given bin can be predicted from the number of signal, S, and background, B, in each universe, n, as in Equation 8.9:

$$P_i^n = \frac{S_i^n}{S_i^n + B_i^n} \tag{8.9}$$

This purity takes into account correlations between how the systematic variations affect the signal and background. This reduces the impact of normalisation uncertainties as the affect on the signal and background will largely cancel, particularly in cases of low purity. However, by including the correlations between signal and background there is a risk that some model dependent based assumptions can be baked into the measurement. As such, care must be taken to ensure that the systematic variations used to evaluate the uncertainties on the purity are sufficient to cover any potential model dependencies.

Once the purity has been estimated, a correction can then be applied to the data rate and a cross section extracted as in Equation 8.10:

$$\left(\frac{d\sigma}{dX}\right)^n_i = \frac{N_i \cdot P^n_i}{\epsilon^n_i \cdot n_t \cdot \Phi^n \cdot (\Delta X)_i} \tag{8.10}$$

Both options will be assessed in this thesis so the impact of the systematic uncertainties can be quantified.

8.2.2 Normalisation

The number of target nucleons, n_t, can be derived from the detector mass, M_{det}, the atomic mass of argon, M_{Ar}, and nucleons per argon atom, N_{Ar}, as described in Equation 8.11:

$$n_t = \frac{N_{Ar} \cdot M_{det}}{M_{Ar}} = \frac{40 \cdot V_{det} \cdot \rho}{39.95 \cdot AMU} \tag{8.11}$$

The mass of the detector is calculated from the Fiducial Volume (FV), V_{det}, defined in Chapter 7.2.2 multiplied by the density of liquid argon ρ[125]. The detector is assumed to comprise entirely of argon 40, with impurities comprising a negligible fraction of the total detector mass, with a mass of 39.948 Atomic Mass Units (AMU)[125]. Therefore the total number of target nucleons is calculated in Equation 8.12:

$$n_t = \frac{40 \cdot 5.67 \times 10^7 \cdot 1.40 \times 10^{-3}}{39.95 \cdot 1.66 \times 10^{-27}} = 4.78 \times 10^{31} \tag{8.12}$$

Uncertainties in the number of target nucleons can be introduced by lack of knowledge of the true detector mass. In particular, the Space Charge Effect (SCE) will shrink the detector volume causing an underestimation of the true mass as discussed in Chapter 3. Although this effect is corrected for in the current simulation, imperfections in this correction procedure will lead to some residual uncertainty. At the time of writing only the nominal SCE is simulated and no systematic variations are available, thus this uncertainty is not included in this analysis.

The total flux is calculated using the simulation discussed in Section 4, scaled to the expected 6.6×10^{20} POT[2]. The total flux through the face of the Fiducial

Volume (FV) defined in Chapter 7.2.2 is shown in Figure 8.1, separated for ν_e and $\overline{\nu}_e$. This shows that the ν_e dominates over the $\overline{\nu}_e$ as discussed in Chapter 4. The total integrated flux is 5.95×10^{10} ν_e and 6.68×10^9 $\overline{\nu}_e$ for a total of 6.61×10^{10} $\nu_e + \overline{\nu}_e$. There will be an overall normalisation uncertainty associated with this integrated flux due to systematic uncertainties in the flux prediction, as will be discussed in Section 8.3.

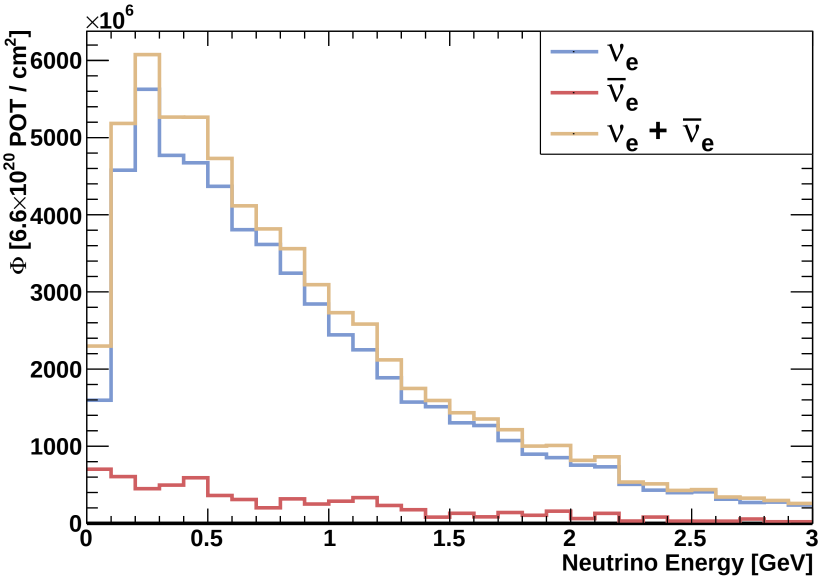


Figure 8.1: Prediction of the ν_e (blue), $\overline{\nu}_e$ (red), and combined (yellow) flux at the front face of SBND FV as a function of true neutrino energy.

8.2.3 Folded and Unfolded Measurements

The previously outlined procedure assumed that a differential cross section measurement would be made in some observable, reconstructed parameter X. However, cross section predictions produced by event generators will typically be presented in terms of the true kinematics of the interaction. In order to compare the measured cross section and a generator prediction one must either “forward fold” the generator prediction into reconstructed space or “unfold” the measured cross section into true space.

The relationship between an observed quantity ν and a true quantity μ is given in Equation 8.13:

$$\nu_i = \sum_j R_{ij} \mu_j \tag{8.13}$$

Where the forward folding matrix R_{ij} describes the probability to observe an event in reconstructed bin i given that it was generated in true bin j. This matrix encapsulates smearing that occurs due to imperfect detector resolution and reconstruction but not efficiency losses, as these will be corrected elsewhere. As such, the matrix is column normalised so every true interaction will be included in the reconstructed prediction. The underflow and overflow bins are included to ensure that events that would be reconstructed but excluded from the final measurement are correctly handled. The off diagonal terms describe the amount of smearing between bins that will occur, in turn smearing out features of the generator predictions and making it harder to effectively differentiate between models.

Alternatively, the forward folding matrix given in Equation 8.13 can be inverted to obtain the true distribution in terms of the reconstructed:

$$\mu_j = \sum_i R_{ij}^{-1} \nu_i \tag{8.14}$$

This unfolding allows the measured cross section to be compared directly to the generator predictions. Whilst these methods are mathematically equivalent, statistical fluctuations in the measured rate ν_i can cause catastrophic instabilities in the unfolding procedure[177]. Whilst regularisation can alleviate these issues, it also carries the risk of baking in model specific assumptions to the unfolding procedure. This can, in turn, limit the ability of the experiment to differentiate between generator predictions.

One advantage of unfolded cross section measurements is that it removes any detector specific smearing and inefficiencies which can inhibit comparisons of data between experiments. However, as the measured cross section is flux-integrated it is not possible to directly compare the measurements to those performed by other experiments which experience a different flux. Due to the proximity to the neutrino beam source, each of the Short-Baseline Neutrino (SBN) detectors experiences a different flux from the BNB and thus measurements are not directly comparable. This also applies to measurements made in other neutrino beams.

As such, the most general solution is to present a forwards folded measurement in reconstructed space. The forward folding matrix can additionally be provided to ensure that future theoretical predictions can be compared to the data. The predicted cross section from an event generator in some true parameter Y can then be folded into reconstructed space X where the cross section can be calculated, as described in Equation 8.15:

$$\left(\frac{d\sigma}{dX}\right)_i = \sum_j R_{ij} \frac{N_j}{n_t \cdot \Phi \cdot (\Delta Y)_j} \tag{8.15}$$

This forward folded prediction can then be compared to the measured data to assess the agreement with the theoretical model. Model dependent assumptions can create a systematic error on the forward folding matrix, as will be discussed in Section 8.4.

8.3 Systematic Uncertainties

In order to assess the impact of systematic uncertainties on a physics measurement, the simplest approach is to simulate a number of different samples each with a given parameter tweaked within its uncertainty range. The physics measurement can then be performed in each of these universes and any biases or smearing across the universes describes the uncertainty that the tweaked parameter has on that measurement. However, fully simulating and reconstructing large numbers of events for each systematic parameter throw is computationally expensive. Therefore, a technique referred to as reweighting can be used to apply a weight to each event based on some tweaked input parameter and recover the desired distribution, saving significant computational expense[177, 178].

Reweighting begins with some input physics parameter P and transforms it into a modified version P' as described by Equation 8.16:

$$P \to P' = P\left(1 + x_p \cdot \frac{\delta P}{P}\right) \tag{8.16}$$

The standard deviation of P is described by δP and thus x_p describes how many standard deviations the parameter is to be shifted, with $x_p = 0$ recovering the unweighted parameter. In the case of neutrino interactions, the probability for an interaction to occur is given by the cross section $\sigma(P)$. As such, a neutrino interaction can be assigned a weight W that describes whether it is more or less likely to occur based on some tweak of x_p to an input parameter P:

$$W = \frac{\sigma'}{\sigma} = \frac{\sigma(P')}{\sigma(P)} \tag{8.17}$$

By applying this weight to every neutrino interaction, it is possible to produce a distribution of events as if they were generated with the given input parameter tweak, providing the sample has sufficient statistics and phase space coverage.

In order to assess the impact of a given uncertainty on a measurement made by an experiment, a series of “universes” are generated where a given input parameter is thrown from a unit Gaussian, where both $\mu = 1$ and $\sigma = 1$. The weight for each interaction is calculated in each universe and the spread, and any biases, of the physics measurement made across the universes are then used to quantify the uncertainty that the given parameters have on the measurement, as will be formalised in Section 8.3.1. If multiple systematic parameters are thrown, the product of weights in a universe can be used to assess the combined impact of all of the parameters, effectively adding the uncertainties in quadrature. This technique reduces the computational requirements for the assessment of the impact of the systematic uncertainties.

However, the systematic parameter tweak may not always be analytically convertable to a probability, as discussed above. In some cases it may still be possible to apply reweighting; for example in the case of hadronic scattering within the nucleus where a parameter tweak will impact the Mean Free Path (MFP) λ as[178]:

$$\lambda \rightarrow \lambda' = \lambda \left(1 + x_{mfp} \cdot \frac{\delta\lambda}{\lambda}\right) \tag{8.18}$$

Unlike in the case of cross sections, this MFP cannot be directly interpreted as a probability for a given interaction to occur. The probability for the hadron to interact within the nucleus P is calculated from the MFP and the distance that it must travel before exiting the nucleus. Therefore, the overall weight given to the neutrino interaction is calculated based on whether the hadron interacts or survives in the nominal simulation:

$$W = \begin{cases} \frac{1-P'}{1-P} & \text{for interacting hadrons} \\ \frac{P'}{P} & \text{for surviving hadrons} \end{cases} \tag{8.19}$$

In other cases it may not be possible to reweight the event and thus the simulation and reconstruction must be rerun. For example, when modelling detector effects such as the wire noise level it may be feasible, but non-trivial, to determine whether a given energy deposition would be reconstructed as a hit. However the effect that removing a single hit would have on the pattern recognition, high level reconstruction, and Particle IDentification (PID) is not analytically calculable. Therefore the downstream reconstruction would still have to be run for each systematic variation, adding significant computational expense.

8.3.1 Error Propagation Formulation

In order to asses the impact of uncertainties a covariance matrix is constructed from a series of observations N, as described in Equation 8.20[177]:

$$V_{ij} = \frac{1}{U}\sum_{n}^{U}(N_i^n - N_i^{CV})(N_j^n - N_j^{CV}) \tag{8.20}$$

This matrix encodes the average deviation between the value in bins i and j from the Cental Value (CV) across U universes, each denoted by ^{n}. The diagonal terms represent the variance in a given bin and thus are used to derive the errors shown in the forthcoming sections:

$$V_{ii} = \sigma_i^2 \tag{8.21}$$

The off-diagonal elements are used to encode the correlations between bins, and are used when calculating a χ^2 between the data and a prediction:

$$\chi^2 = \sum_{ij}(N_i^{Data} - N_i^{Pred})V_{ij}^{-1}(N_j^{Data} - N_j^{Pred}) \tag{8.22}$$

This assumes that the CV, computed as the average across the universes, and the nominal value used in a simulation are consistent. Whilst this is generally true, this is not the case for the flux uncertainties, as will be detailed later in this section. As such, the total error matrix E_{ij} can be decomposed into separate bias and resolution matrices, where the resolution matrix is the standard covariance matrix:

$$\begin{aligned} E_{ij} &= V_{ij}^{Bias} + V_{ij}^{Res}, \\ V_{ij}^{Bias} &= (N_i^{Nom} - N_i^{CV})(N_j^{Nom} - N_j^{CV}), \\ V_{ij}^{Res} &= \frac{1}{U}\sum_{n}^{U}(N_i^{n} - N_i^{CV})(N_j^{n} - N_j^{CV}) \end{aligned} \tag{8.23}$$

The two covariance matrices V_{ij}^{Bias} and V_{ij}^{Res} represent the bias between the nominal and CV and the variance around the CV across the universes respectively. This decomposition separates the effect of biases and resolutions so that their impacts can be assessed individually.

The total error matrix is computed by adding the individual covariance matrices from the individual sources together, effectively adding the errors in quadrature:

$$E_{ij}^{Tot} = \sum_{Sources} E_{ij}^{Source} = E_{ij}^{Stat} + E_{ij}^{Genie} + E_{ij}^{Flux} + ... \tag{8.24}$$

In order to aid in the visualisation of the uncertainties it is useful to show the correlation matrix which shows the relative level of correlation between bins:

$$Corr_{ij} = \frac{V_{ij}}{\sqrt{V_{ii}V_{jj}}} = \frac{V_{ij}}{\sigma_i\sigma_j} \tag{8.25}$$

Alternatively, the fractional covariance matrix shows the relative error in each bin, making it easier to compare across bins with different numbers of entries:

$$V_{ij}^{Frac} = \frac{V_{ij}}{N_i^{Nom}N_j^{Nom}} \tag{8.26}$$

8.3.2 Systematic Error Sources

There are three primary sources of systematic errors that will effect measurements made by SBND: detector, flux, and interaction. The impact of both the flux and interaction systematics can be by measured by using reweighting techniques, provided by MiniBooNE and GENIE respectively[2, 178]. However, as the detector systematics are not reweightable, these must be measured by simulating separate samples, as previously discussed. At the time of writing, SBND does not have the computational or personnel resources to perform multiple versions of a production. Moreover, as the SBND simulation chain is yet to utilise the most up to date detector simulation models developed in other Liquid Argon Time Projection Chambers (LAr TPCs) such a production campaign is difficult to justify. Therefore, the detector systematics are not included in the measurement presented here, but should be included in future iterations of this work.

The uncertainties arising from the detector systematics typically arise from imperfections in the calibration of the detector and and mismodeling of the underlying physics models. Mismodelling of, for example, the drift velocity, diffusion, or recombination can result in discrepancies between simulation and data, resulting in an uncertainty to encapsulate the disagreement. However, this uncertainty can be reduced by in-situ measurements of these processes which, when integrated into the simulation, reduce the tensions. These residual differences can be assessed by looking at side-bands which are insensitive to the signal that is being measured, most commonly cosmics.

The purity of the argon must be calibrated to understand both the charge attenuation and quenching of light in the detector[179]. Similarly, the uniformity of the electric field in the detector must be understood, as this affects both the spatial positions and recombination of energy depositions, the SCE is expected to be the dominant contribution[146]. Finally, the detector response, both the wire readouts for the charge and the Photon Detection System (PDS) for the light, must be characterised and understood, this is expected to be the largest contribution to the detector uncertainties[180]. This includes understanding the noise, electronics response, and gain for each channel individually, in addition to the removal of any defective channels.

The flux simulation developed by MiniBooNE provides not only a nominal prediction but a reweighting framework to assess the uncertainties[2]. The systematic uncertainties used are described in detail in Appendix F.1 and are outlined below:

- **Proton Delivery:** The number of protons delivered to the target is measured by a pair of toroids, which have an associated 2% calibration uncertainty that effects the overall flux normalisation.

- **Particle Production:** The number of particles produced in the target has an uncertainty for each particle type:
  - $\pi^{\pm}$: Pion production in the target is modelled using the Sanford-Wang (SW) parameterisation tuned to HAdron Production Experiment (HARP) and E910 data[2]. Where there is poor agreement between the SW fit and data, fitted splines are instead used.
  - $\mathbf{K^+}$: Due to a lack of available data, Feynman Scaling (FS) is employed to translate data from higher energies to the relevant BNB energy.
  - $\mathbf{K^0}$: Similar to the pions, as SW parameterisation is fitted to E910 and KEK data[2].
  - $\mathbf{K^-}$: Due to a lack of available data and the small contribution to the neutrino flux, only an overall normalisation uncertainty is applied.
- **Hadronic Interactions:** Hadrons produced in the target may interact, either elastically or inelastically, before leaving the target, effecting the kinematics and number of neutrinos produced respectively. The uncertainty in these interaction cross sections is propagated through to the neutrino flux prediction.
- **Horn Magnetic Field:** Uncertainties in the magnetic field of the horn impact the focusing of charged particles produced within the target. Uncertainties in both the current pulsed through the horn and the "skin current" induced on the surface of the target are modelled.

The impact of these uncertainties to the flux prediction can be seen in Figure 8.2 which shows the predicted flux of $\nu_e + \overline{\nu}_e$ in SBND. This compares the nominal flux prediction without uncertainties, red, to the flux prediction from 1,000 systematic universes, blue, where all parameters are randomly thrown. The blue line represents the CV and the error bars the standard deviation across the universes, calculated via a covariance matrix. The fractional uncertainty in across the universes, referred to as the resolution, is additionally shown on the bottom of the plot in blue. Typically, the nominal simulation and CV of the universes would be consistent, as is the case at high energy. However, due to discrepancies between the SW parameterisation used to predict the nominal flux and the fitted splines used in the universes, this is not the case for the π^+ uncertainty[2]. As such, an additional bias term is included in the uncertainties, shown in red on the bottom plot, which is added in quadrature with the resolution to yield the total systematic uncertainty, shown in yellow.

The impact of this uncertainty on the predicted event rate in SBND is shown in Figure 8.3. Due to the signal definition requiring the electron energy above 200 MeV, as defined in Chapter 7, the bias in the flux prediction has a relatively small impact

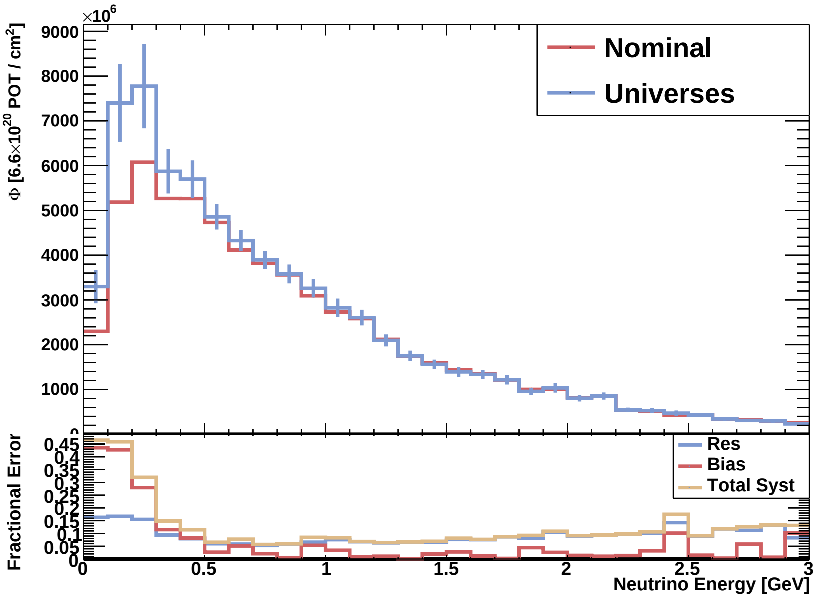


Figure 8.2: Prediction of the combined ν_e and $\overline{\nu}_e$ flux at the front face of the SBND FV for the nominal simulation (red) and the reweighted universes (blue). The fractional uncertainty across the universes is shown at the bottom, split into the resolution across the universes and the bias between the CV and the nominal predictions, added in quadrature for the total.

.

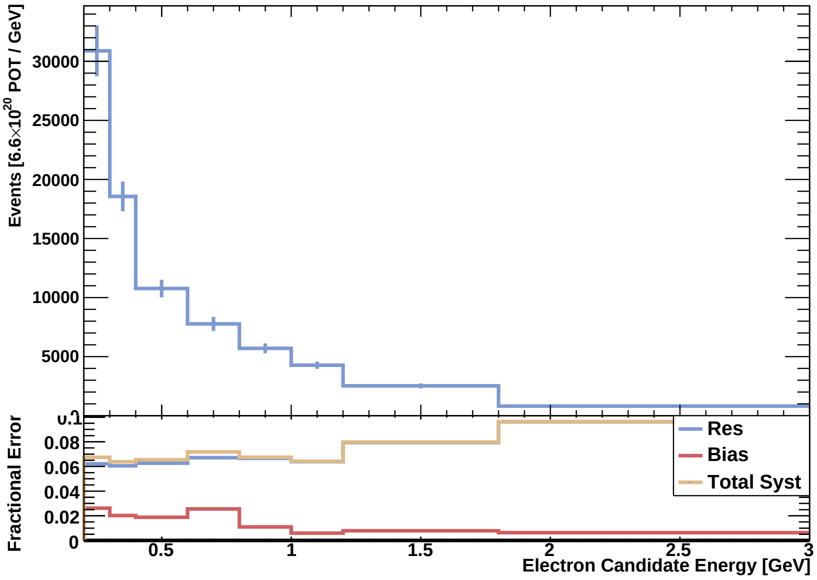


Figure 8.3: Selected event rate distribution with the systematic errors from the flux uncertainties. The fractional uncertainty across the universes is shown at the bottom, split into the resolution across the universes and the bias between the CV and the nominal predictions, added in quadrature for the total.

on the event rate, as the bias is largely concentrated at low neutrino energy. Instead, the event rate is dominated by the resolution across the universes and, as such, is largest at high energy. Therefore, the total uncertainty on the event rate from the flux uncertainty is 6.4%, with 1.8% and 6.2% contributions from the bias and resolution terms respectively.

The integrated flux used to normalise the cross section measurement will similarly be effected by the systematic uncertainties. The distribution of the integrated flux across the 1,000 universes is shown in blue in Figure 8.4, with the nominal prediction indicated by the red line. The bias between the SW parameterisation used in the nominal simulation and splines used in the universes is clearly visible as as the offset between the histogram centre and the red line. There is a 9.6% uncertainty corresponding to this bias which is added in quadrature with the 6.4% resolution uncertainty to yield an overall 11.6% flux normalisation uncertainty. This normalisation uncertainty is particularly important as it effects all of the bins simultaneously, shifting the entire measured cross section.

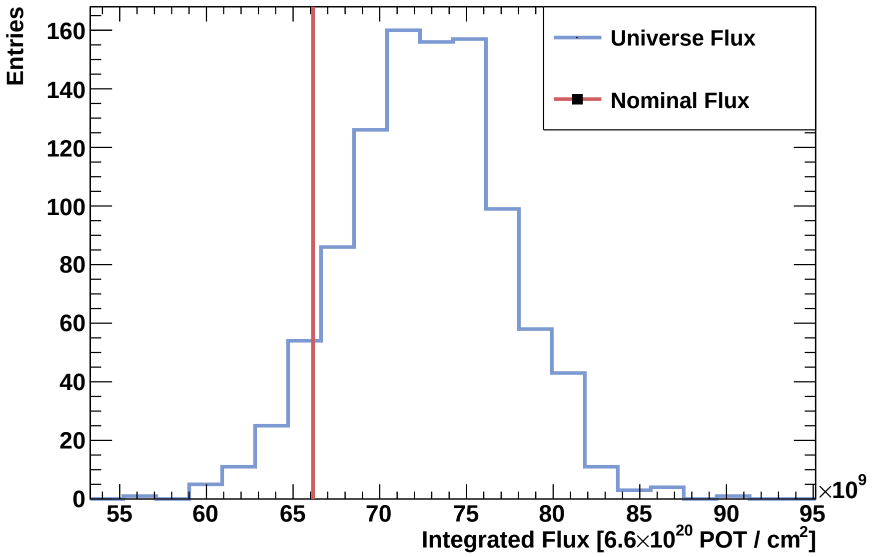


Figure 8.4: Integrated flux through the SBND fiducial volume face for 6.6×10^{20} POT for both the nominal simulation (red) and across the reweighted universes (blue).

Similarly to the flux simulation, the GENIE generator provides a series of tuneable knobs and a framework in which to reweight them[178]. The systematic uncertainties used are described in detail in Appendix F.2 and are outlined below:

- **$\mathbf{M_a}$:** Axial mass is varied for CC Quasi-Elastic (QE), CC Resonant Scattering (RES), NC Elastic, and NC RES interactions independently.

- **$\mathbf{M_v}$:** Vector mass is varied for CC and NC RES interactions independently.
- **$\mathbf{M_\eta}$:** Strange axial form factor η forNC Elastic.
- **NonRES Backgrounds:** Non-resonant backgrounds are varied for CC and NC, ν and $\overline{\nu}$, neutron and proton, and 1π and 2π interactions for a total of 16 systematics.
- **BY Twist:** AHT and BHT parameters in Bodek-Yang (BY) model scaling.
- **CV:** CV1u valence GRV98 PDF correction in BY model.
- **Hadron Transport MFP:** The Mean Free Path (MFP) is varied for both pions and nucleons.
- **Hadron Transport Interactions:** The absorption, charge exchange, and inelastic interaction cross is varied for both pions and nucleons.

The effect of the interaction systematics on the event rate can be seen in Figure 8.5. Unlike in the case of the flux uncertainties, the bias between the nominal simulation and CV across the universes is negligible and the uncertainties are dominated by the resolution term. The uncertainties are largest at low energy but remain significant across the entire energy range. The uncertainty on the integrated rate is 8.9%, with

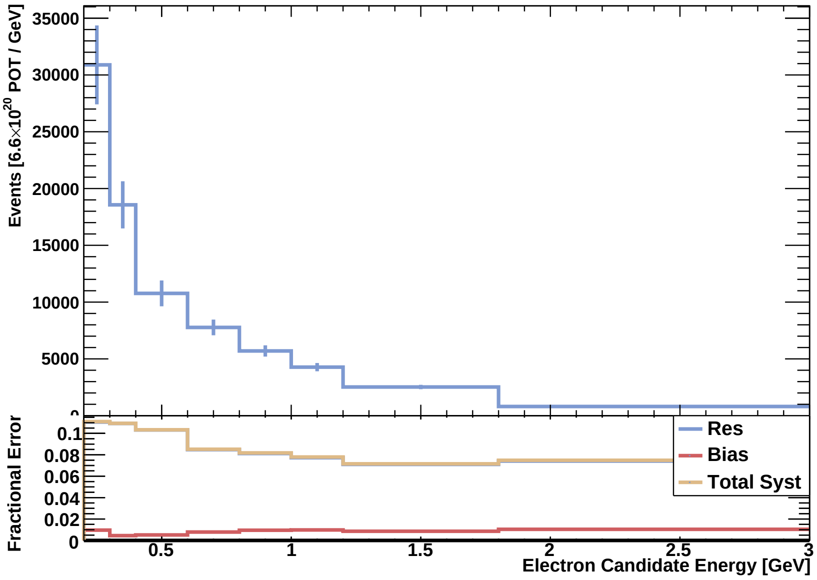


Figure 8.5: Selected event rate distribution with the systematic errors from the interaction uncertainties. The fractional uncertainty across the universes is shown at the bottom, split into the resolution across the universes and the bias between the CV and the nominal predictions, added in quadrature for the total.

only a 0.2% contribution from the bias term and the remaining 8.8% coming from the resolution. Fractional covariance matrices are shown in Appendix G for the flux, interaction, and total uncertainties.

Another potential source of uncertainty is finite simulation statistics, which can impact the efficiency and background estimations. The Poissonian uncertainty in the simulation limits the resolution of the estimates that are derived from the simulated samples. Experiments typically simulate multiple times the number of expected data events such that this uncertainty is below that statistical uncertainty in the data. Due to production constraints, the simulation used for this analysis only has a fraction of the total expected events compared to the full exposure, let alone the final simulation targets. Therefore, an attempt to measure the impact of finite statistics uncertainty would represent the uncertainty on the current statistical sample, not that of the ultimate sample that would be used once SBND has data. As such, any such attempts would have limited utility and thus this uncertainty was not included in this analysis.

Whilst the impact of the included sources of systematic uncertainties on the observed event rate and integrated flux has been presented these do not directly translate to an uncertainty in the measured cross section. Instead, the cross section is calculated in each universe to ensure any cancellations from (anti)correlations are included. This assessment is performed for the flux and interaction systematics individually, with the extracted covariance matrices later combined, as presented in Equation 8.24.

8.4 Extracted Cross Section Results

The starting point for a cross section measurement is a spectrum of observed neutrino interactions after the selection discussed in Chapter 7 is applied. This input is shown in Figure 8.6 and represents data that would be taken by the detector, although in this case it is taken from simulation, and thus has only a statistical uncertainty. In order to maximise the performance of the cross section measurement, adjacent 100 MeV bins were merged until each contained at least 400 expected events and thus had a statistical uncertainty below 5%. This 5% target was chosen such that it is below half the flux normalisation uncertainty and thus is negligible when added in quadrature. The bin edges, in units of GeV, are given in Equation 8.27:

$$[0.2, 0.3, 0.4, 0.6, 0.8, 1.0, 1.2, 1.8, 3.0] \tag{8.27}$$

The fractional statistical uncertainty for each bin is shown in the bottom part of Figure 8.6, with an uncertainty on the integrated rate of 0.8%.

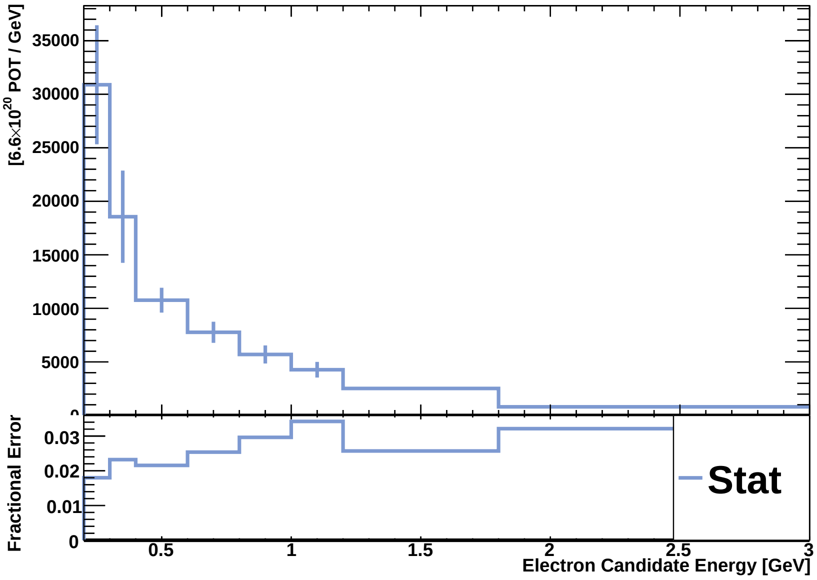
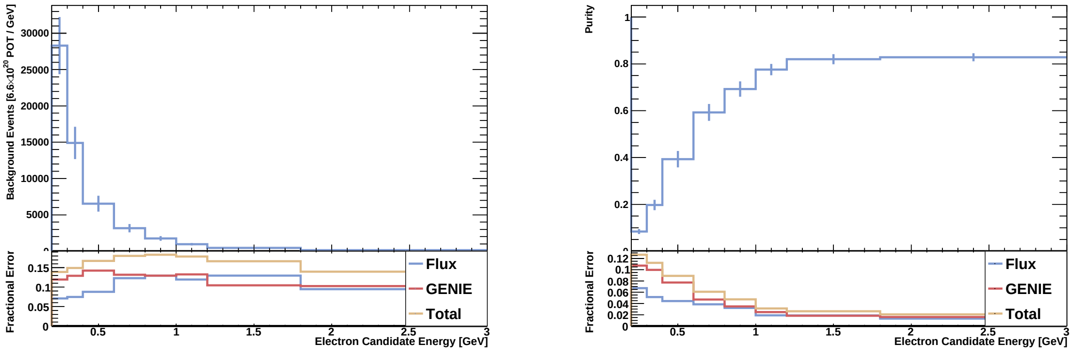


Figure 8.6: Statistical uncertainties on the selected event rate

.

The next part of the cross section extraction is the background removal, where both background subtraction and purity corrections are presented. The background subtraction begins with isolating the background events in simulation, as shown in Figure 8.7a. The backgrounds are largely concentrated at low energy, as discussed in Chapter 7. The effects of both the flux and interaction systematics are shown in the bottom part of the plot, alongside the total uncertainty from adding them in quadrature used for the error bars. This shows significant contributions from both the flux and interaction systematics, although the latter dominated at low energy.



(a) Predicted Background Rate (b) Predicted Purity

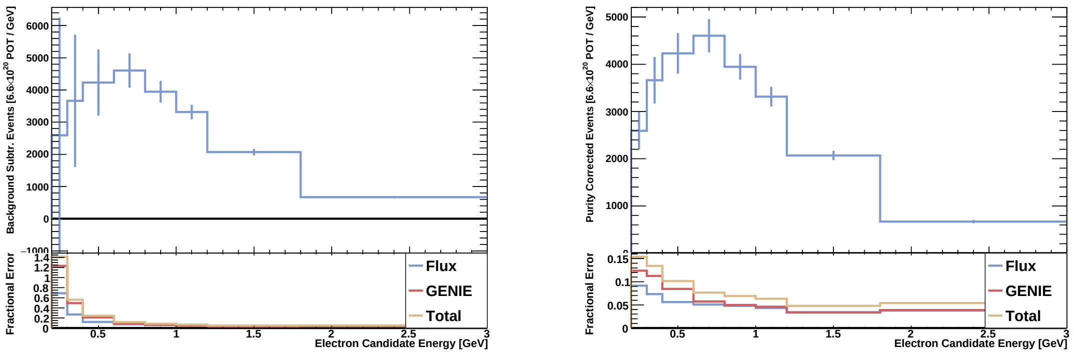
Figure 8.7: Predicted background rate (left) and purity (right) of the selection. The fractional uncertainty from the flux, interaction, and total systematic uncertainties is shown at the bottom.

The total uncertainty on the integrated background rate is 14.3%, with 7.8% and 12.0% contributions from flux and interaction systematics respectively.

Similarly, the purity, as previously defined in Equation 8.9, is shown in Figure 8.7b. The purity is lowest at low energy but gradually rises to a plateau of 83%. Again, the error bars represent the total systematic error from adding the flux and interaction systematics in quadrature. Here, the fractional uncertainty is highest at low energy, corresponding to the area with the lowest purity, and lower at high energy. The uncertainty on the integrated purity is only 7.2%, with 3.7% and 6.2% contributions from flux and interaction systematics respectively. The lower uncertainties compared to the background rate demonstrates the reduced impact of normalisation uncertainties.

Next, the backgrounds must be removed by either subtracting the expected background rate or multiplying by the purity. When subtracting the backgrounds, the error on the expected background rate is assigned to the surviving events, as can be seen in Figure 8.8a. For bins with low purity, this can lead to extremely large errors, with the lowest energy bin having a fractional error above 100%. For example, in this lowest energy bin contains 2,830 background events out of a total of 3,089. The 393 (13.9%) uncertainty on the number of background events is transferred to the remaining 259 signal, yielding an uncertainty of 152%. This inhibits the usefulness of these low purity bins in the final cross section measurement. Overall, the integrated background subtracted event rate has a 17.3% uncertainty with 9.5% and 14.5% contributions from flux and interaction systematics respectively.

In contrast, when correcting for purity it is the fractional error that is passed on to the remaining events, yielding a significantly smaller error in the case of low purity bins as shown in Figure 8.8b. However, the uncertainty increases on the high



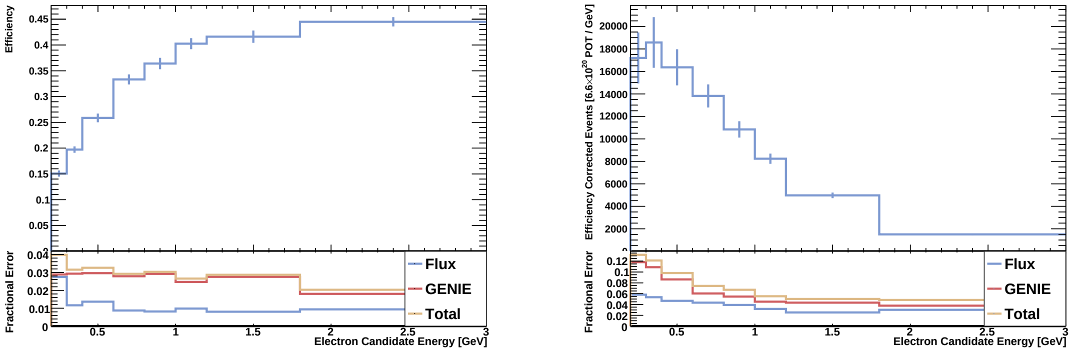
(a) Background Subtraction

(b) Purity Correction

Figure 8.8: Selected event rate after background removal via background subtraction (left) and purity correction (right). The fractional uncertainty from the flux, interaction, and total systematic uncertainties is shown at the bottom.

energy, lower background bins. Overall, the purity corrected event rate has a 7.5% uncertainty with 3.9% and 6.3% contributions from flux and interaction systematics respectively. In the absence of any dedicated side-band samples, which are not available at the time of writing for SBND, and with the current achievable purity of the selection the purity correction reduces the systematic by a factor of 57%. This is especially the case in the low energy bins which are particularly important for distinguishing between interaction models, as will be later discussed. Therefore, the purity correction method was chosen for the remainder of this analysis.

Once backgrounds have been removed, the next stage of the cross section extraction is to correct for the efficiency losses of the reconstruction and selection. This is performed similarly to the purity correction, with the efficiency curve shown in Figure 8.9a. The efficiency is assessed across the 1,000 universes and is found to be relatively consistent, with an uncertainty on the integrated efficiency of 2.5%. This comprises of contributions of 0.7% and 2.4% from flux and interaction systematics respectively. Each background subtracted data bin is then divided by the corresponding efficiency and the errors are combined with the existing errors from background removal, as is shown in Figure 8.9b. The total uncertainty on the efficiency corrected integrated rate is 7.9%, with contributions of 3.6% and 7.1% from flux and interaction systematics respectively.



(a) Efficiency

(b) Efficiency Corrected Rate

Figure 8.9: Signal efficiency curve (left) and efficiency corrected event rate (right). The fractional uncertainty from the flux, interaction, and total systematic uncertainties is shown at the bottom.

Finally, the normalisation is applied to the selection to correct for the number of targets and neutrino flux. The resultant measured cross section is shown in Figure 8.10, alongside all of the statistical and systematic errors, including the 2% POT normalisation uncertainty. The flux normalisation is the largest overall uncertainty, especially at high energy as there is some reduction at low energy due to anti-correlations between the purity correction and the normalisation. However,

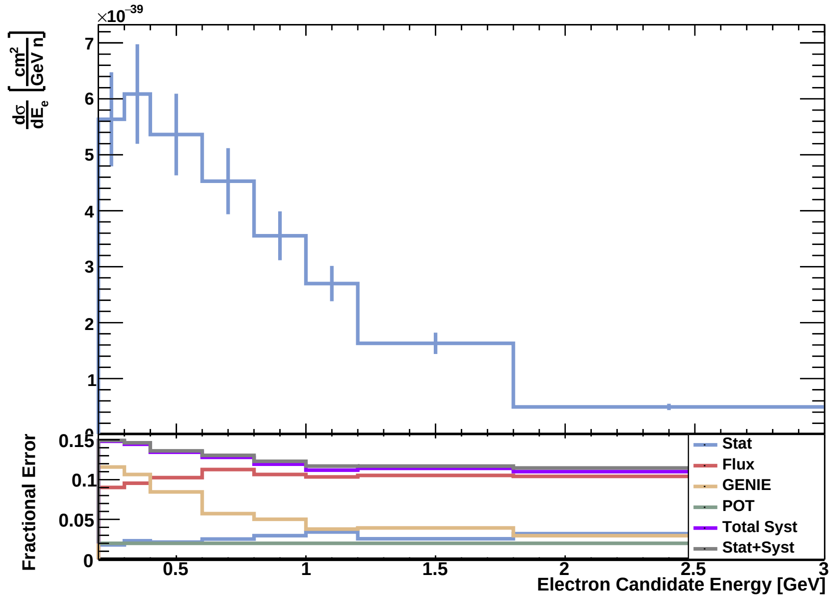


Figure 8.10: The extracted ν_e and $\overline{\nu}_e$ cross section with combined statistical and systematic errors. The contributions from the individual error sources are shown at the bottom.

the interaction systematics grow in this low energy region due to this lower purity, yielding a higher total uncertainty.

The total uncertainty on the integrated cross section is shown in Table 8.1. The flux uncertainties dominate over the other systematics, largely due to the impact of the normalisation uncertainty. Nevertheless, the interaction systematics contribute a significant amount, predominantly via the background removal. The statistical uncertainty on the integrated rate is negligible compared to the systematics.

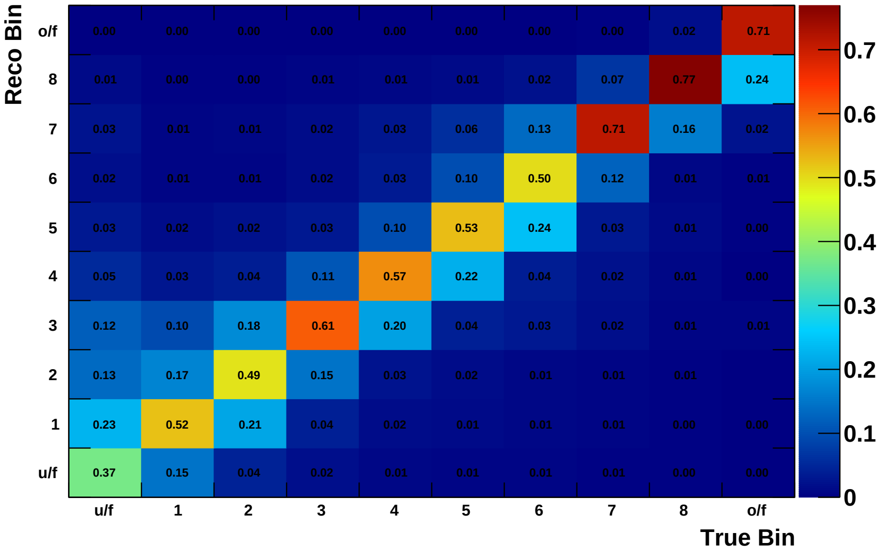
Source	Uncertainty
Flux	10.0%
Interaction	7.1%
POT	2.0%
Total Systematic	12.4%
Statistical	0.8%
Total	12.5%

Table 8.1: Systematic and statistical error contributions to the measured integrated cross section.

8.5 Model Comparisons

Once the cross section has been extracted, the next stage is to compare this to different theoretical models and assess the ability of the measurement to differentiate between said models. This process begins with generating events in SBND using GENIE tunes G00_00b_00_000 and G18_10a_02_11a, referred to as GENIEv2 and GENIEv3 respectively[181]. The key differences between these models are the use of a Local Fermi Gas (LFG) nuclear model and Valencia QE and Meson Exchange Currents (MEC) models used in GENIEv3 compared to Relativistic Fermi Gas (RFG), Llewellyn-Smith QE, and empirical MEC models in GENIEv2, the details of these are discussed in Chapter 2. These represent legacy configurations to reproduce GENIEv2_12 "Default+MEC" used in the proposal and a more modern tune respectively. GENIEv3 is the nominal tune used in the work presented, so good agreement is expected between the measured cross section and prediction as this is, effectively, a closure test. The generated events are subjected to the same lepton energy and FV cuts as defined in the signal definition in Chapter 7.1.

In order to compare the predictions to the measured cross section these simulated events, generated in true space, are mapped to reconstructed space using the forward folding matrix shown in Figure 8.11, as discussed in Section 8.2.3. The bin edges are as defined in Equation 8.27, for both truth and reconstructed space. The matrix is mostly diagonal which demonstrates that the reconstruction is correctly calibrated and operating as expected. However, the relatively large off-diagonal terms show that there is a moderate amount of smearing, with the electron energy often being underestimated by the reconstruction. This is consistent with the performance presented in Chapter 6, but demonstrates the need for continued development of the reconstruction tools.



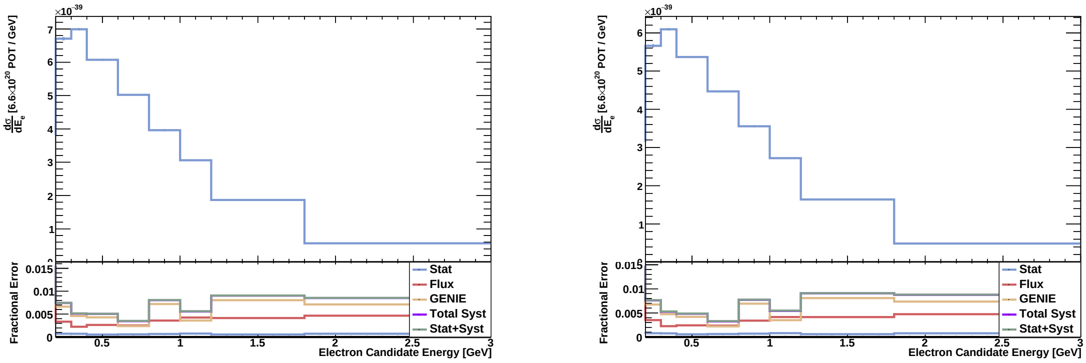
	u/f	1	2	3	4	5	6	7	8	o/f
o/f	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.71
8	0.01	0.00	0.00	0.01	0.01	0.01	0.02	0.07	0.77	0.24
7	0.03	0.01	0.01	0.02	0.03	0.06	0.13	0.71	0.16	0.02
6	0.02	0.01	0.01	0.02	0.03	0.10	0.50	0.12	0.01	0.01
5	0.03	0.02	0.02	0.03	0.10	0.53	0.24	0.03	0.01	0.00
4	0.05	0.03	0.04	0.11	0.57	0.22	0.04	0.02	0.01	0.00
3	0.12	0.10	0.18	0.61	0.20	0.04	0.03	0.02	0.01	0.01
2	0.13	0.17	0.49	0.15	0.03	0.02	0.01	0.01	0.01	
1	0.23	0.52	0.21	0.04	0.02	0.01	0.01	0.01	0.00	0.00
u/f	0.37	0.15	0.04	0.02	0.01	0.01	0.01	0.01	0.00	0.00

Figure 8.11: Forward folding matrix between true lepton energy and reconstructed electron candidate energy, column normalised.

The, nominal, flux and target nucleon normalisation corrections are then applied to convert the rate into a predicted cross section, as laid out in Equation 8.15. The extracted cross sections for each of these generator tunes is shown in Figure 8.12. GENIEv2 predicts an overall higher cross section than GENIEv3, with a 14% increase in the integrated cross section. The enhancement present in GENIEv2 is lowest at around 0.8-1 GeV with only a 11% increase whilst the increase is largest at the lowest and highest energy bins of 22% and 17% respectively.

The effect of systematic uncertainties on this folding matrix are not easily visualised in 2D, but can be seen on the predicted rate prediction shown in Figure 8.12. This uncertainty was found by constructing a forwards folding matrix in each of the systematic universes and then applying that to the predicted rate, with the errors quantified by covariance matrices. The small errors in the extracted cross section demonstrate that the forward folding procedure is largely unaffected by the systematic uncertainties, with a total uncertainty on the integrated cross section of only 0.7% for both tunes. The statistical uncertainties were reduced by generating samples larger than those expected for full exposure, corresponding to 2.5×10^{21} and 2.8×10^{21} POT for v2 and v3 respectively, yielding a statistical error below 0.1% in both cases.

Finally, these predictions are compared to the measured cross section in Figure 8.13. The agreement between the measured cross section and prediction is quantified by a χ^2, using the covariance matrix to include correlations between bins as defined in Equation 8.22. The fractional covariance and correlation matrices are shown in Appendix G for the flux, interaction, and combined uncertainties. The total error matrix used is the sum of the error matrices for the measured cross section and the generator prediction, although the former dominates.



(a) GENIEv2

(b) GENIEv3

Figure 8.12: Forward folded cross section prediction from GENIE v2 (left) and v3 (right) tunes with the systematic and statistical errors on the forward folding matrix shown at the bottom.

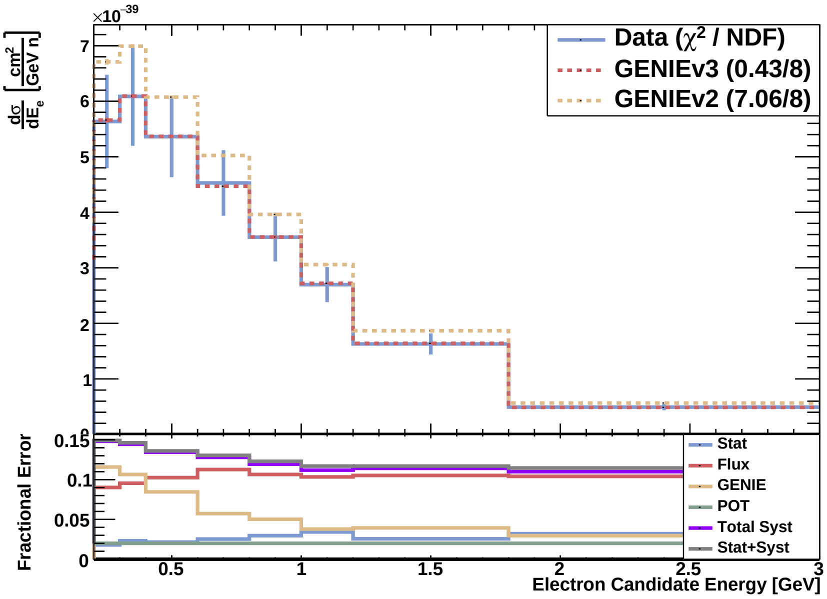


Figure 8.13: The extracted ν_e and $\overline{\nu}_e$ cross section with combined statistical and systematic errors compared to predictions from GENIE v2 and v3. χ^2 are calculated between the data and predictions using covariance matrices as outlined in Equation 8.22. The contributions from the individual error sources are shown at the bottom.

Overall, the GENIEv3 prediction shows good agreement in with the measured cross section, as expected as that was the nominal model used to generate the data. Converting the reduced χ^2 to a P value, they are consistent within 99.99%, thus successfully completing the closure test.

The GENIEv2 prediction shows significantly worse agreement with the measured data across the whole energy range, and can be rejected with 47% confidence. The largest difference between the GENIE model predictions is at lower energies but the ability to differentiate between the models is limited due to the large uncertainties in this area of phase space. Specifically, the uncertainties on the background modelling, particularly from the interaction systematics, contribute significantly at lower energies due to the lower purity. As the purity increases at higher energies the impact of the interaction systematics is reduced. However, the flux normalisation uncertainties limit the ability to effectively distinguish between the generator predictions across the entire energy range.

8.6 Discussion

Whilst the 47% confidence with which the GENIEv2 prediction can be rejected may initially seem low, it is important to view this measurement in the wider context of the existing ν_e cross section measurements. The total uncertainty on the integrated cross section of 12.5% represents a dramatic reduction from 22.2% of the current leading ν_e-Ar cross section measurement from μBooNE[157]. The historic scarcity of ν_e cross section measurements mean that this measurement, despite the seemingly low confidence, represents a large step forward.

However, this work represents the first iteration of a ν_e cross section measurement within SBND and as such there remains outstanding features that must be included into the measurement. In particular, the lack of detector uncertainties limits the conclusions that can be drawn from this work, in particular in comparison to the μBooNE measurement. It is the opinion of the author that including these should be the highest priority for future developments on this measurement.

The flux and interaction systematics that are included additionally require more in depth investigation. The large bias in the flux prediction resulting from disagreement between the fitted splines and SW parameterisation contributes a large amount of the systematic errors in this measurement[2]. In particular, the flux normalisation uncertainty that arises from this bias limits the ability of cross section models to differentiate between models across the entire energy range. The interaction systematics are also incomplete and fail to sufficiently cover all interaction modes, for example there are currently no systematics assigned to MEC events. Other experiments have expanded the list of systematics to provide more comprehensive coverage, such as those developed by μBooNE[182]. Efforts should be made to incorporate these updated systematics as they become available to ensure that the interaction systematics are as accurate as possible. Nevertheless, these interaction systematics alone are insufficient to provide insight into the potential model dependence of this measurement. Performing the measurement presented here using multiple simulated datasets utilising different generators, or at least generator tunes, as inputs would allow this model dependence to be investigated.

These improvements would enhance the realism of this measurement and thus more accurately describe the ability of SBND to differentiate between these interaction models. However, this is likely to further degrade the ability of the measurement presented here to differentiate between the models.

Nevertheless, improvements to the reconstruction and selection chains can improve the differentiating power of this measurement. Firstly, the 200 MeV threshold on electron candidate energy limits the ability to differentiate between interaction models, as the differences are often most pronounced at low energy. Continued efforts to improve the reconstruction of low energy showers can help to reduce this threshold and thus improve the sensitivity between models. However, the ability to reject backgrounds will need to improve in conjunction with these reduced thresholds, as they are more numerous at lower energies. The lowest energy bin considered in this analysis has only a 10% purity which subsequently yields a high uncertainty that in turn limits the ability to differentiate between models, demonstrating the importance of the selection performance.

Whilst performing a purity correction currently is more performant, the risk of introducing model dependence makes it generally less preferable than a background subtraction. As the reconstruction and selection continue to improve the purity of the selection, the performance advantage of the purity correction will most likely diminish. By utilising side-bands to constrain the number of backgrounds, particularly from π^0 interactions, the uncertainty associated with a background subtraction can be further reduced[68]. Additionally, a measurement of $\nu - e$ scattering could provide an in-situ constraint on the flux, potentially reducing the flux normalisation uncertainty dramatically[158].

8.7 Concluding Remarks

The extraction procedure for a $\nu_e + \overline{\nu}_e$ CC inclusive cross section measurement has been presented, considering both background subtraction and purity correction methods with the latter ultimately chosen. Statistical and systematic uncertainties have been included via covariance matrices, yielding uncertainties of 0.8% and 12.4% respectively and a total of 12.5% on the integrated measurement. The flux normalisation uncertainties dominate with a 10.0% contribution, largely due to a bias between the reweighted Cental Value (CV) and the nominal flux simulation. Interaction systematics contribute a significant 7.1% due to the low purity of the selection, particularly at low energy. Detector systematics are not available at the time of writing and are thus not included. Comparisons of the measured cross section to the forward folded prediction from the nominal GENIEv3 model found good agreement, with a P value of 99.99%, completing the closure test. In contrast, the GENIEv2 model showed comparatively poor agreement, and can be rejected with 47% confidence, demonstrating the ability of SBND to reject this model. Future developments to the reconstruction and selection alongside in-situ constraints on the flux and backgrounds can enhance the differentiating power achieved.

Chapter 9

Conclusions

In this thesis the first assessment of the ability of the the Short-Baseline Near Detector (SBND) to perform a $\nu_e + \overline{\nu}_e$ Charged Current (CC) inclusive cross section measurement has been presented. This measurement is vital to the continued development of neutrino nucleus interaction modelling to unlock the potential of future experiments, particularly Liquid Argon Time Projection Chambers (LAr TPCs). Due to the proximity to the neutrino beam source, SBND aims to collect the largest number of neutrino interactions of any LAr TPC. This unprecedented level of statistics allows SBND not only extremely low statistical uncertainties but also the ability to probe rare phenomenon and exclusive channels. Although ν_e constitute only 0.5% of the total Booster Neutrino Beam (BNB) flux, SBND still expects to observe the largest number of ν_e interactions of any LAr TPC.

Developments to the reconstruction, at both the pattern recognition and shower characterisation levels, have been presented which improve the ability of SBND to perform such a measurement, discussed in Chapters 5 and 6 respectively. The improved vertexing, clustering, and characterisation resulting from the implementation of both Multi Variate Analyses (MVAs) and new algorithms improves the reconstruction performance. This is particularly the case for showers with a dramatic reduction in segmentation resulting in improved selection efficiency and energy resolution. The new shower characterisation that was developed, PandoraModularShowerCreation, allows for rapid development of tools to reconstruct the individual characteristics of the shower. Building upon the improvements in pattern recognition, this shower reconstruction provides a solid foundation for which subsequent selections and measurements can be based.

A selection has additionally been developed to identify the $\nu_e + \overline{\nu}_e$ CC signal whilst rejecting both cosmogenic and neutrino induced backgrounds that dramatically outnumber the signal initially, as presented in Chapter 7. This initial low purity is overcome by rejecting over 99.999% of cosmogenic and 99.85% of neutrino induced backgrounds yielding a final purity of 45.1% whilst maintaining an efficiency of 28.1%. This not only improves the purity of the selected sample but also ensures the containment, and thus energy resolution, of the selected events.

Finally, building upon the reconstruction and selection previously discussed, the ability of SBND to measure the $\nu_e + \overline{\nu}_e$ CC inclusive cross section was assessed in Chapter 8. In the absence of real data, the cross section was calculated from a simulated data, using GENIEv3 as the nominal interaction model. It was found that the GENIEv2 model can be rejected with 47% confidence whilst the nominal GENIEv3 model was accepted with over 99.99% confidence, successfully completing the closure test. The uncertainty in the flux was found to be the largest contribution to the overall uncertainty, largely due to disagreement between the nominal value and reweighted universes. The 12.5% total error on the integrated cross section is dominated by this flux uncertainty with a contribution of 10.0%, with a 7.1% from uncertainty in the underlying interaction model and only a 0.8% contribution from statistical uncertainties.

As a first iteration of this analysis, the lack of detector systematics and alternative model data limit the conclusions that can be drawn from this study. However, the continued improvements to the pattern recognition, reconstruction, and selection can undoubtedly reduce the remaining backgrounds, in turn reducing the systematic uncertainty on the measurement. Nevertheless, even at this early stage of the experiments lifetime SBND demonstrates the ability to make world leading ν_e CC inclusive cross section measurements.

Bibliography

[1] R. Acciarri et al. “A Proposal for a Three Detector Short-Baseline Neutrino Oscillation Program in the Fermilab Booster Neutrino Beam” (Mar. 2015). arXiv: 1503.01520.

[2] MiniBooNE Collaboration et al. “Neutrino flux prediction at MiniBooNE”. *Physical Review D* 79.7 (Apr. 2009), p. 072002. DOI: 10.1103/PhysRevD.79.072002.

[3] J. S. Marshall and M. A. Thomson. “The Pandora software development kit for pattern recognition”. *The European Physical Journal C* 75.9 (Sept. 2015), p. 439. DOI: 10.1140/epjc/s10052-015-3659-3.

[4] R. Acciarri et al. “The Pandora multi-algorithm approach to automated pattern recognition of cosmic-ray muon and neutrino events in the MicroBooNE detector”. *The European Physical Journal C* 78.1 (Jan. 2018), p. 82. DOI: 10.1140/epjc/s10052-017-5481-6.

[5] E. L. Snider and G. Petrillo. “LArSoft: toolkit for simulation, reconstruction and analysis of liquid argon TPC neutrino detectors”. *Journal of Physics: Conference Series* 898 (Oct. 2017), p. 042057. DOI: 10.1088/1742-6596/898/4/042057.

[6] Dominic Barker. “Developments Towards a Electron Neutrino CC Sterile Appearance Sensitivity in the Short-Baseline Neutrino Programme”. PhD thesis. University of Sheffield, July 2020.

[7] Michael Wallbank. “Reconstruction and Analysis for the DUNE 35-ton Liquid Argon Prototype”. PhD thesis. University of Sheffield, Feb. 2018.

[8] C. Backhouse. “The CAFAna framework for neutrino analysis” (Mar. 2022). DOI: 10.48550/arXiv.2203.13768. arXiv: 2203.13768.

[9] Thomas Brooks. “Selecting Charged Current Muon Neutrino Interactions on Argon with the Short-Baseline Near Detector”. PhD thesis. University of Sheffield, Apr. 2020.

[10] Charles Drummond Ellis, W. A. Wooster, and Ernest Rutherford. “The average energy of disintegration of radium E”. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 117.776 (Dec. 1927), pp. 109–123. DOI: 10.1098/rspa.1927.0168.

[11] Laurie M. Brown. “The idea of the neutrino”. *Physics Today* 31.9 (Sept. 1978), pp. 23–28. DOI: 10.1063/1.2995181.

[12] Fred L. Wilson. “Fermi’s Theory of Beta Decay”. *American Journal of Physics* 36.12 (Dec. 1968), pp. 1150–1160. DOI: 10.1119/1.1974382.

[13] C. L. Cowan, F. Reines, F. B. Harrison, H. W. Kruse, and A. D. McGuire. “Detection of the Free Neutrino: a Confirmation”. *Science* 124.3212 (July 1956), pp. 103–104. DOI: 10.1126/science.124.3212.103.

[14] B. Pontekorvo. “ELECTRON AND MUON NEUTRINOS”. *Zhur. Eksptl’. i Teoret. Fiz.* Vol: 37 (Dec. 1959).

[15] G. Danby et al. “Observation of High-Energy Neutrino Reactions and the Existence of Two Kinds of Neutrinos”. *Physical Review Letters* 9.1 (July 1962), pp. 36–44. DOI: 10.1103/PhysRevLett.9.36.

[16] M. L. Perl et al. “Evidence for Anomalous Lepton Production in $e^+ - e^-$ Annihilation”. *Physical Review Letters* 35.22 (Dec. 1975), pp. 1489–1492. DOI: 10.1103/PhysRevLett.35.1489.

[17] M. L. Perl and W. T. Kirk. “Heavy leptons. [History of discovery]”. *Sci. Am.; (United States)* 238:3 (1978).

[18] D. DeCamp et al. “Determination of the number of light neutrino species”. *Physics Letters B* 231.4 (Nov. 1989), pp. 519–529. DOI: 10.1016/0370-2693(89)90704-1.

[19] “Precision electroweak measurements on the Z resonance”. *Physics Reports* 427.5 (May 2006), pp. 257–454. DOI: 10.1016/j.physrep.2005.12.006.

[20] K. Kodama et al. “Observation of tau neutrino interactions”. *Physics Letters B* 504.3 (Apr. 2001), pp. 218–224. DOI: 10.1016/S0370-2693(01)00307-0.

[21] Raymond Davis, Don S. Harmer, and Kenneth C. Hoffman. “Search for Neutrinos from the Sun”. *Physical Review Letters* 20.21 (May 1968), pp. 1205–1209. DOI: 10.1103/PhysRevLett.20.1205.

[22] K. S. Hirata et al. “Results from one thousand days of real-time, directional solar-neutrino data”. *Physical Review Letters* 65.11 (Sept. 1990), pp. 1297–1300. DOI: 10.1103/PhysRevLett.65.1297.

[23] A. I. Abazov et al. “Search for neutrinos from the Sun using the reaction 71Ga(νe,e-)71Ge”. *Physical Review Letters* 67.24 (Dec. 1991), pp. 3332–3335. DOI: 10.1103/PhysRevLett.67.3332.

[24] P. Anselmann et al. “Solar neutrinos observed by GALLEX at Gran Sasso”. *Physics Letters B* 285.4 (July 1992), pp. 376–389. DOI: 10.1016/0370-2693(92)91521-A.

[25] J. N. Bahcall. “Solar Models and Solar Neutrinos”. *Physica Scripta* T121 (Jan. 2005), pp. 46–50. DOI: 10.1088/0031-8949/2005/T121/006.

[26] T. K. Gaisser and M. Honda. “Flux of Atmospheric Neutrinos”. *Annual Review of Nuclear and Particle Science* 52.1 (2002), pp. 153–199. DOI: 10.1146/annurev.nucl.52.050102.090645.

[27] T. J. Haines et al. “Calculation of Atmospheric Neutrino-Induced Backgrounds in a Nucleon-Decay Search”. *Physical Review Letters* 57.16 (Oct. 1986), pp. 1986–1989. DOI: 10.1103/PhysRevLett.57.1986.

[28] R. Becker-Szendy et al. “Electron- and muon-neutrino content of the atmospheric flux”. *Physical Review D* 46.9 (Nov. 1992), pp. 3720–3724. DOI: 10.1103/PhysRevD.46.3720.

[29] K. S. Hirata et al. “Experimental study of the atmospheric neutrino flux”. *Physics Letters B* 205.2 (Apr. 1988), pp. 416–420. DOI: 10.1016/0370-2693(88)91690-5.

[30] K. S. Hirata et al. “Observation of a small atmospheric ν_μ/ν_e ratio in Kamiokande”. *Physics Letters B* 280.1 (Apr. 1992), pp. 146–152. DOI: 10.1016/0370-2693(92)90788-6.

[31] B. Pontecorvo. “Inverse beta processes and nonconservation of lepton charge”. *Zh. Eksp. Teor. Fiz.* 34 (1957), p. 247.

[32] Ziro Maki, Masami Nakagawa, and Shoichi Sakata. “Remarks on the Unified Model of Elementary Particles”. *Progress of Theoretical Physics* 28.5 (Nov. 1962), pp. 870–880. DOI: 10.1143/PTP.28.870.

[33] S. M. Bilenky, J. Hošek, and S. T. Petcov. “On the oscillations of neutrinos with Dirac and Majorana masses”. *Physics Letters B* 94.4 (Aug. 1980), pp. 495–498. DOI: 10.1016/0370-2693(80)90927-2.

[34] L. Wolfenstein. “Neutrino oscillations in matter”. *Physical Review D* 17.9 (May 1978), pp. 2369–2374. DOI: 10.1103/PhysRevD.17.2369.

[35] The T2K Collaboration et al. “Constraint on the matter–antimatter symmetry-violating phase in neutrino oscillations”. *Nature* 580.7803 (Apr. 2020), pp. 339–344. DOI: 10.1038/s41586-020-2177-0.

[36] Y. Fukuda et al. “Evidence for Oscillation of Atmospheric Neutrinos”. *Physical Review Letters* 81.8 (Aug. 1998), pp. 1562–1567. DOI: 10.1103/PhysRevLett.81.1562.

[37] Q. R. Ahmad et al. “Direct Evidence for Neutrino Flavor Transformation from Neutral-Current Interactions in the Sudbury Neutrino Observatory”. *Physical Review Letters* 89.1 (June 2002), p. 011301. DOI: 10.1103/PhysRevLett.89.011301.

[38] M. Aker et al. “Direct neutrino-mass measurement with sub-electronvolt sensitivity”. *Nature Physics* 18.2 (Feb. 2022), pp. 160–166. DOI: 10.1038/s41567-021-01463-1.

[39] *mass-hierarchy-jgu-mainz-web.jpg (1200×676).* URL: https://neutrinos.fnal.gov/wp-content/uploads/2018/04/mass-hierarchy-jgu-mainz-web.jpg (visited on 01/10/2022).

[40] Borexino Collaboration. “Neutrinos from the primary proton–proton fusion process in the Sun”. *Nature* 512.7515 (Aug. 2014), pp. 383–386. DOI: 10.1038/nature13702.

[41] B. Aharmim et al. “Combined analysis of all three phases of solar neutrino data from the Sudbury Neutrino Observatory”. *Physical Review C* 88.2 (Aug. 2013), p. 025501. DOI: 10.1103/PhysRevC.88.025501.

[42] K. Abe et al. “Solar neutrino results in Super-Kamiokande-III”. *Physical Review D* 83.5 (Mar. 2011), p. 052010. DOI: 10.1103/PhysRevD.83.052010.

[43] A. Gando et al. “Reactor on-off antineutrino measurement with KamLAND”. *Physical Review D* 88.3 (Aug. 2013), p. 033001. DOI: 10.1103/PhysRevD.88.033001.

[44] D. Adey et al. “Measurement of the Electron Antineutrino Oscillation with 1958 Days of Operation at Daya Bay”. *Physical Review Letters* 121.24 (Dec. 2018), p. 241805. DOI: 10.1103/PhysRevLett.121.241805.

[45] G. Bak et al. “Measurement of Reactor Antineutrino Oscillation Amplitude and Frequency at RENO”. *Physical Review Letters* 121.20 (Nov. 2018), p. 201801. DOI: 10.1103/PhysRevLett.121.201801.

[46] The Double Chooz Collaboration. “Double Chooz $\theta 13$ measurement via total neutron capture detection”. *Nature Physics* 16.5 (May 2020), pp. 558–564. DOI: 10.1038/s41567-020-0831-y.

[47] M. A. Acero et al. “First measurement of neutrino oscillation parameters using neutrinos and antineutrinos by NOvA”. *Physical Review Letters* 123.15 (Oct. 2019), p. 151803. DOI: 10.1103/PhysRevLett.123.151803.

[48] Super-Kamiokande Collaboration et al. “Atmospheric neutrino oscillation analysis with external constraints in Super-Kamiokande I-IV”. *Physical Review D* 97.7 (Apr. 2018), p. 072001. DOI: 10.1103/PhysRevD.97.072001.

[49] IceCube Collaboration et al. “Determining neutrino oscillation parameters from atmospheric muon neutrino disappearance with three years of IceCube DeepCore data”. *Physical Review D* 91.7 (Apr. 2015), p. 072004. DOI: 10.1103/PhysRevD.91.072004.

[50] Ivan Esteban, M.C. Gonzalez-Garcia, Michele Maltoni, Thomas Schwetz, and Albert Zhou. “The fate of hints: updated global analysis of three-flavor neutrino oscillations”. *Journal of High Energy Physics* 2020.9 (Sept. 2020), p. 178. DOI: 10.1007/JHEP09(2020)178.

[51] B. Abi et al. “Deep Underground Neutrino Experiment (DUNE), Far Detector Technical Design Report, Volume II: DUNE Physics” (Mar. 2020). arXiv: 2002.03005.

[52] Hyper-Kamiokande Proto-Collaboration et al. “Hyper-Kamiokande Design Report” (Nov. 2018). arXiv: 1805.04163.

[53] Fengpeng An et al. “Neutrino physics with JUNO”. *Journal of Physics G: Nuclear and Particle Physics* 43.3 (Feb. 2016), p. 030401. DOI: 10.1088/0954-3899/43/3/030401.

[54] A. Diaz, C. A. Argüelles, G. H. Collin, J. M. Conrad, and M. H. Shaevitz. “Where are we with light sterile neutrinos?” *Physics Reports*. Where are we with light sterile neutrinos? 884 (Nov. 2020), pp. 1–59. DOI: 10.1016/j.physrep.2020.08.005.

[55] S. Gariazzo, C. Giunti, M. Laveder, Y. F. Li, and E. M. Zavanin. “Light sterile neutrinos”. *Journal of Physics G: Nuclear and Particle Physics* 43.3 (Mar. 2015), p. 033001. DOI: 10.1088/0954-3899/43/3/033001.

[56] A. Aguilar et al. “Evidence for neutrino oscillations from the observation of ν ¯ e appearance in a ν ¯ μ beam”. *Physical Review D* 64.11 (Nov. 2001), p. 112007. DOI: 10.1103/PhysRevD.64.112007.

[57] A. A. Aguilar-Arevalo et al. “Significant Excess of Electronlike Events in the MiniBooNE Short-Baseline Neutrino Experiment”. *Physical Review Letters* 121.22 (Nov. 2018), p. 221801. DOI: 10.1103/PhysRevLett.121.221801.

[58] A. A. Aguilar-Arevalo et al. “Updated MiniBooNE neutrino oscillation results with increased data and new background studies”. *Physical Review D* 103.5 (Mar. 2021), p. 052002. DOI: 10.1103/PhysRevD.103.052002.

[59] KARMEN Collaboration et al. “Upper limits for neutrino oscillations $\overline{\nu}_\mu \to \overline{\nu}_e$ from muon decay at rest”. *Physical Review D* 65.11 (June 2002), p. 112001. DOI: 10.1103/PhysRevD.65.112001.

[60] The OPERA collaboration et al. “Final results of the search for $\nu_\mu \to \nu_e$ oscillations with the OPERA detector in the CNGS beam”. *Journal of High Energy Physics* 2018.6 (June 2018), p. 151. DOI: 10.1007/JHEP06(2018)151.

[61] Mona Dentler et al. “Updated global analysis of neutrino oscillations in the presence of eV-scale sterile neutrinos”. *Journal of High Energy Physics* 2018.8 (Aug. 2018), p. 10. DOI: 10.1007/JHEP08(2018)010.

[62] MINOS+ Collaboration et al. “Search for Sterile Neutrinos in MINOS and MINOS+ Using a Two-Detector Fit”. *Physical Review Letters* 122.9 (Mar. 2019), p. 091803. DOI: 10.1103/PhysRevLett.122.091803.

[63] MiniBooNE and SciBooNE Collaborations et al. “Dual baseline search for muon antineutrino disappearance at $0.1\ \text{eV}^2 < \Delta m^2 < 100\ \text{eV}^2$”. *Physical Review D* 86.5 (Sept. 2012), p. 052009. DOI: 10.1103/PhysRevD.86.052009.

[64] The Super-Kamiokande Collaboration et al. “Atmospheric neutrino oscillation analysis with subleading effects in Super-Kamiokande I, II, and III”. *Physical Review D* 81.9 (May 2010), p. 092004. DOI: 10.1103/PhysRevD.81.092004.

[65] IceCube Collaboration et al. “Searches for Sterile Neutrinos with the IceCube Detector”. *Physical Review Letters* 117.7 (Aug. 2016), p. 071801. DOI: 10.1103/PhysRevLett.117.071801.

[66] M. Harada et al. “Proposal: A Search for Sterile Neutrino at J-PARC Materials and Life Science Experimental Facility” (Oct. 2013). arXiv: 1310.1437.

[67] MicroBooNE collaboration et al. “Search for Neutrino-Induced Neutral Current Δ Radiative Decay in MicroBooNE and a First Test of the MiniBooNE Low Energy Excess Under a Single-Photon Hypothesis” (Oct. 2021). arXiv: 2110.00409.

[68] MicroBooNE collaboration et al. “Search for an Excess of Electron Neutrino Interactions in MicroBooNE Using Multiple Final State Topologies” (Oct. 2021). arXiv: 2110.14054.

[69] I. Alekseev et al. “Search for sterile neutrinos at the DANSS experiment”. *Physics Letters B* 787 (Dec. 2018), pp. 56–63. DOI: 10.1016/j.physletb.2018.10.038.

[70] A. P. Serebrov et al. “Search for sterile neutrinos with the Neutrino-4 experiment and measurement results”. *Physical Review D* 104.3 (Aug. 2021), p. 032003. DOI: 10.1103/PhysRevD.104.032003.

[71] PROSPECT Collaboration et al. “Improved short-baseline neutrino oscillation search and energy spectrum measurement with the PROSPECT experiment at HFIR”. *Physical Review D* 103.3 (Feb. 2021), p. 032001. DOI: 10.1103/PhysRevD.103.032001.

[72] Y. Abreu et al. “SoLid: a short baseline reactor neutrino experiment”. *Journal of Instrumentation* 16.2 (Feb. 2021), P02025–P02025. DOI: 10.1088/1748-0221/16/02/P02025.

[73] STEREO Collaboration et al. “Improved sterile neutrino constraints from the STEREO experiment with 179 days of reactor-on data”. *Physical Review D* 102.5 (Sept. 2020), p. 052002. DOI: 10.1103/PhysRevD.102.052002.

[74] M. A. Acero et al. “Search for Active-Sterile Antineutrino Mixing Using Neutral-Current Interactions with the NOvA Experiment”. *Physical Review Letters* 127.20 (Nov. 2021), p. 201801. DOI: 10.1103/PhysRevLett.127.201801.

[75] K. Abe et al. “Search for light sterile neutrinos with the T2K far detector Super-Kamiokande at a baseline of 295 km”. *Physical Review D* 99.7 (Apr. 2019), p. 071103. DOI: 10.1103/PhysRevD.99.071103.

[76] M. G. Aartsen et al. “Searching for eV-scale sterile neutrinos with eight years of atmospheric neutrinos at the IceCube Neutrino Telescope”. *Physical Review D* 102.5 (Sept. 2020), p. 052009. DOI: 10.1103/PhysRevD.102.052009.

[77] NOvA Collaboration et al. “New constraints on oscillation parameters from ν_e appearance and ν_μ disappearance in the NOvA experiment”. *Physical Review D* 98.3 (Aug. 2018), p. 032012. DOI: 10.1103/PhysRevD.98.032012.

[78] K. Abe et al. “Improved constraints on neutrino mixing from the T2K experiment with 3.13 × 10 21 protons on target”. *Physical Review D* 103.11 (June 2021), p. 112008. DOI: 10.1103/PhysRevD.103.112008.

[79] J. Nieves, F. Sánchez, I. Ruiz Simo, and M. J. Vicente Vacas. “Neutrino energy reconstruction and the shape of the charged current quasielastic-like total cross section”. *Physical Review D* 85.11 (June 2012), p. 113008. DOI: 10.1103/PhysRevD.85.113008.

[80] Melanie Day and Kevin S. McFarland. “Differences in quasielastic cross sections of muon and electron neutrinos”. *Physical Review D* 86.5 (Sept. 2012), p. 053003. DOI: 10.1103/PhysRevD.86.053003.

[81] The T2K Collaboration et al. “Measurement of the muon neutrino charged-current single π^+ production on hydrocarbon using the T2K off-axis near detector ND280”. *Physical Review D* 101.1 (Jan. 2020), p. 012007. DOI: 10.1103/PhysRevD.101.012007.

[82] MINERvA Collaboration et al. “Measurement of Electron Neutrino Quasielastic and Quasielasticlike Scattering on Hydrocarbon at $E_\nu = 3.6$ GeV”. *Physical Review Letters* 116.8 (Feb. 2016), p. 081802. DOI: 10.1103/PhysRevLett.116.081802.

[83] C. Andreopoulos et al. “The GENIE neutrino Monte Carlo generator”. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 614.1 (Feb. 2010), pp. 87–104. DOI: 10.1016/j.nima.2009.12.009.

[84] J. A. Formaggio and G. P. Zeller. “From eV to EeV: Neutrino cross sections across energy scales”. *Reviews of Modern Physics* 84.3 (Sept. 2012), pp. 1307–1341. DOI: 10.1103/RevModPhys.84.1307.

[85] C. H. Llewellyn Smith. “Neutrino reactions at accelerator energies”. *Physics Reports* 3.5 (June 1972), pp. 261–379. DOI: 10.1016/0370-1573(72)90010-5.

[86] Teppei Katori. “Meson Exchange Current (MEC) Models in Neutrino Interaction Generators”. 1663 (Apr. 2013). DOI: 10.1063/1.4919465.

[87] J. Nieves, I. Ruiz Simo, and M. J. Vicente Vacas. “Inclusive charged-current neutrino-nucleus reactions”. *Physical Review C* 83.4 (Apr. 2011), p. 045501. DOI: 10.1103/PhysRevC.83.045501.

[88] Dieter Rein and Lalit M. Sehgal. “Coherent $\pi 0$ production in neutrino reactions”. *Nuclear Physics B* 223.1 (Aug. 1983), pp. 29–44. DOI: 10.1016/0550-3213(83)90090-1.

[89] Ch. Berger and L. M. Sehgal. “Lepton mass effects in single pion production by neutrinos”. *Physical Review D* 76.11 (Dec. 2007), p. 113004. DOI: 10.1103/PhysRevD.76.113004.

[90] A. Bodek and U. K. Yang. “Higher twist, ξ_W, and effective LO PDFs for lepton scattering in the few GeV region”. *Journal of Physics G: Nuclear and Particle Physics* 29.8 (July 2003), pp. 1899–1905. DOI: 10.1088/0954-3899/29/8/369.

[91] T. Yang, C. Andreopoulos, H. Gallagher, K. Hofmann, and P. Kehayias. “A hadronization model for few-GeV neutrino interactions”. *The European Physical Journal C* 63.1 (Sept. 2009), pp. 1–10. DOI: 10.1140/epjc/s10052-009-1094-z.

[92] J. E. Amaro et al. “Using electron scattering superscaling to predict charge-changing neutrino cross sections in nuclei”. *Physical Review C* 71.1 (Jan. 2005), p. 015501. DOI: 10.1103/PhysRevC.71.015501.

[93] M. Khachatryan et al. “Electron-beam energy reconstruction for neutrino oscillation measurements”. *Nature* 599.7886 (Nov. 2021), pp. 565–570. DOI: 10.1038/s41586-021-04046-5.

[94] Luis Alvarez-Ruso et al. “Recent highlights from GENIE v3”. *The European Physical Journal Special Topics* 230.24 (Dec. 2021), pp. 4449–4467. DOI: 10.1140/epjs/s11734-021-00295-7.

[95] T. Cai et al. “Nucleon binding energy and transverse momentum imbalance in neutrino-nucleus reactions”. *Physical Review D* 101.9 (May 2020), p. 092001. DOI: 10.1103/PhysRevD.101.092001.

[96] R. Gran, J. Nieves, F. Sanchez, and M. J. Vicente Vacas. “Neutrino-nucleus quasi-elastic and 2p2h interactions up to 10 GeV”. *Physical Review D* 88.11 (Dec. 2013), p. 113007. DOI: 10.1103/PhysRevD.88.113007.

[97] G. D. Megias et al. “Meson-exchange currents and quasielastic predictions for charged-current neutrino- C 12 scattering in the superscaling approach”. *Physical Review D* 91.7 (Apr. 2015), p. 073004. DOI: 10.1103/PhysRevD.91.073004.

[98] Sowjanya Gollapinni. “Neutrino Cross section Future” (Feb. 2016). arXiv: 1602.05299.

[99] Júlia Tena-Vidal et al. “Neutrino-nucleon cross-section model tuning in GENIE v3”. *Physical Review D* 104.7 (Oct. 2021), p. 072009. DOI: 10.1103/PhysRevD.104.072009.

[100] S. J. Barish et al. “Study of neutrino interactions in hydrogen and deuterium: Description of the experiment and study of the reaction $\nu + d \rightarrow \mu^- + p + p_s$”. *Physical Review D* 16.11 (Dec. 1977), pp. 3103–3121. DOI: 10.1103/PhysRevD.16.3103.

[101] J. Blietschau et al. “Total cross sections for ve and ve interactions and search for neutrino oscillations and decay”. *Nuclear Physics B* 133.2 (Feb. 1978), pp. 205–219. DOI: 10.1016/0550-3213(78)90299-7.

[102] Véronique Bernard, Latifa Elouadrhiri, and Ulf-G. Meißner. “Axial structure of the nucleon”. *Journal of Physics G: Nuclear and Particle Physics* 28.1 (Nov. 2001), R1–R35. DOI: 10.1088/0954-3899/28/1/201.

[103] MiniBooNE Collaboration et al. “First measurement of the muon neutrino charged current quasielastic double differential cross section”. *Physical Review D* 81.9 (May 2010), p. 092005. DOI: 10.1103/PhysRevD.81.092005.

[104] M. Martini, M. Ericson, G. Chanfray, and J. Marteau. “Neutrino and antineutrino quasielastic interactions with nuclei”. *Physical Review C* 81.4 (Apr. 2010), p. 045502. DOI: 10.1103/PhysRevC.81.045502.

[105] L. Aliaga et al. “Design, calibration, and performance of the MINERvA detector”. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 743 (Apr. 2014), pp. 130–159. DOI: 10.1016/j.nima.2013.12.053.

[106] MINERνA Collaboration et al. “Double-differential inclusive charged-current ν_μ cross sections on hydrocarbon in MINERvA at $E_\nu \sim 3.5$ GeV”. *Physical Review D* 101.11 (June 2020), p. 112007. DOI: 10.1103/PhysRevD.101.112007.

[107] D. Ruterbories et al. “Measurement of inclusive charged-current ν_μ cross sections as a function of muon kinematics at $E_\nu \sim 6$ GeV on hydrocarbon”. *Physical Review D* 104.9 (Nov. 2021), p. 092007. DOI: 10.1103/PhysRevD.104.092007.

[108] MINERvA Collaboration et al. “Measurement of Ratios of ν_μ Charged-Current Cross Sections on C, Fe, and Pb to CH at Neutrino Energies 2–20 GeV”. *Physical Review Letters* 112.23 (June 2014), p. 231801. DOI: 10.1103/PhysRevLett.112.231801.

[109] The MINERνA Collaboration et al. “Nucleon binding energy and transverse momentum imbalance in neutrino-nucleus reactions”. *Physical Review D* 101.9 (May 2020), p. 092001. DOI: 10.1103/PhysRevD.101.092001.

[110] MINERvA Collaboration et al. “Direct Measurement of Nuclear Dependence of Charged Current Quasielasticlike Neutrino Interactions Using MINERvA”. *Physical Review Letters* 119.8 (Aug. 2017), p. 082001. DOI: 10.1103/PhysRevLett.119.082001.

[111] T2K Collaboration et al. “Measurement of inclusive double-differential ν_μ charged-current cross section with improved acceptance in the T2K off-axis near detector”. *Physical Review D* 98.1 (July 2018), p. 012004. DOI: 10.1103/PhysRevD.98.012004.

[112] T2K Collaboration et al. “First combined measurement of the muon neutrino and antineutrino charged-current cross section without pions in the final state at T2K”. *Physical Review D* 101.11 (June 2020), p. 112001. DOI: 10.1103/PhysRevD.101.112001.

[113] T2K Collaboration et al. “First T2K measurement of transverse kinematic imbalance in the muon-neutrino charged-current single-π^+ production channel containing at least one proton”. *Physical Review D* 103.11 (June 2021), p. 112009. DOI: 10.1103/PhysRevD.103.112009.

[114] K. Abe et al. “Measurement of the charged-current electron (anti-)neutrino inclusive cross-sections at the T2K off-axis near detector ND280”. *Journal of High Energy Physics* 2020.10 (Oct. 2020), p. 114. DOI: 10.1007/JHEP10(2020)114.

[115] C. Rubbia. “The Liquid Argon Time Projection Chamber: A New Concept for Neutrino Detectors” (May 1977).

[116] ArgoNeuT Collaboration et al. “First Measurements of Inclusive Muon Neutrino Charged Current Differential Cross Sections on Argon”. *Physical Review Letters* 108.16 (Apr. 2012), p. 161802. DOI: 10.1103/PhysRevLett.108.161802.

[117] R. Acciarri et al. “First Measurement of Neutrino and Antineutrino Coherent Charged Pion Production on Argon”. *Physical Review Letters* 113.26 (Dec. 2014), p. 261801. DOI: 10.1103/PhysRevLett.113.261801.

[118] R. Acciarri et al. “Detection of back-to-back proton pairs in charged-current neutrino interactions with the ArgoNeuT detector in the NuMI low energy beam line”. *Physical Review D* 90.1 (July 2014), p. 012008. DOI: 10.1103/PhysRevD.90.012008.

[119] ArgoNeuT Collaboration et al. “First measurement of electron neutrino scattering cross section on argon”. *Physical Review D* 102.1 (July 2020), p. 011101. DOI: 10.1103/PhysRevD.102.011101.

[120] P. Abratenko et al. “Measurement of space charge effects in the MicroBooNE LArTPC using cosmic muons”. *Journal of Instrumentation* 15.12 (Dec. 2020), P12037–P12037. DOI: 10.1088/1748-0221/15/12/P12037.

[121] P. Abratenko et al. “First Measurement of Inclusive Muon Neutrino Charged Current Differential Cross Sections on Argon at E ν 0.8 GeV with the MicroBooNE Detector”. *Physical Review Letters* 123.13 (Sept. 2019), p. 131801. DOI: 10.1103/PhysRevLett.123.131801.

[122] MicroBooNE Collaboration et al. “First Measurement of Differential Charged Current Quasielasticlike ν_μ-Argon Scattering Cross Sections with the MicroBooNE Detector”. *Physical Review Letters* 125.20 (Nov. 2020), p. 201803. DOI: 10.1103/PhysRevLett.125.201803.

[123] MicroBooNE Collaboration et al. "First Measurement of Energy-Dependent Inclusive Muon Neutrino Charged-Current Cross Sections on Argon with the MicroBooNE Detector". *Physical Review Letters* 128.15 (Apr. 2022), p. 151801. DOI: 10.1103/PhysRevLett.128.151801.

[124] R. Acciarri et al. "Design and Construction of the MicroBooNE Detector". *Journal of Instrumentation* 12 (Dec. 2016). DOI: 10.1088/1748-0221/12/02/P02017.

[125] Ch Tegeler, Roland Span, and Wolfgang Wagner. "A New Equation of State for Argon Covering the Fluid Region for Temperatures From the Melting Line to 700 K at Pressures up to 1000 MPa". *Journal of Physical and Chemical Reference Data* 28.3 (Aug. 1999), p. 779. DOI: 10.1063/1.556037.

[126] M. Antonello et al. "Experimental observation of an extremely high electron lifetime with the ICARUS-T600 LAr-TPC". *Journal of Instrumentation* 9.12 (Dec. 2014), P12006–P12006. DOI: 10.1088/1748-0221/9/12/P12006.

[127] T. Heindl et al. "The scintillation of liquid argon". *EPL (Europhysics Letters)* 91.6 (Sept. 2010), p. 62002. DOI: 10.1209/0295-5075/91/62002.

[128] S. Bonetti et al. "A study of the electron image due to ionizing events in a two-dimensional liquid argon TPC with a 24 cm drift gap". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 286.1 (Jan. 1990), pp. 135–146. DOI: 10.1016/0168-9002(90)90215-R.

[129] S. Amerio et al. "Design, construction and tests of the ICARUS T600 detector". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 527.3 (July 2004), pp. 329–410. DOI: 10.1016/j.nima.2004.02.044.

[130] A. Ankowski et al. "Measurement of through-going particle momentum by means of multiple scattering with the ICARUS T600 TPC". *The European Physical Journal C - Particles and Fields* 48.2 (Nov. 2006), pp. 667–676. DOI: 10.1140/epjc/s10052-006-0051-3.

[131] M. Antonello et al. "Experimental search for the "LSND anomaly" with the ICARUS detector in the CNGS neutrino beam". *The European Physical Journal C* 73.3 (Mar. 2013), p. 2345. DOI: 10.1140/epjc/s10052-013-2345-6.

[132] C. Anderson et al. "The ArgoNeuT detector in the NuMI low-energy beam line at Fermilab". *Journal of Instrumentation* 7.10 (Oct. 2012), P10019–P10019. DOI: 10.1088/1748-0221/7/10/P10019.

[133] R. Acciarri et al. “The Liquid Argon In A Testbeam (LArIAT) experiment”. *Journal of Instrumentation* 15.4 (Apr. 2020), P04026–P04026. DOI: 10.1088/1748-0221/15/04/P04026.

[134] E. Gramellini et al. *Measurement of the (π^-, Ar) total hadronic cross section at the LArIAT experiment.* arXiv:2108.00040. arXiv, July 2021. DOI: 10.48550/arXiv.2108.00040. arXiv: 2108.00040.

[135] Jonathan Asaadi et al. “First Demonstration of a Pixelated Charge Readout for Single-Phase Liquid Argon Time Projection Chambers”. *Instruments* 4.1 (Mar. 2020), p. 9. DOI: 10.3390/instruments4010009.

[136] D. A. Dwyer et al. “LArPix: demonstration of low-power 3D pixelated charge readout for liquid argon time projection chambers”. *Journal of Instrumentation* 13.10 (Oct. 2018), P10007–P10007. DOI: 10.1088/1748-0221/13/10/P10007.

[137] D. E. Groom and S. R. Klein. “Passage of particles through matter”. *The European Physical Journal C* 15.1 (Mar. 2000), pp. 163–173. DOI: 10.1007/BF02683419.

[138] C. Adams et al. “Reconstruction and measurement of $\mathcal{O}(100)$ MeV energy electromagnetic activity from $\pi^0 \rightarrow \gamma\gamma$ decays in the MicroBooNE LArTPC”. *Journal of Instrumentation* 15.2 (Feb. 2020), P02007–P02007. DOI: 10.1088/1748-0221/15/02/P02007.

[139] LArIAT Collaboration et al. “Calorimetry for low-energy electrons using charge and light in liquid argon”. *Physical Review D* 101.1 (Jan. 2020), p. 012010. DOI: 10.1103/PhysRevD.101.012010.

[140] R. Acciarri et al. “A study of electron recombination using highly ionizing particles in the ArgoNeuT Liquid Argon TPC”. *Journal of Instrumentation* 8.8 (Aug. 2013), P08005–P08005. DOI: 10.1088/1748-0221/8/08/P08005. arXiv: 1306.1712.

[141] M. Miyajima et al. “Average energy expended per ion pair in liquid argon”. *Physical Review A* 9.3 (Mar. 1974), pp. 1438–1443. DOI: 10.1103/PhysRevA.9.1438.

[142] R. Acciarri et al. “Effects of Nitrogen contamination in liquid Argon”. *Journal of Instrumentation* 5.6 (June 2010), P06003–P06003. DOI: 10.1088/1748-0221/5/06/P06003.

[143] M. Antonello et al. “Detection of Cherenkov light emission in liquid argon”. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 516.2 (Jan. 2004), pp. 348–363. DOI: 10.1016/j.nima.2003.08.177.

[144] Yichen Li et al. “Measurement of longitudinal electron diffusion in liquid argon”. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 816 (Apr. 2016), pp. 160–170. DOI: 10.1016/j.nima.2016.01.094.

[145] P. Abratenko et al. “Measurement of the longitudinal diffusion of ionization electrons in the MicroBooNE detector”. *Journal of Instrumentation* 16.9 (Sept. 2021), P09025. DOI: 10.1088/1748-0221/16/09/P09025.

[146] Michael Mooney. “The MicroBooNE Experiment and the Impact of Space Charge Effects” (Nov. 2015). arXiv: 1511.01563.

[147] G. M. Seidel, R. E. Lanou, and W. Yao. “Rayleigh scattering in rare-gas liquids”. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 489.1 (Aug. 2002), pp. 189–194. DOI: 10.1016/S0168-9002(02)00890-2.

[148] N. Ishida et al. “Attenuation length measurements of scintillation light in liquid rare gases and their mixtures using an improved reflection suppresser”. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 384.2 (Jan. 1997), pp. 380–386. DOI: 10.1016/S0168-9002(96)00740-1.

[149] Diego Garcia-Gamez, Patrick Green, and Andrzej M. Szelc. “Predicting transport effects of scintillation light signals in large-scale liquid argon detectors”. *The European Physical Journal C* 81.4 (Apr. 2021), p. 349. DOI: 10.1140/epjc/s10052-021-09119-3.

[150] R. Acciarri et al. “Noise Characterization and Filtering in the MicroBooNE Liquid Argon TPC”. *Journal of Instrumentation* 12.8 (Aug. 2017), P08003–P08003. DOI: 10.1088/1748-0221/12/08/P08003.

[151] B. Aimard et al. “A 4 tonne demonstrator for large-scale dual-phase liquid argon time projection chambers”. *Journal of Instrumentation* 13.11 (Nov. 2018), P11003–P11003. DOI: 10.1088/1748-0221/13/11/P11003.

[152] E. Segreto. “Evidence of delayed light emission of tetraphenyl-butadiene excited by liquid-argon scintillation light”. *Physical Review C* 91.3 (Mar. 2015), p. 035503. DOI: 10.1103/PhysRevC.91.035503.

[153] A. A. Machado and E. Segreto. “ARAPUCA a new device for liquid argon scintillation light detection”. *Journal of Instrumentation* 11.2 (Feb. 2016), pp. C02004–C02004. DOI: 10.1088/1748-0221/11/02/C02004.

[154] A. A. Machado et al. “The X-ARAPUCA: an improvement of the ARAPUCA device”. *Journal of Instrumentation* 13.4 (Apr. 2018), pp. C04026–C04026. DOI: 10.1088/1748-0221/13/04/C04026.

[155] Pedro A.N. Machado, Ornella Palamara, and David W. Schmitz. “The Short-Baseline Neutrino Program at Fermilab”. *Annual Review of Nuclear and Particle Science* 69.1 (2019), pp. 363–387. DOI: 10.1146/annurev-nucl-101917-020949.

[156] *SBN-doc-27037-v1: SBN Oscillation Sensitivity Technical Note (2021).* URL: https://sbn-docdb.fnal.gov/cgi-bin/sso/ShowDocument?docid=27037 (visited on 07/07/2022).

[157] MicroBooNE Collaboration et al. “First measurement of inclusive electron-neutrino and antineutrino charged current differential cross sections in charged lepton energy on argon in MicroBooNE”. *Physical Review D* 105.5 (Mar. 2022), p. L051102. DOI: 10.1103/PhysRevD.105.L051102.

[158] MINERνA Collaboration et al. “Constraint of the MINERνA medium energy neutrino flux using neutrino-electron elastic scattering”. *Physical Review D* 100.9 (Nov. 2019), p. 092001. DOI: 10.1103/PhysRevD.100.092001.

[159] The HARP CollaborationM.G. Catanesi et al. “Measurement of the production cross-section of positive pions in the collision of 8.9 GeV/c protons on beryllium”. *The European Physical Journal C* 52.1 (Sept. 2007), pp. 29–53. DOI: 10.1140/epjc/s10052-007-0382-8.

[160] J. Spitz. “Sterile neutrino search with kaon decay at rest”. *Physical Review D* 85.9 (May 2012), p. 093020. DOI: 10.1103/PhysRevD.85.093020.

[161] R. Acciarri et al. “Construction of precision wire readout planes for the Short-Baseline Near Detector (SBND)”. *Journal of Instrumentation* 15.6 (June 2020), P06033–P06033. DOI: 10.1088/1748-0221/15/06/P06033.

[162] Martin Auger et al. “A Novel Cosmic Ray Tagger System for Liquid Argon TPC Neutrino Detectors”. *Instruments* 1.1 (Dec. 2017), p. 2. DOI: 10.3390/instruments1010002.

[163] S. Agostinelli et al. “Geant4—a simulation toolkit”. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 506.3 (July 2003), pp. 250–303. DOI: 10.1016/S0168-9002(03)01368-8.

[164] C. Adams et al. "Ionization electron signal processing in single phase LArTPCs. Part I. Algorithm Description and quantitative evaluation with MicroBooNE simulation". *Journal of Instrumentation* 13.7 (July 2018), P07006–P07006. DOI: 10.1088/1748-0221/13/07/P07006.

[165] C. Adams et al. "Ionization electron signal processing in single phase LArTPCs. Part II. Data/simulation comparison and performance in MicroBooNE". *Journal of Instrumentation* 13.7 (July 2018), P07007–P07007. DOI: 10.1088/1748-0221/13/07/P07007.

[166] *GausHitFinder*. Dec. 2021. URL: https://github.com/LArSoft/larreco/blob/86b7154f11685e67d424b40fe1be97a41bb87473/larreco/HitFinder/GausHitFinder_module.cc (visited on 06/17/2022).

[167] F. Pedregosa et al. "Scikit-learn: Machine Learning in Python". *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830.

[168] Jerome Friedman, Trevor Hastie, and Robert Tibshirani. "Additive Logistic Regression: a Statistical View of Boosting". *Annals of Statistics* 28 (1998), p. 2000.

[169] Ian T. Jolliffe and Jorge Cadima. "Principal component analysis: a review and recent developments". *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 374.2065 (Apr. 2016), p. 20150202. DOI: 10.1098/rsta.2015.0202.

[170] *LArPandoraTrackCreation Module*. Dec. 2021. URL: https://github.com/LArSoft/larpandora/blob/master/larpandora/LArPandoraEventBuilding/LArPandoraTrackCreation_module.cc (visited on 05/05/2022).

[171] *Calorimetry Module*. Dec. 2021. URL: https://github.com/LArSoft/larpandora/blob/master/larpandora/LArPandoraEventBuilding/LArPandoraTrackCreation_module.cc (visited on 05/05/2022).

[172] R. Acciarri et al. "Michel electron reconstruction using cosmic-ray data from the MicroBooNE LArTPC". *Journal of Instrumentation* 12.9 (Sept. 2017), P09014–P09014. DOI: 10.1088/1748-0221/12/09/P09014.

[173] P. Abratenko et al. "Determination of muon momentum in the MicroBooNE LArTPC using an improved model of multiple Coulomb scattering". *Journal of Instrumentation* 12.10 (Oct. 2017), P10010–P10010. DOI: 10.1088/1748-0221/12/10/P10010.

[174] *Pandora Modular Shower Creation*. Dec. 2021. URL: https://github.com/LArSoft/larpandora/tree/master/larpandora/LArPandoraEventBuilding/LArPandoraShower (visited on 05/05/2022).

[175] R. Acciarri et al. “First Observation of Low Energy Electron Neutrinos in a Liquid Argon Time Projection Chamber”. *Physical Review D* 95.7 (Apr. 2017), p. 072005. DOI: 10.1103/PhysRevD.95.072005. arXiv: 1610.04102.

[176] Andreas Hocker et al. *TMVA - Toolkit for Multivariate Data Analysis with ROOT: Users guide. TMVA - Toolkit for Multivariate Data Analysis*. Geneva: CERN, 2007.

[177] Glen Cowan. *Statistical data analysis*. Oxford science publications. Oxford: Clarendon Press, 1998.

[178] Costas Andreopoulos et al. “The GENIE Neutrino Monte Carlo Generator: Physics and User Manual” (2015).

[179] C. Adams et al. “Calibration of the charge and energy loss per unit length of the MicroBooNE liquid argon time projection chamber using muons and protons”. *Journal of Instrumentation* 15.3 (Mar. 2020), P03022. DOI: 10.1088/1748-0221/15/03/P03022.

[180] P. Abratenko et al. “Novel approach for evaluating detector-related uncertainties in a LArTPC using MicroBooNE data”. *The European Physical Journal C* 82.5 (May 2022), p. 454. DOI: 10.1140/epjc/s10052-022-10270-8.

[181] *genie-mc.org*. URL: http://www.genie-mc.org/ (visited on 07/27/2022).

[182] MicroBooNE collaboration. *Neutrino Interaction Model and Uncertainties for MicroBooNE Analyses*. May 2020. URL: https://microboone.fnal.gov/wp-content/uploads/MICROBOONE-NOTE-1074-PUB.pdf.

Glossary

μBooNE Micro Booster Neutrino Experiment. iii, xi, 16, 26, 27, 32, 41–44, 54, 58, 61, 67, 77, 78, 174

dE/dx Energy loss per unit length. viii, xii, 33, 35, 57, 92–97, 100–102, 108, 114, 126, 128–130, 134, 141–143, 146

dQ/dx Charge deposited per unit length. 35, 94, 101

CORSIKA COsmic Ray SImulations for KAscade. 54, 58, 59, 70, 150, 151

FermiLab Fermi National Accelerator Laboratory. 50

GEANT4 GEometry ANd Tracking 4. 46, 56, 58, 59

GENIE GENIE event generator. iii, 17, 18, 21, 25–27, 44, 54–56, 161, 164, 171–173, 175, 177, 211

LArSoft Liquid Argon Software Package. iii, 54, 62, 92

MiniBooNE Mini Booster Neutrino Experiment. iii, 14, 15, 23, 32, 41–43, 65, 161

NEUT NEUT event generator. 25, 26

Acronyms

M_a Axial Mass. 17, 23

JSNS^2 J-PARC Sterile Neutrino Search at J-PARC Spallation Neutron Source. 15

ADC Analogue Digital Converter. 49

AGKY Andreopoulos-Gallagher-Kehayias-Yang. 19

ALEPH Apparatus for LEP PHysics. 4

AMU Atomic Mass Units. 155

APA Anode Plane Assembly. xi, 47–49, 51

ARAPUCA Argon R&D Advanced Program at UniCAmp. 39, 51, 52, 63

ArgoNeuT Argon Neutrino Teststand. 26, 27, 31, 32, 54, 94

ASIC Application-Specific Integrated Circuit. 49, 50

ATF Assembly Transport Fixture. 50

BDT Boosted Decision Tree. 75, 76, 88, 91, 126, 133–135, 137–139, 143, 145–148

BNB Booster Neutrino Beam. ii, iii, 1, 14, 15, 23, 25, 26, 32, 41, 42, 45–47, 52, 53, 65, 70, 72, 112, 131, 135, 149–151, 157, 162, 176, 210

BSM Beyond Standard Model. 42, 44, 54, 65

BY Bodek-Yang. 19, 165

CC Charged Current. ii, x–xii, 1, 2, 9, 10, 17, 26, 27, 30, 31, 44, 69, 72, 74, 110–112, 114, 115, 117–119, 121, 123–129, 132, 133, 137–142, 145–150, 154, 164, 165, 175–177, 211

CFG Correlated Fermi Gas. 20

CNGS CERN Neutrinos to Gran Sasso. 31

COH Coherent Scattering. 19, 55

CPA Cathode Plane Assembly. 47, 48, 50

CRT Cosmic Ray Tagger. 41, 47, 52–54, 60, 61, 64, 65, 88, 93, 110, 111, 119, 123–125

CV Cental Value. 159, 160, 162, 163, 165, 175

DAQ Data AcQuisition. 38, 50, 54

DCA Distance of Closest Approach. 64, 123, 134

DIS Deep Inealstic Scattering. 19, 55

DONUT Direct Observation of the NU Tau. 4

DUNE Deep Underground Neutrino Experiment. iii, 12, 27–29, 32, 38, 39, 52, 54

ES Electron Scatter. 9, 10

FPGA Field Programmable Gate Arrays. 50

FS Feynman Scaling. 162

FSI Final State Interactions. 21, 22, 24, 27, 55, 151

FV Fiducial Volume. 111–118, 121, 126, 129, 155, 156, 163, 171

HARP HAdron Production Experiment. 162

HIP Highly Ionising Particle. 130–132, 134

ICARUS Imaging Cosmic And Rare Underground Signals. 31, 32, 41, 42, 54

ILC International Linear Collider. 67

IMB Irvine Michigan Brookhaven. 5

JUNO Jiangmen Underground Neutrino Observatory. 12

Kamiokande Kamioka Neutrino Detection Experiment. 5

KATRIN Karlsruhe Tritium Neutrino Experiment. 10

KS test Kolmogorov-Smirnov Test. 135

LAr Liquid Argon. 38, 49

LAr TPC Liquid Argon Time Projection Chamber. ii, iii, x, 1, 26, 29–32, 36, 38–44, 47, 54, 64–67, 92, 95, 119, 120, 140, 141, 161, 176

LArIAT Liquid Argon In A Testbeam. 32

LEP Large Electron-Positron Collider. 4, 12

LFG Local Fermi Gas. 20, 171

LGC Langau-Gaussian Convolution. 34

LSND Liquid Scintillator Neutrino Detector. 12, 14, 15, 31, 32, 41–43

MC Monte Carlo. 9, 54, 59, 102, 153, 154

MCS Multiple Coulomb Scattering. 33, 94, 134, 135

MEC Meson Exchange Currents. 18, 21, 171, 174

MFP Mean Free Path. 159, 165

MINERνA Main Injector Neutrino ExpeRiment to study ν-A interactions. 24, 25

MINOS Main Injector Neutrino Oscillation Search. 15, 26, 31

MIP Minimum Ionising Particle. 33, 34, 100, 101, 124, 130–132, 134, 141–143

MPV Most Probable Value. 34

MVA Multi Variate Analysis. iii, 1, 66, 75, 77–80, 82, 84–88, 91, 107, 108, 119, 121, 125, 176

NC Neutral Current. 9, 10, 17, 44, 111, 114, 115, 117, 118, 121, 123–129, 132, 137, 139, 141, 142, 145–147, 149, 154, 164, 165, 211

ND280 Near Detector at 280 m. 25

NOνA NuMI Off-axis ν_e Appearance experiment. iii, 11, 12, 16

np-nh n-particle n-hole. 18, 23

NuMI Neutrinos at the Main Injector. 26, 31, 32, 42

PCA Principal Component Analysis. 82, 93, 97–100, 105–107

PCB Printed Circuit Board. 38

PDF Parton Distribution Fuction. 19

PDS Photon Detection System. 41, 47, 51, 61, 64, 65, 88, 110, 111, 119, 122, 123, 125, 161

PE Photo-Electrons. 63, 64, 124

PFO Particle Flow Object. 68, 70–72, 75, 80–82, 86, 88–93, 97, 107, 108

PID Particle IDentification. 23, 26, 31, 33, 36, 43, 60, 92–94, 109, 126, 129, 131–133, 137–139, 143, 145, 159

PMNS Pontecorvo-Maki-Nakagawa-Sakata. 6, 7, 12

PMT Photo-Multiplier Tube. 39, 51–54, 63, 122

POT Protons On Target. 44, 111, 112, 147, 151, 155, 164, 169, 172

PROSPECT Precision Reactor Oscillation and Spectrum Experiment. 16

QE Quasi-Elastic. 17–19, 23, 24, 55, 164, 171

RES Resonant Scattering. 19, 55, 164, 165

RFG Relativistic Fermi Gas. 20, 171

ROI Region Of Interest. 62

RPA Random Phase Approximation. 21, 23

SAGE Soviet–American Gallium Experiment. 5

SBN Short-Baseline Neutrino. ii–iv, 1, 15, 16, 26, 27, 29, 32, 40–43, 45, 65, 157

SBND Short-Baseline Near Detector. ii–iv, xi, 1, 2, 16, 19, 28, 32, 37, 41–47, 50–56, 58, 60, 62, 65, 66, 70, 72, 77, 78, 80, 86, 92, 95, 107–110, 119, 121, 122, 131, 140, 147, 150, 151, 156, 161–164, 166, 169, 171, 174–177

SCE Space Charge Effect. 36, 59, 155, 161

SERDES SERializer/DESerializer. 50

SiPM Silicon Photo-Multiplier. 39, 52, 60, 64

SK Super-Kamiokande. 8, 9, 11–13, 15

SNO Sudbury Neutrino Observatory. 9–11

SNR Signal-to-Noise Ratio. 38, 62, 64

SRC Short Range Correlations. 20

SSM Standard Solar Model. 9, 10

STEREO Search for Sterile Reactor Neutrino Oscillations. 16

SW Sanford-Wang. 162, 164, 174

T2HK Tokai to Hyper-Kamiokande. 12

T2K Tokai to Super-Kamiokande. 11, 12, 16, 25

TPB TetraPhenyl Butadiene. 37, 39, 47, 50, 51, 122

TPC Time Projection Chamber. 38, 41, 47–50, 52, 54, 56, 61, 63–65, 110–114, 119, 121, 122

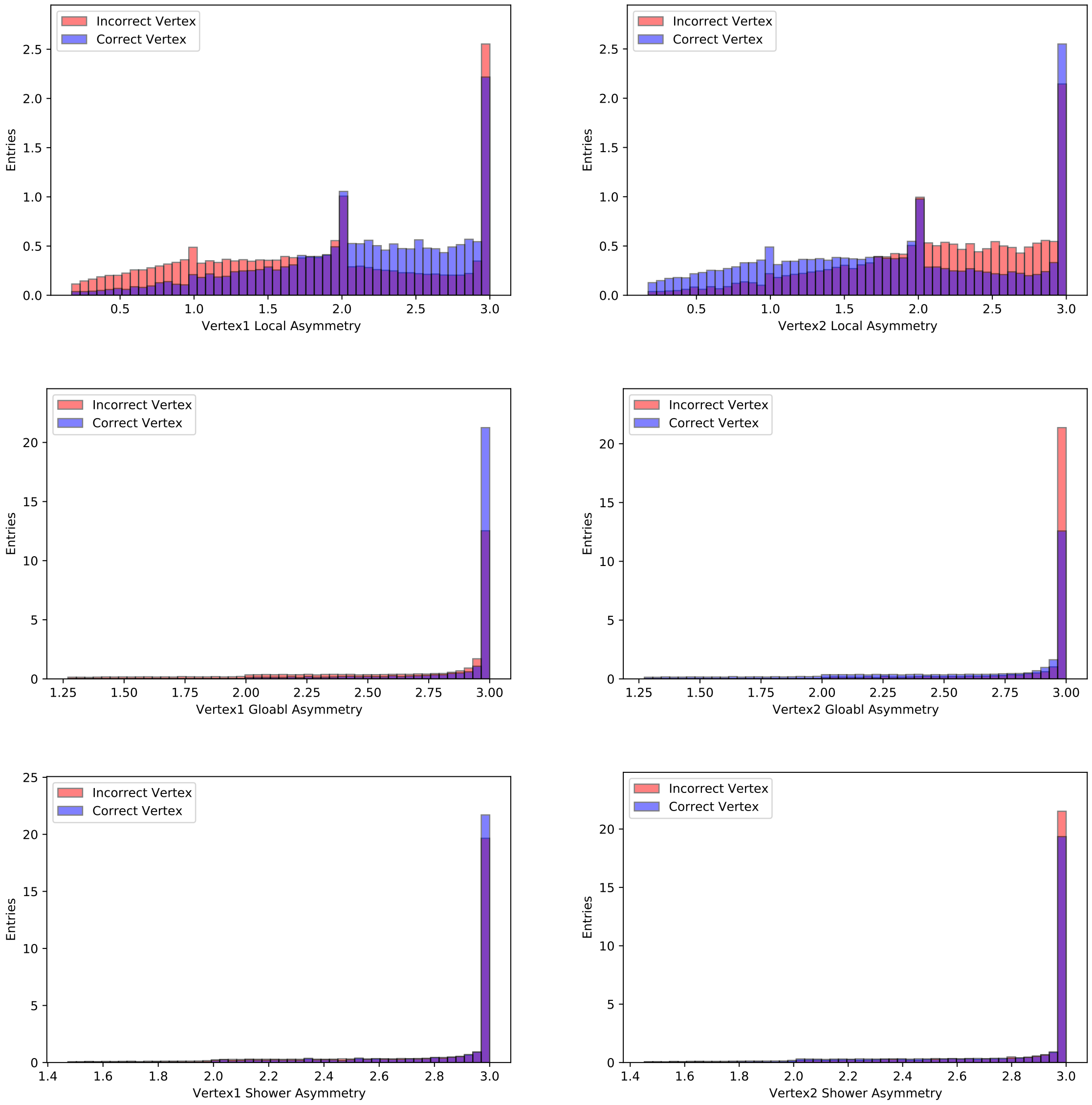
VUV Vacuum Ultra-Violet. 34, 39, 51, 122

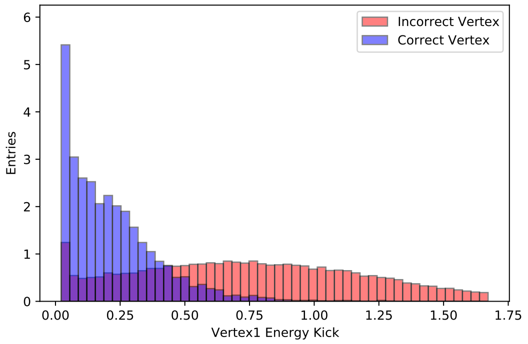
WIB Warm Interface Board. 50

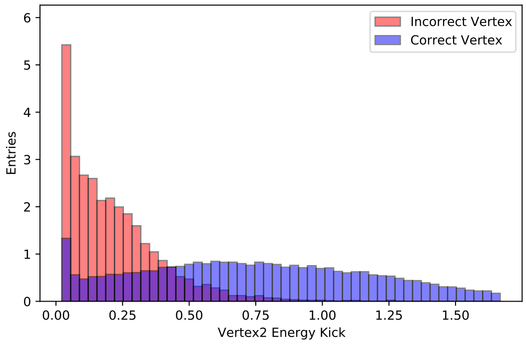
Appendix A

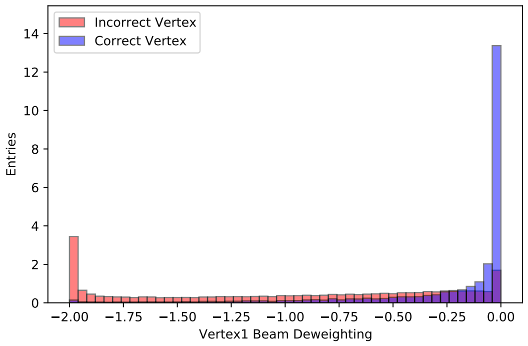
Vertex Selection MVA

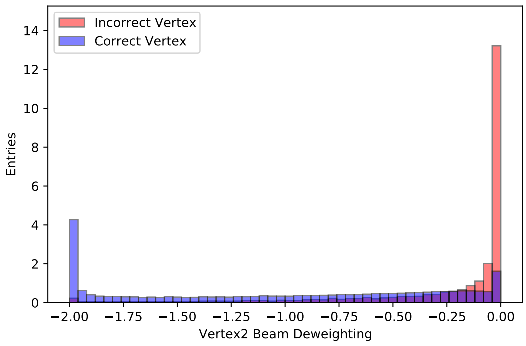
Figure A.1: Input variables used in the Pandora Vertex Selection MVA.

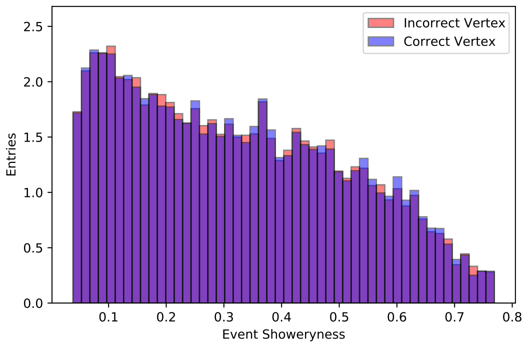


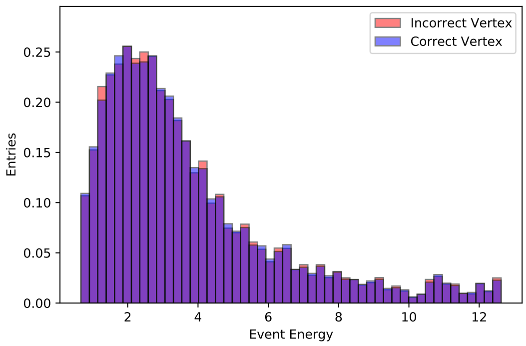

Incorrect Vertex
Correct Vertex
Entries
0
1
2
3
4
5
6
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
Vertex1 Energy Kick

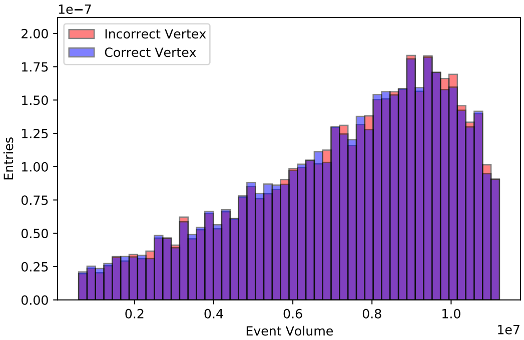

Incorrect Vertex
Correct Vertex
Entries
0
1
2
3
4
5
6
0.00
0.25
0.50
0.75
1.00
1.25
1.50
Vertex2 Energy Kick

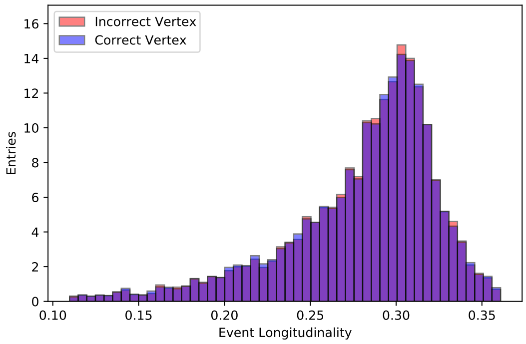

Incorrect Vertex
Correct Vertex
Entries
0
2
4
6
8
10
12
14
−2.00
−1.75
−1.50
−1.25
−1.00
−0.75
−0.50
−0.25
0.00
Vertex1 Beam Deweighting

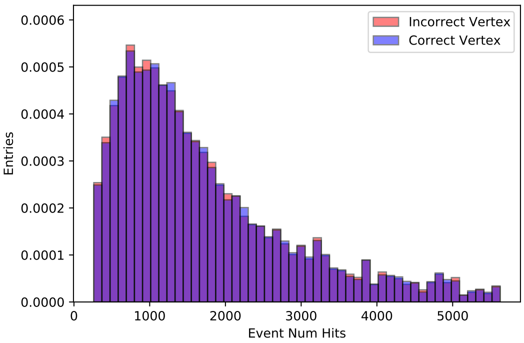

Incorrect Vertex
Correct Vertex
Entries
0
2
4
6
8
10
12
14
−2.00
−1.75
−1.50
−1.25
−1.00
−0.75
−0.50
−0.25
0.00
Vertex2 Beam Deweighting

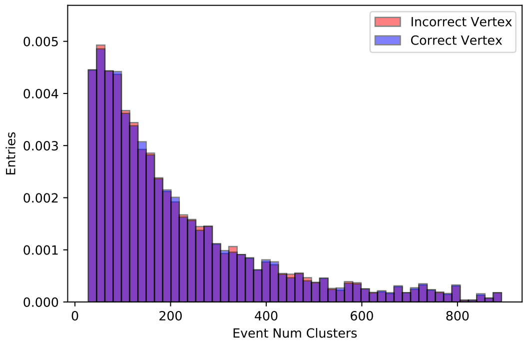

Incorrect Vertex
Correct Vertex
Entries
0.0
0.5
1.0
1.5
2.0
2.5
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
Event Showeryness

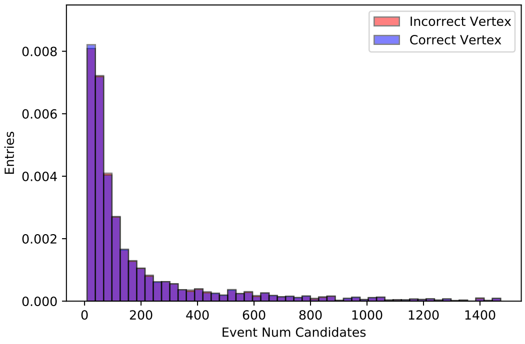

Incorrect Vertex
Correct Vertex
Entries
0.00
0.05
0.10
0.15
0.20
0.25
2
4
6
8
10
12
Event Energy


1e−7
Incorrect Vertex
Correct Vertex
Entries
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
0.2
0.4
0.6
0.8
1.0
1e7
Event Volume


Incorrect Vertex
Correct Vertex
Entries
16
14
12
10
8
6
4
2
0
0.10
0.15
0.20
0.25
0.30
0.35
Event Longitudinality


Incorrect Vertex
Correct Vertex
Entries
0.0006
0.0005
0.0004
0.0003
0.0002
0.0001
0.0000
0
1000
2000
3000
4000
5000
Event Num Hits

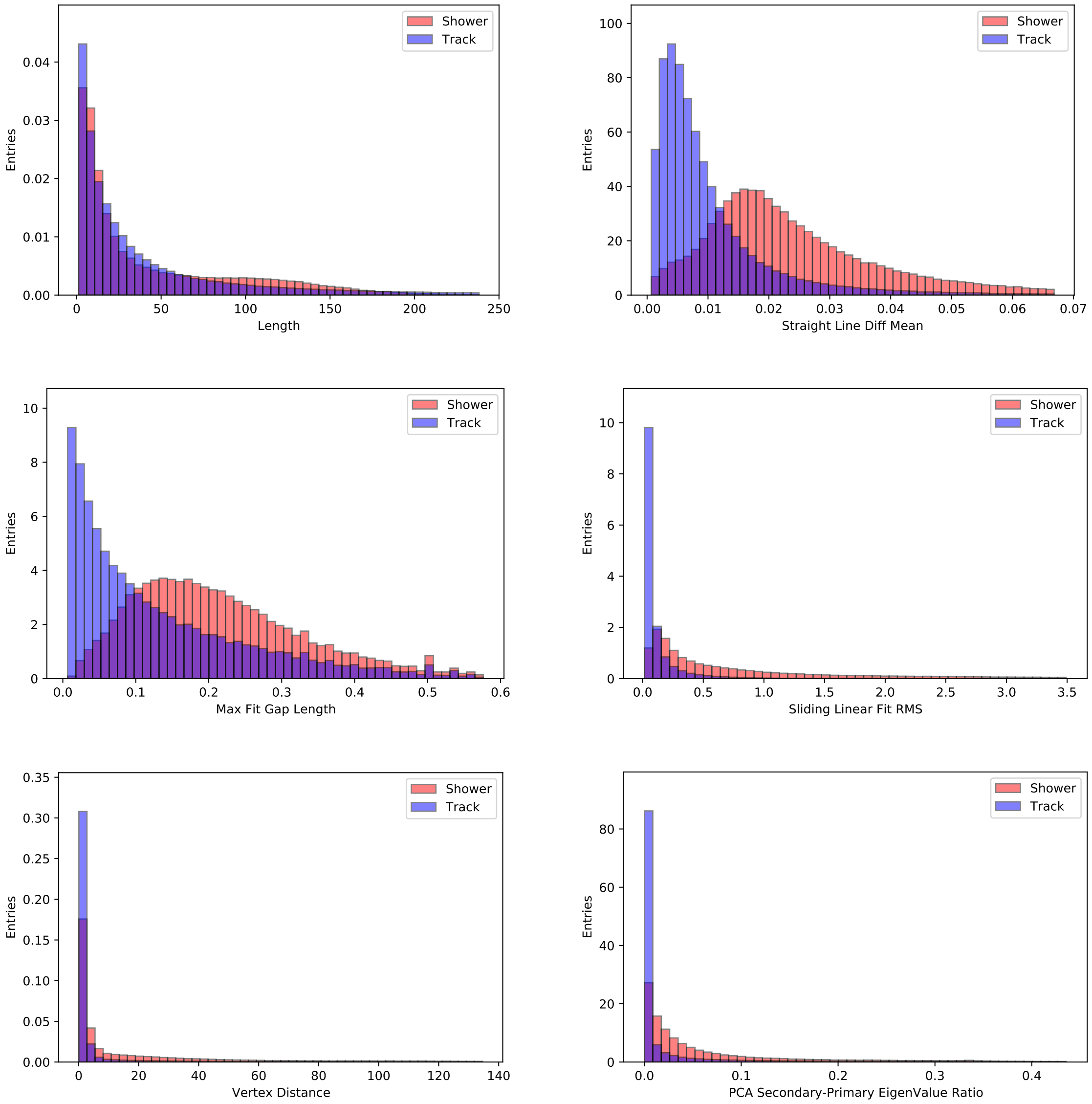

Incorrect Vertex
Correct Vertex
Entries
0.005
0.004
0.003
0.002
0.001
0.000
0
200
400
600
800
Event Num Clusters

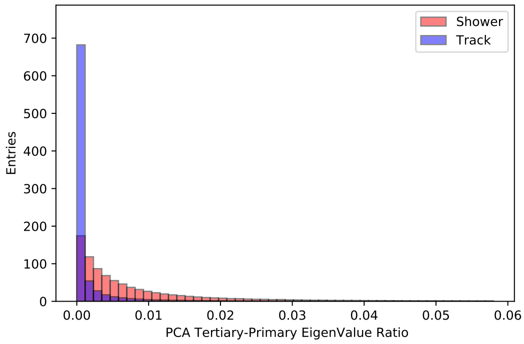

Incorrect Vertex
Correct Vertex
Entries
0.008
0.006
0.004
0.002
0.000
0
200
400
600
800
1000
1200
1400
Event Num Candidates

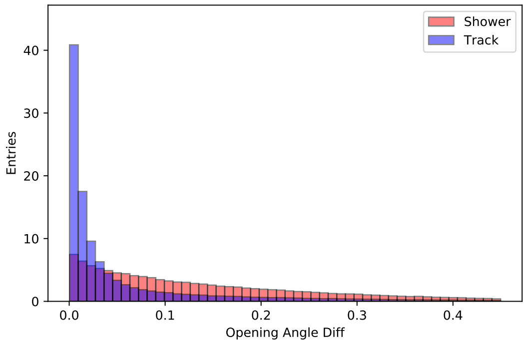
Appendix B

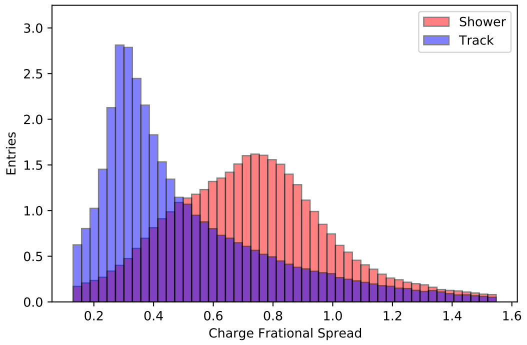
PFO Characterisation MVA

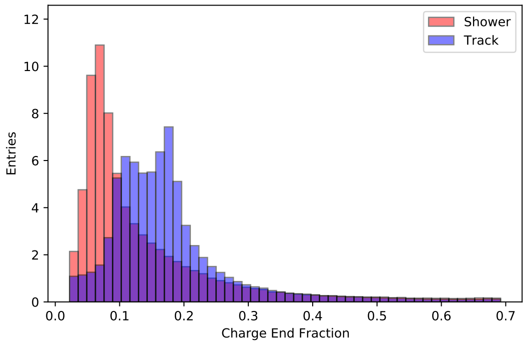
Figure B.1: Input variables used in the Pandora PFO Characterisation MVA.




Shower
Track
Entries
700
600
500
400
300
200
100
0
0.00
0.01
0.02
0.03
0.04
0.05
0.06
PCA Tertiary-Primary EigenValue Ratio


Shower
Track
Entries
40
30
20
10
0
0.0
0.1
0.2
0.3
0.4
Opening Angle Diff

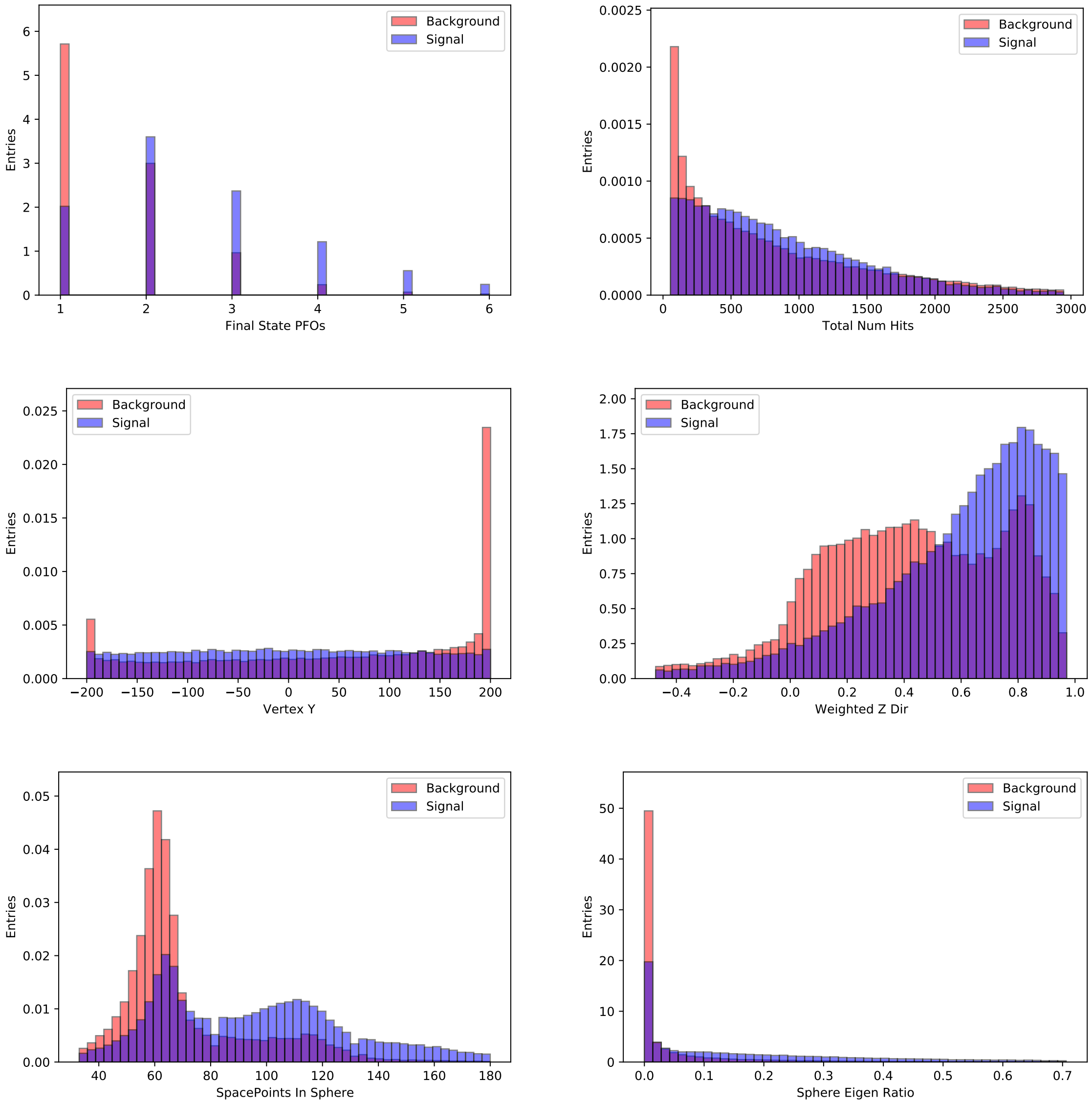

Shower
Track
Entries
3.0
2.5
2.0
1.5
1.0
0.5
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
Charge Frational Spread

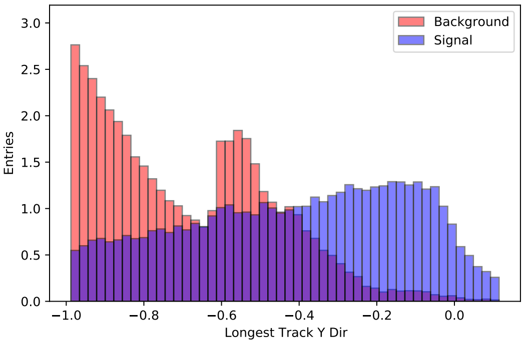

Shower
Track
Entries
12
10
8
6
4
2
0
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
Charge End Fraction

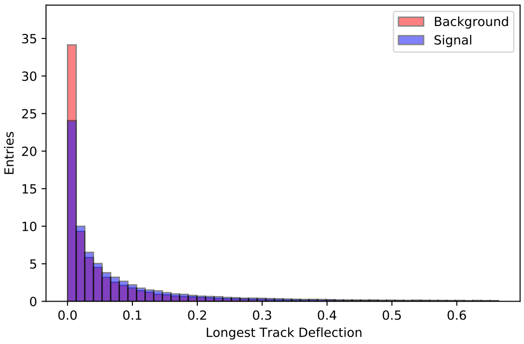
Appendix C

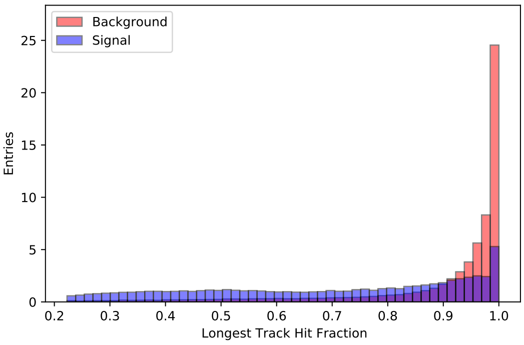
Slice ID MVA

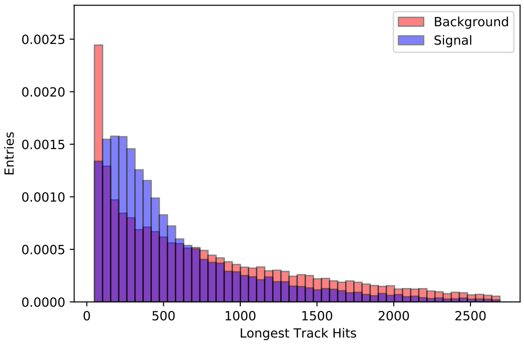
Figure C.1: Input variables used in the Pandora Slice ID MVA.




Background
Signal
Entries
3.0
2.5
2.0
1.5
1.0
0.5
0.0
−1.0
−0.8
−0.6
−0.4
−0.2
0.0
Longest Track Y Dir


Background
Signal
Entries
35
30
25
20
15
10
5
0
0.0
0.1
0.2
0.3
0.4
0.5
0.6
Longest Track Deflection


Background
Signal
Entries
25
20
15
10
5
0
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
Longest Track Hit Fraction


Background
Signal
Entries
0.0025
0.0020
0.0015
0.0010
0.0005
0.0000
0
500
1000
1500
2000
2500
Longest Track Hits

Appendix D

Dazzle Track PID

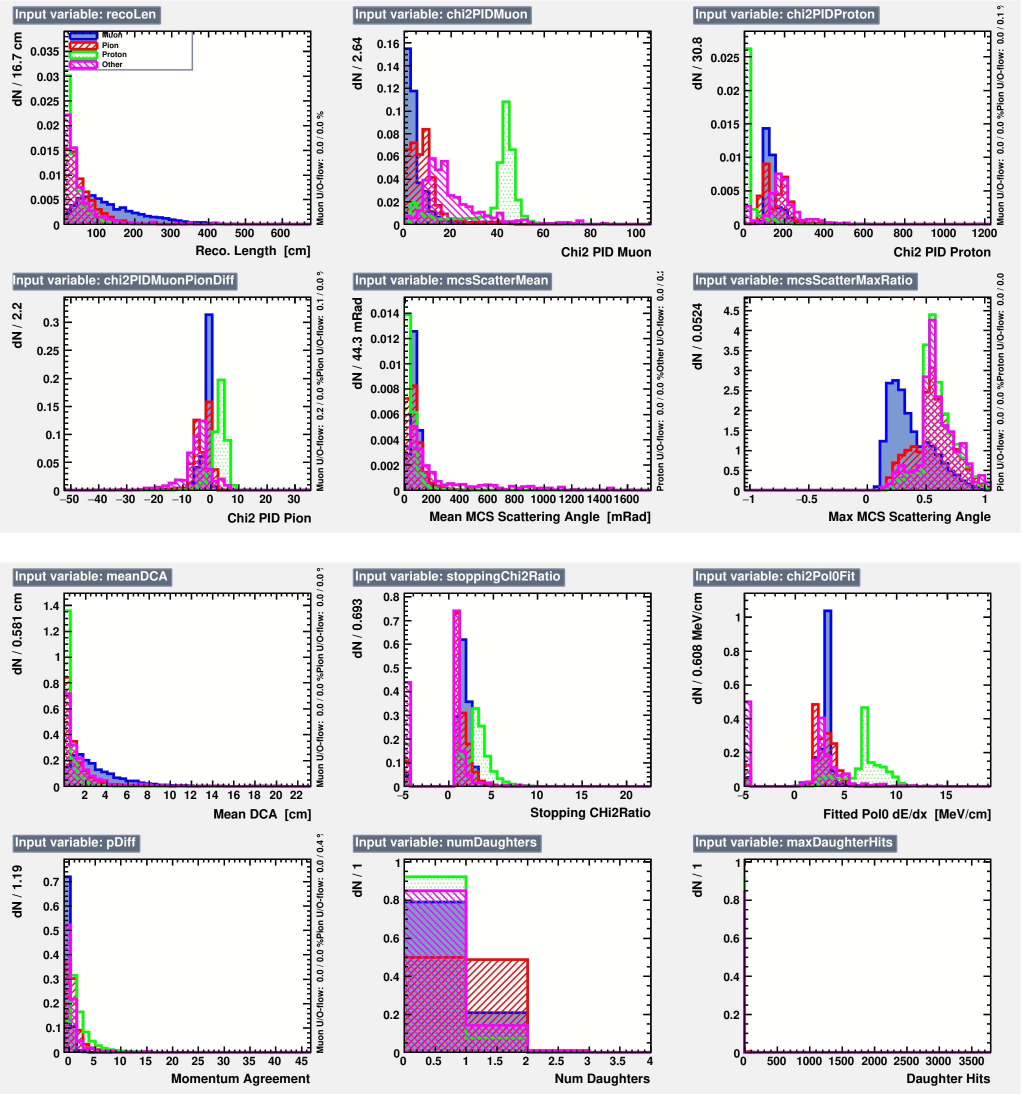


Figure D.1: Input variables used in the Dazzle Track PID MVA.

Appendix E

Razzle Shower PID

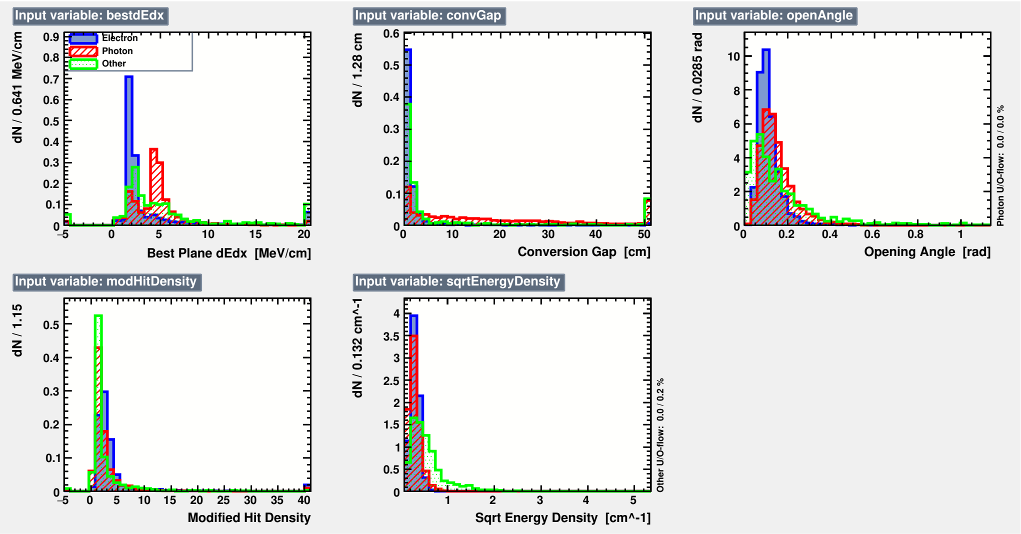


Figure E.1: Input variables used in the Razzle Shower PID MVA.

Appendix F

Cross Section Uncertainties

F.1 Flux Uncertainties

Parameter	Description	$\delta P/P$
Horn Current	Current pulsed through the focusing horn	±1 kA
Skin Current	Current induced on the surface of the horn	On/Off
σ_{π^+}	π^+ production in the target	Spline Variation
σ_{π^-}	π^- production in the target	Spline Variation
σ_{k^+}	k^+ production in the target	Feynman Scaling
σ_{k^0}	k^0 production in the target	Sanford Wang
σ_{k^-}	k^- production in the target	±100% Normalisation
σ^{π}_{Tot}	Total interaction cross section for pions	±11.9 mb
σ^{π}_{QE}	Quasi-Elastic interaction cross section for pions	±11.2 mb
σ^{π}_{Ine}	Inelastic interaction cross section for pions	±10 mb
σ^{n}_{Tot}	Total interaction cross section for nucleons	±15 mb
σ^{n}_{QE}	Quasi-Elastic interaction cross section for nucleons	±20 mb
σ^{n}_{Ine}	Inelastic interaction cross section for nucleons	±5 mb

Table F.1: Flux systematic uncertainties from the BNB simulation[2].

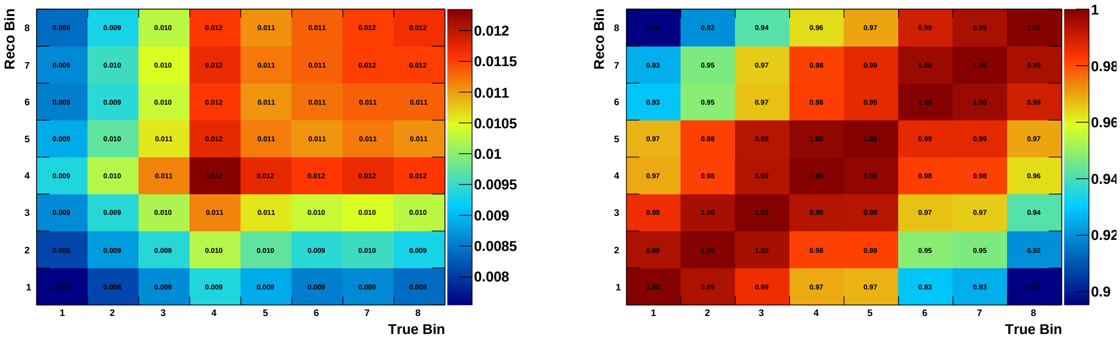
F.2 Interaction Uncertainties

Parameter	Description	$\delta P/P$
M_A^{CCQE}	Axial mass for CC Quasi-Elastic interactions	±10%
M_A^{NCE}	Axial mass for Neutral Current (NC) Elastic interactions	±25%
η^{NCE}	Strange axial form factor η for NC Elastic interactions	±30%
M_A^{CCRes}	Axial mass for CC Resonant interactions	±20%
M_A^{NCRes}	Axial mass for NC Resonant interactions	±20%
M_V^{CCRes}	Vector mass for CC Resonant interactions	±10%
M_V^{NCRes}	Vector mass for NC Resonant interactions	±10%
$N_{Bk}^{\nu p,CC1\pi}$	Non-resonant background normalisation in ν-p CC 1π interactions	±50%
$N_{Bk}^{\nu n,CC1\pi}$	Non-resonant background normalisation in ν-n CC 1π interactions	±50%
$N_{Bk}^{\nu p,CC2\pi}$	Non-resonant background normalisation in ν-p CC 2π interactions	±50%
$N_{Bk}^{\nu n,CC2\pi}$	Non-resonant background normalisation in ν-n CC 2π interactions	±50%
$N_{Bk}^{\nu n,NC1\pi}$	Non-resonant background normalisation in ν-n NC 1π interactions	±50%
$N_{Bk}^{\nu p,NC1\pi}$	Non-resonant background normalisation in ν-p NC 1π interactions	±50%
$N_{Bk}^{\nu n,NC2\pi}$	Non-resonant background normalisation in ν-n NC 2π interactions	±50%
$N_{Bk}^{\nu p,NC2\pi}$	Non-resonant background normalisation in ν-p NC 2π interactions	±50%
$N_{Bk}^{\overline{\nu}p,CC1\pi}$	Non-resonant background normalisation in $\overline{\nu}$-p CC 1π interactions	±50%
$N_{Bk}^{\overline{\nu}n,CC1\pi}$	Non-resonant background normalisation in $\overline{\nu}$-n CC 1π interactions	±50%
$N_{Bk}^{\overline{\nu}p,CC2\pi}$	Non-resonant background normalisation in $\overline{\nu}$-p CC 2π interactions	±50%
$N_{Bk}^{\overline{\nu}n,CC2\pi}$	Non-resonant background normalisation in $\overline{\nu}$-n CC 2π interactions	±50%
$N_{Bk}^{\overline{\nu}n,NC1\pi}$	Non-resonant background normalisation in $\overline{\nu}$-n NC 1π interactions	±50%
$N_{Bk}^{\overline{\nu}p,NC1\pi}$	Non-resonant background normalisation in $\overline{\nu}$-p NC 1π interactions	±50%
$N_{Bk}^{\overline{\nu}n,NC2\pi}$	Non-resonant background normalisation in $\overline{\nu}$-n NC 2π interactions	±50%
$N_{Bk}^{\overline{\nu}p,NC2\pi}$	Non-resonant background normalisation in $\overline{\nu}$-p NC 2π interactions	±50%
A^{HT}	A Higher Twist parameter in Bodek-Yang model	±25%
B^{HT}	B Higher Twist parameter in Bodek-Yang model	±25%
$Cv1u$	Valence PDF correction factor Cv1u in Bodek-Yang model	±30%
$Cv2u$	Valence PDF correction factor Cv2u in Bodek-Yang model	±40%
λ_{Tot}^{n}	Intranuclear Nucleon total mean free path	±20%
P_{CEx}^{n}	Intranuclear Charge Exchange probability for nucleons	±50%
P_{Ine}^{n}	Intranuclear Inelastic scatter probability for nucleons	±40%
P_{π}^{n}	Intranuclear π production probability for nucleons	±20%
λ_{Tot}^{π}	Intranuclear Pion total mean free path	±20%
P_{CEx}^{π}	Intranuclear Charge Exchange probability for pions	±50%
P_{Ine}^{π}	Intranuclear Inelastic scatter probability for pions	±40%
P_{π}^{π}	Intranuclear π production probability for pions	±20%

Table F.2: Interaction systematic uncertainties provided by GENIE[178].

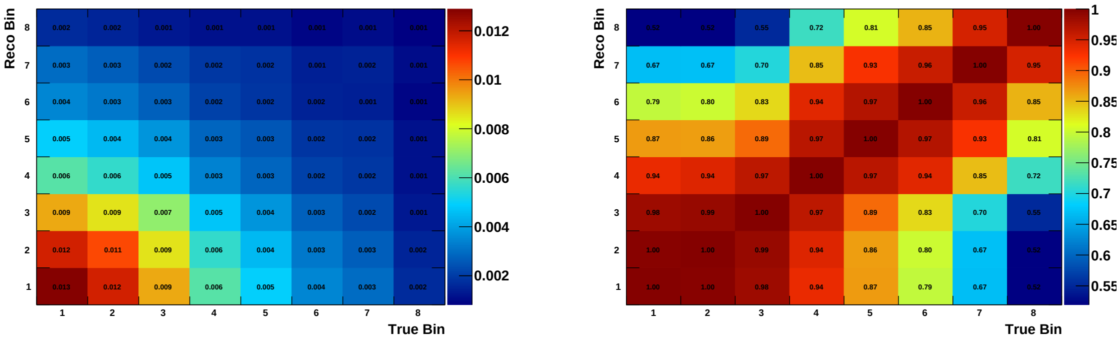
Appendix G

Cross Section Covariance Matrices



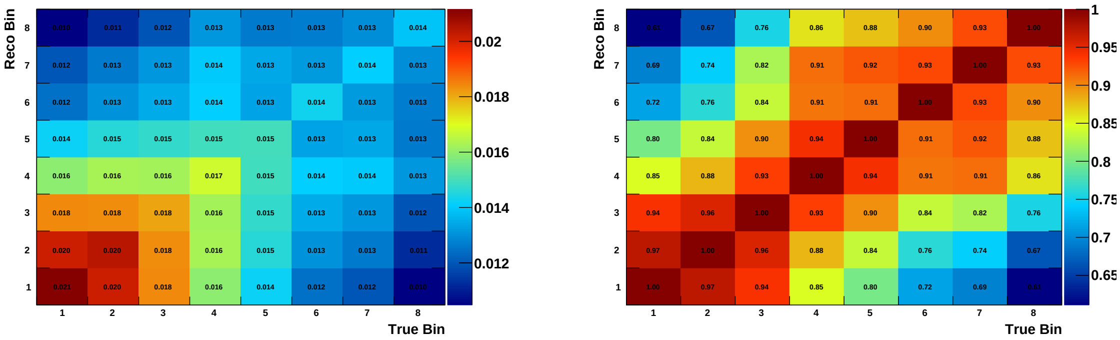
(a) Fractional Covariance Matrix (b) Correlation Matrix

Figure G.1: Covariance Matrices for the Flux Uncertainties on the Extracted Cross Section.



(a) Fractional Covariance Matrix (b) Correlation Matrix

Figure G.2: Covariance Matrices for the Interaction Uncertainties on the Extracted Cross Section.



Reco Bin	1	2	3	4	5	6	7	8
8	0.010	0.011	0.012	0.013	0.013	0.013	0.013	0.014
7	0.012	0.013	0.013	0.014	0.013	0.013	0.014	0.013
6	0.012	0.013	0.013	0.014	0.013	0.014	0.013	0.013
5	0.014	0.015	0.015	0.015	0.015	0.013	0.013	0.013
4	0.016	0.016	0.016	0.017	0.015	0.014	0.014	0.013
3	0.018	0.018	0.018	0.016	0.015	0.013	0.013	0.012
2	0.020	0.020	0.018	0.016	0.015	0.013	0.013	0.011
1	0.021	0.020	0.018	0.016	0.014	0.012	0.012	0.010

(a) Fractional Covariance Matrix

Reco Bin	1	2	3	4	5	6	7	8
8	0.61	0.67	0.76	0.86	0.88	0.90	0.93	1.00
7	0.69	0.74	0.82	0.91	0.92	0.93	1.00	0.93
6	0.72	0.76	0.84	0.91	0.91	1.00	0.93	0.90
5	0.80	0.84	0.90	0.94	1.00	0.91	0.92	0.88
4	0.85	0.88	0.93	1.00	0.94	0.91	0.91	0.86
3	0.94	0.96	1.00	0.93	0.90	0.84	0.82	0.76
2	0.97	1.00	0.96	0.88	0.84	0.76	0.74	0.67
1	1.00	0.97	0.94	0.85	0.80	0.72	0.69	0.61

(b) Correlation Matrix

Figure G.3: Covariance Matrices for the Total Uncertainties on the Extracted Cross Section.