Preparing for Precision: Neutral-Current Neutral Pion Production in the Short-Baseline Near Detector

Henry Gurton Lay, MPhys (Hons)

Department of Physics

Lancaster University

A thesis submitted for the degree of

Doctor of Philosophy

January, 2025

Preparing for Precision: Neutral-Current Neutral Pion Production in the Short-Baseline Near Detector

Henry Gurton Lay, MPhys (Hons).

Department of Physics, Lancaster University

A thesis submitted for the degree of *Doctor of Philosophy.* January, 2025.

Abstract

Precision measurements of neutrino interaction cross sections will be of critical importance to the next generation of neutrino oscillation experiments. The Short-Baseline Near Detector (SBND) is a 112 ton Liquid Argon Time Projection Chamber experiment in the Booster Neutrino Beam at the Fermi National Accelerator Laboratory in Illinois. It will collect the largest ever dataset of neutrino-argon interactions with which it will pursue a rich cross section measurement program. This thesis presents the development of a number of tools for the reconstruction and analysis of this data to prepare SBND for precision measurements. The ability of SBND to make a measurement of the challenging NC $1\pi^0$ channel is then explored. This channel is of particular interest due to its similar signature to the charged-current electron neutrino interactions that represent the signal of interest in the MiniBooNE and LSND low energy excess. SBND will investigate this excess as part of the Short-Baseline Neutrino program. Powerful selections reduce the background to this channel by over 99 %, whilst retaining 34 % of the signal. A method for extracting the interaction cross section is then developed, including an assessment of a range of systematic uncertainties associated with the simulation chain. By comparing the extracted cross section from a sample of simulated SBND data to the raw predictions from a couple of different generator models, it is clear that significant differences will be resolvable with SBND data. Critical further work is detailed that will enable full exploitation of the generational advancement in precision measurements of neutrino-argon interaction cross sections possible with SBND.

Acknowledgements

Four years of your life is a *really long time*, and PhD life comes with a few challenges to overcome. I could write a whole chapter on the hundreds of people who deserve my thanks for getting me here. I had to promise not to do so, so here is the abridged version.

I'll start in Lancaster with all the members of our fantastic LAr group, past and present. In particular, Andy, you spent years convincing me not to abandon physics before we even got to the start of this PhD. Thanks for backing me, for spending a summer teaching me ROOT and for providing me with years of wise supervision and guidance. Jarek, there is no chance I would have done any of this without a brilliant master's year in which you introduced me to SBND, to LArSoft and to the wonderful world of neutrino interactions. Thanks for your reminders on focus and prioritisation, I'd still be on chapter 1 without them! To the inhabitants of B06a, Hannah, Ruby, Niam, Ryan, Chris, Ishanee, Beth, Rachel, Adam, Rebecca, Beltran - the constant support and distraction was a pleasure throughout, *it does in fact tea*. Elliot, you get a special mention for being a consistent buddy from the moment they let us into the office post-covid, from F1 dissections to the darticle gun, I'll miss it all!

Dom, you shouldn't be surprised to get your own paragraph. I, and this thesis, owe an incredible amount to you. Your seemingly limitless knowledge of the LAr codebase is only matched by your patience with young students. You taught me so much whilst your trademark cynicism never failed to make me smile, even on the worst of PhD days. You're not rid of me yet, I intend to continue bugging you for a long time to come.

I feel genuinely honoured to have been part of the incredible SBND community these past few years, most impactfully whilst living at Fermilab. I'd especially like to thank Ornella Palamara for being a great mentor, Joseph Zennamo for never failing to make every last one of us feel valuable, Lauren Yates for always looking out for the students, Roberto Acciarri, Monica Nunes and Anne Schukraft for patiently answering my many stupid questions at SBN-ND, Dave Schmitz for a healthy balance of questions that make you think and pints for you to drink, and, of course, Michelle Stancari for trusting this software nerd with

a handful of cables and CRT modules and teaching me quite literally everything I know about hardware.

I was so lucky to have such an amazing cohort of fellow SBND students starting out together. To Lan, Jiaoyang, Fran, Rodrigo, Lynn, Mun, Bear and Lane - I consider you all friends for life. You made the daunting world of PhD studenting a much more enjoyable and supportive place, my only regret is how little time we spent all on the same continent! Lan and Jiaoyang - we were so lucky to have our core UK SBND crew from the start, I would never have got to the end of this without all the time we spent sharing rants, pepping each other up and laughing at the ridiculousness of all of this! I'd also like to thank all the other amazing people I met living at Fermilab. Tania, the User's Centre is what brings us all together, the Fermilab community owes so much to you! In particular to Namitha Chithirasreemadam, thanks for being the first friendly face way back in our dorm 5 era and then a constant friend through your many cameo appearances at the lab. Also to Veera Mikola, thanks for being a brilliant friend, my knowledge (and playlists) of Eurovision qualifying rounds have been greatly enhanced by our friendship.

Despite the best attempts of academia to fill all gaps, work is not everything. Thanks to the members of Pilling Jubilee Silver Band, the Fox Valley Brass Band and all the other brilliant musicians in both Lancaster and Illinois for filling my evenings and weekends with the best company and fantastic performances. Randy Micheletti and Mark Fenne, you really took me under your wings when I turned up at FVBB, thanks for taking me into your homes and for all our hilarious cultural comparisons. Tom Fell you also get a mention here, as much as I genuinely love running on my own, our silly adventures have been some of my true highlights of the last few years. Thanks for the motivation and companionship. On the subject of silly adventures, Charlie, you're a wonderful friend and the person I can always rely on to have a madder suggestion, thanks for encouraging the optimistic idiot in me, to Harry's eternal exasperation. Harry, thank you for being such a close friend for so many years. Moving to Southampton when I was still in Lancaster was frankly just rude, but I'm always delighted to see you and am looking forward to many more years setting the world to rights over a few craft beers. The rest of Newton's illegitimate children also deserve a mention, it is rare to find such a fantastically mad bunch of people, even rarer to

hold on to that once scattered across the country - thank you all for the effort that takes, it's well worth it! To Dibs, coping living me with four years was already impressive, adding a further three even more so. You're by far the smartest man I know, I will deeply miss our hours of aimless conversation sustained by copious tea and our shared obsession with (useless?) facts.

To my entire wonderful family, thank you all for supporting me, encouraging me and keeping me grounded. Mum and Dad, there are a million ways in which you made it possible for me to be here and I'm grateful for every single one. To Granny, Grampy and Rosie, one day I'll get a proper job but today is not that day! Granny G, we lost you whilst I was writing the final version of these acknowledgements, I'll never get to hear what you think I look like in a floppy hat, but I can certainly imagine!

Finally, and most importantly, to Meg. There are a thousand cliches I could write here but none of them would suffice. Simply, thanks for being you!

Declaration

I declare that the work presented in this thesis is, to the best of my knowledge and belief, original and my own work. The material has not been submitted, either in whole or in part, for a degree at this, or any other university. This thesis does not exceed the maximum permitted word length of 80,000 words including appendices and footnotes, but excluding the bibliography. A rough estimate of the word count is: 51521

Henry Gurton Lay

Contents

List of Tables

List of Figures

Glossary

A non-exhaustive list of the many acronyms used in this thesis - many experiment names have been omitted as their somewhat contrived acronyms are not illuminating.

SBND	Short-Baseline Near Detector
SBN	Short-Baseline Neutrino
BNB	Booster Neutrino Beam
LAr	Liquid Argon
PDG	Particle Data Group
CERN	Conseil européen pour la Recherche nucléaire (European Organization for Nuclear Research)
SLAC	Stanford Linear Accelerator Centre
FNAL	Fermi National Accelerator Laboratory
PMNS	Pontecorvo-Maki-Nakagawa-Sakata
CP	Charge-Parity
CC	Charged-Current
NC	Neutral-Curret
RES	Resonance
DIS	Deep Inelastic Scattering
QE	Quasi-Elastic
FSI	Final State Interactions
LEE	Low Energy Excess
SCE	Space Charge Effect
LNGS	Laboratori Nazionali del Gran Sasso (Gran Sasso National Laboratory)
CNGS	CERN Neutrinos to Gran Sasso

CRT	Cosmic Ray Tagger
PDS	Photon Detection System
TPC	Time Projection Chamber
PMT	Photomultiplier Tube
LArSoft	Liquid Argon Software
PFO	Particle Flow Object
BDT	Boosted Decision Tree
MVA	Multi-Variate Analysis
PD	Photodetector
CRUMBS	Cosmic Rejection Using Multi-system BDT Score
PID	Particle Identification
ROC	Receiver-Operator-Characteristics
MC	Monte Carlo
POT	Protons on Target
CRT♭	"CRT Flat" - the CRT modules under the SBND cryostat
CRT##	"CRT Sharps" - the CRT and DAQ teststand at SBN-ND 2022-2023

Chapter 1

Introduction

The opening line of Kane's Modern Elementary Particle Physics [1] reads:

"The Standard Model of particle physics is an awesome theory, providing a description and explanation of the world we see, in a full relativistic quantum field theory. It leaves no puzzles in its domain. It achieves the goals of four centuries of physics."

Similar lines can be found in the opening of almost all modern particle physics textbooks [2, 3, 4]. The Standard Model is indeed a remarkable theory which, over 40 years after its codification continues to be reinforced by more precise tests of its most obscure predictions [5]. So far particle physicists have only found one definitive piece of evidence for physics beyond this standard model, neutrino mass. Results around the turn of the century from the SNO [6] and Super-Kamiokande [7] collaborations conclusively demonstrated the existence of neutrino oscillations. A phenomenon that relies on mixing between neutrino states and the existence of neutrino masses, the latter deviating from the massless prediction of the Standard Model. This discovery won the 2015 Nobel Prize and ushered in an exciting era of precision neutrino physics pushing the boundaries of our most accurate model yet.

Chapter 2 contains a brief overview of the developments in the almost 100 years of neutrino physics followed by a description of the theories underpinning neutrino interactions

and specifically the neutral-current interactions relevant to this thesis in chapter 3. Chapter 4 describes the general operating principle of liquid argon time projection chambers (LArTPCs) as well as the specifics of the Short-Baseline Near Detector (SBND) within the Short-Baseline Neutrino (SBN) program. The processing of SBND Monte Carlo via the simulation of neutrino interactions, propagation of particles in the detector medium, simulation of the detector and electronics responses, and reconstruction is all covered in chapter 5. Chapter 6 describes the development and utilisation of a tool used to reject cosmic-ray backgrounds. Finally, chapters 7 and 8 assess SBND's ability to select and reconstruct NC $1\pi^0$ events and use them to make a differential measurement of the cross section.

Chapter 2

Neutrinos

2.1 Postulation and Discovery

2.1.1 Little Neutral One

Wolfgang Pauli was the first to, grudgingly, suggest the existence of the neutrino in 1930 [8]. Experiments studying the kinematics of beta decay had demonstrated a continuous energy spectrum of the emitted electron [9], a result totally incompatible with the predicted two-body decay channel. Pauli's letter to the Tübingen conference suggested a solution could be found by introducing a neutral, spin-$\frac{1}{2}$ third body termed the 'neutron'. It was Enrico Fermi who took Pauli's proposal and developed it into a full quantitative theory of beta decay. In lieu of Chadwick's use of 'neutron' in his 1932 discovery of a distinct nuclear particle [10], Fermi used the Italian diminutive suffix to rechristen Pauli's particle, the neutrino.

Pauli supposedly observed that "[he had] done something very bad today by proposing a particle that cannot be detected; it is something no theorist should ever do" [11]. Thankfully this statement was overly pessimistic, although Pauli was right in so far as his prediction that neutrino observation would be incredibly challenging. The developments

in neutrino physics over the last century have been in spite of these challenges and have been all the more interesting as a result.

It would be over two decades later that another communication, this time a telegram to Pauli, heralded the first detection of neutrinos by Frederick Reines and Clyde Cowan in 1956 [12]. Their setup consisted of two cadmium-doped water target tanks sandwiched by three tanks containing liquid scintillator, making use of an intense neutrino flux from a fission reactor at the Savannah River Plant, South Carolina. Electron antineutrinos interact with nuclear protons via inverse beta decay to produce a positron and a neutron:

$$\bar{\nu}_e \; p \; \longrightarrow \; e^+ \; n. \tag{2.1}$$

A dual signature is observed resulting from the prompt annihilation of the positron and a delayed de-excitation of a cadmium nucleus following capture of the neutron. Their observation of what we now know to be the electron antineutrino observed a signal rate 20 times that of the reactor backgrounds. The experiment was repeated with heavy water (lower proton density), greater and lesser cadmium concentrations and with shielding applied (to reduce background neutron and gamma rates), all demonstrating consistency with the expected changes [13]. As they put it in their telegram "we have definitely detected neutrinos".

2.1.2 Not Just One Neutrino

Whilst the hunt for the neutrino had been going on, Neddermeyer and Anderson had identified the existence of a particle with a mass between that of electrons and protons [14, 15] in their studies of cosmic ray particles. This particle, the muon, had also been shown to behave in a manner different to other moderate energy particles [16, 17, 18] by Conversi and Piccioni in the mid-1940s. It would later be understood that this difference stemmed from the fact that the other particles were mesons (mainly pions), quark based composite particles with interactions mediated by the strong nuclear force, whilst the muon was in fact a new lepton, a 'heavy electron'. It was suggested by Pontecorvo that the existence of multiple charged leptons might imply that their associated neutrinos could also be distinct

particles [19]. Early results from Ray Davis' chlorine inverse beta decay experiments at Brookhaven had already implied that neutrinos and antineutrinos underwent non-identical interactions [20, 21].

By the 1960s, physicists were harnessing the properties of the newly discovered pions and muons to stray for the first time into accelerator-based neutrino physics. The Alternating Gradient Synchrotron at Brookhaven was used to collide 15 GeV protons with a beryllium target. The resulting particles included a significant flux of charged pions which would decay in flight via

$$\pi^{\pm} \longrightarrow \mu^{\pm} \overset{(-)}{\nu} \tag{2.2}$$

to produce a beam of neutrinos. The remaining charged elements of the beam were removed using iron shielding. This is the principle still used to this day to produce accelerator neutrino beams. A team led by Leon Lederman, Jack Steinberger and Melvin Schwartz observed neutrino interactions from this beam in a spark chamber. They showed that these neutrino interactions predominantly produced muons not electrons and that, combined with previous observations of the neutrino interaction rates, indicated that the neutrinos produced in pion decay, and associated with the muon were different from the electron associated neutrinos produced in beta decay [22]. This discovery, the existence of the muon neutrino, would go on to win the 1988 Nobel Prize in Physics for the three aforementioned authors.

2.1.3 Good Things (Neutrinos) Come in Threes

Evidence built in the mid 1970s for a third charged lepton, with behaviour similar to that of the electron and muon but a mass in the region of 1.7 GeV [23, 24, 25]. The discovery of the tau lepton by Martin Perl and others at SLAC immediately suggested that it may be the first of a full third generation of quarks and leptons including a corresponding tau neutrino to go with the muon and electron neutrinos [26].

Despite the confidence of these predictions it took until 2000 for a direct observation of ν_τ, primarily due to the large mass and short lifetime of its charged partner. The

detection of ν_e and ν_μ had utilised the production of $e^\pm$ and $\mu^\pm$ respectively via charged-current interactions. The neutrino energy required to produce $\tau^\pm$ via a charged-current interaction was prohibitively high for most accelerator neutrino beam facilities. The DONUT collaboration at Fermilab were the first to pass this hurdle. They observed four candidate tau lepton production events in a neutrino beam produced using 800 GeV protons from the Tevatron accelerator [27]. The distinctive signature of a single charged lepton produced at the vertex and decaying within the first couple of millimetres into a charged daughter lepton left them with an expected background of just 0.34±0.05 events. An analysis of their full dataset in 2008 yielded a total of 9 CC ν_τ events [28].

2.1.4 Is That It?

During the long wait for the observation of tau neutrinos, the ALEPH experiment at CERN's Large Electron-Positron Collider (LEP) had indicated that there would be no "fourth light neutrino flavour". They studied the decay of the Z-boson, the mediator of the weak neutral-current which had been discovered in 1983 [29, 30]. The Z-boson couples to all fermions and thus the decay width of the Z can be expressed as a sum over all kinematically accessible fermionic pairs:

$$\Gamma_Z = \sum_q^{u,d,c,s,b} \Gamma_{Z\to q\bar{q}} + \sum_l^{e,\mu,\tau} \Gamma_{Z\to l\bar{l}} + \sum_l^{e,\mu,\tau} \Gamma_{Z\to \nu_l\bar{\nu}_l}. \tag{2.3}$$

Note the fact that the top quark mass, $m_t > \frac{m_Z}{2}$ makes the decay $Z \to t\bar{t}$ inaccessible and thus the quark term is summed over the five lighter quark flavours. The decay width expressed in 2.3 can equivalently be written as

$$\Gamma_Z = \sum_q^{u,d,c,s,b} \Gamma_{Z\to q\bar{q}} + \sum_l^{e,\mu,\tau} \Gamma_{Z\to l\bar{l}} + N_\nu \Gamma_{Z\to \nu\bar{\nu}}, \tag{2.4}$$

where N_ν is the number of light active neutrino generations, by which we mean, the number of weakly-interacting neutrino flavours with a mass light enough to be accessible in Z-boson decays. By measuring the Z-boson width and using Standard Model predictions

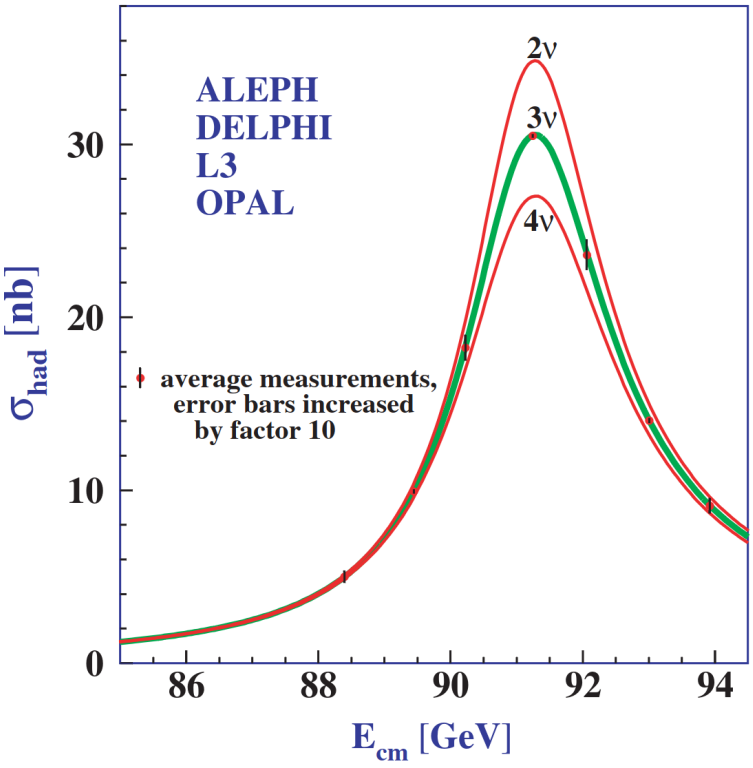


Figure 2.1: A combined measurement of hadronic production in the region of the Z-boson mass resonance, relative to the Standard Model predictions for 2, 3 and 4 light active neutrino flavours respectively. Figure from [33].

of the visible widths, ALEPH measured the value of N_ν to be 3.27±0.30 with just their first three weeks of data, ruling out the existence of a fourth light active neutrino at the 98% confidence level [31]. A later combined analysis of the entire dataset from all four LEP experiments (ALEPH, DELPHI, L3 and OPAL) resulted in a much more stringent constraint (see figure 2.1). The most modern value is given by the particle data group (PDG) to be 2.996±0.007 [32].

Cosmological measurements of the abundances of light elements (D, He and Li) can also be used to constrain the number of light neutrino flavours, although with less precision than the LEP result [34, 35].

2.2 Oscillations

So we have 3 light neutrinos matching the 3 generations of quarks and charged leptons, that all nicely fit together to form the Standard Model. With the discovery of the Higgs boson by the ATLAS [36] and CMS [37] collaborations in 2012, the final predicted piece was in place. However, by this point, neutrinos had already stopped playing nicely and given experimental particle physicists plenty to think about.

2.2.1 Solar Neutrino Problem

Given the original neutrino predictions emanated from anomalies in beta decay results, the Sun was an obvious place to search for a significant flux of neutrinos [38]. Detailed models of solar behaviour, developed principally by John Bahcall, indicated the decay of boron-8 should be the dominant source of solar neutrinos and that a larger version of the radiochemical experiments performed by Ray Davis at Brookhaven should be sensitive to this flux [39].

Davis established an experiment in the Homestake mine in Lead, South Dakota using 390,000 litres of tetrachloroethylene. The experiment observed electron neutrino capture on the 520 tons of chlorine-37 contained in the tetrachloroethylene resulting in the creation of a radioactive argon isotope with a half-life of 35 days.

$$\nu_e \, {}^{37}\mathrm{Cl} \longrightarrow \, {}^{37}\mathrm{Ar}^{+} \, e^{-}. \tag{2.5}$$

The argon was removed by purging the liquid with helium gas, and then its decay was observed using a small proportional counter. The initial results in 1968 showed that the rate of solar neutrinos was significantly lower than that predicted by Bahcall's modelling [40, 41]. Work continued over the course of the next decade with questions being put to both the theoretical predictions and the experimental results, but both held up under scrutiny. As more data was taken with the Homestake setup it became clear that the observed solar electron neutrino flux was consistently around a third of that predicted by the models [42],

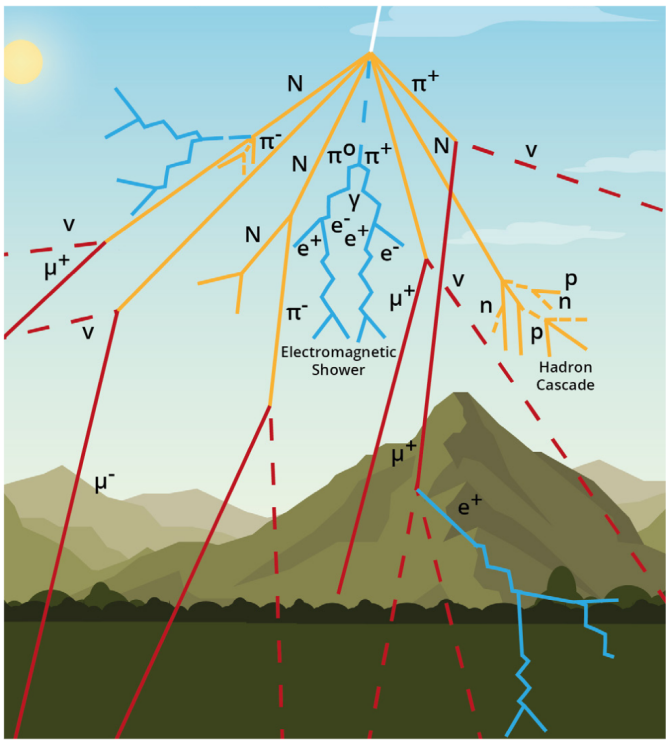


Figure 2.2: An illustration of the evolution of a cosmic ray shower induced by interaction in the earth's atmosphere. Figure from [46].

the tension was real and very much remained.

Further experiments including SAGE [43], GALLEX [44] and Kamiokande [45] further reinforced the Homestake results using both similar radiochemical approaches and water Cherenkov detection.

2.2.2 Atmospheric Neutrino Problem

Another large natural source of neutrinos originates from the top of our atmosphere. The earth is continually bombarded by a large flux of cosmic ray particles which undergo high-energy interactions with the atmospheric medium, creating showers of a variety of particles as indicated in figure 2.2. The hadronic cascades that result from these interactions

primarily finish with the following decays:

$$\begin{aligned} \pi^{\pm} \longrightarrow \overset{(-)}{\nu}_{\mu} \ \mu^{\pm} \\ \hookrightarrow e^{\pm} \ \overset{(-)}{\nu}_{\mu} \ \overset{(-)}{\nu}_{e} . \end{aligned} \tag{2.6}$$

As a result, the flux of atmospheric neutrinos was expected to consist of about double the number of muon neutrinos and antineutrinos relative to electron neutrinos and antineutrinos [47]. This neutrino flux was expected to be a significant source of background to the first generation of large proton decay experiments in the 1980s. Both the IMB and Kamiokande water Cherenkov experiments made measurements of the rate of atmospheric neutrinos in order to constrain this background. Both experiments observed that the ratio between muon-type and electron-type atmospheric events was lower than expected [48, 49]. These results were strengthened over the coming decade with further measurements at IMB-3 [50] and Kamiokande [51] as well as experiments such as MACRO [52] and Soudan-2 [53] using alternative detector technologies. What all of these results demonstrated was that the ratio was in fact significantly less than 2, whilst the deficit in muon-type events was correlated with the zenith angle, i.e. the direction of arrival of the neutrino. Downward going neutrinos, that had come from the atmosphere directly above the experiments, showed rates consistent with the theoretical prediction. However, upward going neutrinos, that had passed all the way through the earth having originated in the atmosphere on the other side of the planet, were the source of the significant deficit.

As with the tension in the observations of solar neutrinos, this difference was thoroughly examined with no issues in the theoretical prediction or the experimental observation being found.

2.2.3 SNO and Super-Kamiokande

The Sudbury Neutrino Observatory (SNO) and Super-Kamiokande experiments both published results around the turn of the millennium which conclusively demonstrated that both the solar neutrino problem [6] and the atmospheric neutrino problem [7] could be explained via the conversion between different neutrino flavour states. The two would later win the 2015 Nobel Prize for this critical advancement.

The SNO experiment consisted of a 1000 tonnes of heavy water surrounded by 9,600 photomultiplier tubes to detect the Cherenkov light emitted by the passage of high energy charged particles. Critically, the use of heavy water (D_2O) allowed SNO to be sensitive to three different neutrino interaction processes [54]:

$$\nu_e \; d \; \rightarrow \; p \; p \; e^- \qquad \text{(CC)} \tag{2.7}$$

$$\nu_x \; d \; \rightarrow \; p \; n \; \nu_x \qquad \text{(NC)} \tag{2.8}$$

$$\nu_x \; e^- \; \rightarrow \; \nu_x \; e^- \qquad \text{(ES).} \tag{2.9}$$

The charged-current (CC) channel is sensitive only to electron flavour neutrinos, while the neutral-current (NC) channel is equally sensitive to all 3 flavours. On top of this, the elastic neutrino electron scattering (ES) channel is also sensitive to all 3 flavours but with a higher sensitivity for electron neutrinos due to contributions from both charged and neutral-current processes.

By combining measurements of all three channels (see figure 2.3), SNO were able to show that the electron neutrino flux was consistent with that measured by the Homestake mine experiment, whilst simultaneously showing that the total neutrino flux was consistent with the predictions made by the standard solar model. This combination strongly suggested a process of flavour change was responsible for the electron neutrino deficit referred to as the solar neutrino problem. Neutrinos produced by beta decay in the Sun would be of electron flavour but during their propagation to Earth some proportion of them had converted to muon and tau flavour neutrinos and thus were only visible in the neutral-current and elastic scattering channels.

Super-Kamiokande's result focused on observations of atmospheric neutrinos, reiterating the existence of a deficit of muon neutrinos with a strong dependence on zenith angle. This can be seen in figure 2.4. Their results also showed that a similar deficit was not present in electron neutrino events. They were therefore able to demonstrate that such results were consistent with a length dependent conversion of muon neutrinos to tau neutrinos during their propagation from production in the upper atmosphere to detection on the Earth's

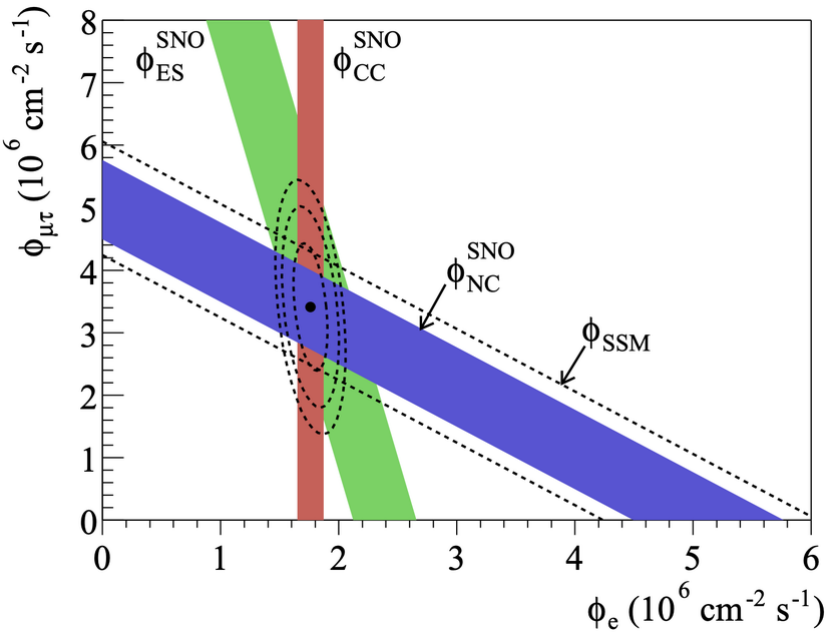


Figure 2.3: The electron neutrino (ϕ_e) and combined muon and tau neutrino fluxes ($\phi_{\mu\tau}$) measured by the SNO collaboration using the charged-current (red), neutral-current (blue) and elastic scattering (green) channels respectively. The charged-current band provides a restriction purely on the electron neutrino flux, leaving the entire $\phi_{\mu\tau}$ phase space unconstrained. However, the neutral-current restricts the sum of the three fluxes and so creates a constraint band of constant $\phi_e + \phi_{\mu\tau}$. Finally, the elastic scattering channel also places a restriction on the combination of the two fluxes but this time with a scaling factor on the electron component to account for the extra mechanism. This creates another sum constraint but with a different slope. The combined fit region as a result of these three constraints is also indicated. Figure from [6].

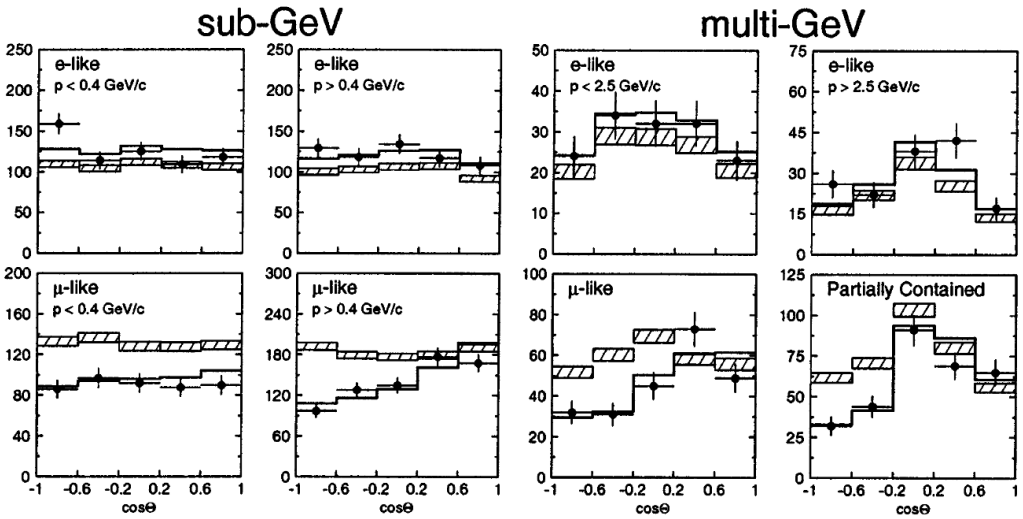


Figure 2.4: The Super-Kamiokande observation of both e-like and μ-like events as a function of their zenith angle. The hatched region indicates the predicted flux in the no oscillation case, whilst the solid line represents the predicted flux assuming $\nu_\mu \rightarrow \nu_\tau$ oscillations with best-fit values of $\sin^2 2\theta = 1$ and $\Delta m^2 = 2.2 \times 10^{-3}\,\mathrm{eV}^2$. Figure from [7].

surface.

Although not included in the Nobel Prize, results from another Japanese experiment, KamLAND, in 2002 confirmed the SNO results via the observation of a deficit in the electron antineutrino flux from nuclear reactors [55]. Finally, the K2K experiment were able to confirm the Super-Kamiokande result of muon neutrino disappearance, becoming the first experiment to observe oscillations in an accelerator beam of neutrinos [56, 57]. These results, using different sources, energies and baselines to SNO or Super-Kamiokande, helped add conclusive weight to the evidence piling up behind the neutrino oscillations picture.

2.2.4 Oscillation Framework

As it had become clear during the 1950s and 1960s that there were different flavours of neutrino and that neutrinos and antineutrinos may be distinct particles, Bruno Pontecorvo had developed a theory of transitions between these states analogous to the observed oscillations in neutral kaons [58, 59, 60]. Pontecorvo pursued this idea over multiple decades, describing a robust framework in which this mixing could occur.

It was this theory that provided the framework to explain the results of SNO and Super-Kamiokande, and from this basis a series of experiments have measured a variety of aspects of this neutrino oscillation phenomenon. These results have populated Pontecorvo's original theory with the required experimentally measured parameters and tested the implications of this on its predictions.

The model developed by Pontecorvo, Maki, Nakagawa and Sakata is predicated on the existence of neutrinos existing in two sets of eigenstates: flavour and mass. These states are related via the superpositions:

$$|\nu_\alpha\rangle = \sum_i U_{i\alpha} |\nu_i\rangle \qquad |\nu_i\rangle = \sum_\alpha U^*_{i\alpha} |\nu_\alpha\rangle \tag{2.10}$$

where $|\nu_\alpha\rangle$ are the flavour eigenstates (e, μ, τ), $|\nu_i\rangle$ are the mass eigenstates (1, 2, 3) and $U_{i\alpha}$ are the elements of the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix which describes the mixing between the two bases [61, 32]. The process of oscillations arises from the way in which the two bases appear in different contexts. Neutrinos only interact via the weak force, which couples exclusively to the flavour eigenstates. Thus the flavour of the neutrino can be directly defined by the electron, muon or tau emitted by its charged-current interaction. However, neutrino propagation occurs as a function of the mass eigenstate, and in the case of a non-diagonal PMNS matrix, these are not the same thing.

Particle propagation, in this case a neutrino, can be described as a plane wave solution to the Schrödinger equation:

$$|\nu_i(t)\rangle = e^{-iE_i t} |\nu_i\rangle \tag{2.11}$$

and thus the flavour eigenstates as:

$$|\nu_\alpha(t)\rangle = \sum_{i=1}^{3} U_{i\alpha}\ e^{-iE_i t}\,|\nu_i\rangle\,. \tag{2.12}$$

The chance of observing a neutrino in flavour state β after time t having initiated as flavour α is then given as:

$$\begin{aligned} P_{\alpha\to\beta}(t) =&\ |\langle\nu_\beta|\nu_\alpha(t)\rangle|^2 \\ =&\ \left|\langle\nu_\beta|\sum_{i=1}^{3} U_{i\alpha}\ e^{-iE_i t}\,|\nu_i\rangle\right|^2 \\ =&\ \left|\langle\nu_\beta|\sum_{i=1}^{3}\left(U_{i\alpha}\ e^{-iE_i t}\sum_{\lambda}^{e,\mu,\tau} U^*_{i\lambda}\,|\nu_\lambda\rangle\right)\right|^2 \\ =&\ \left|\sum_{i=1}^{3}\left(U_{i\alpha}\ e^{-iE_i t}\sum_{\lambda}^{e,\mu,\tau} U^*_{i\lambda}\,\langle\nu_\beta|\nu_\lambda\rangle\right)\right|^2. \end{aligned} \tag{2.13}$$

Given that $|\nu_\beta\rangle$ and $|\nu_\lambda\rangle$ are exact flavour states then $\langle\nu_\beta|\nu_\lambda\rangle = \delta_{\beta\lambda}$ and as such:

$$\begin{aligned} P_{\alpha\to\beta}(t) =&\ \left|\sum_{i=1}^{3}\left(U_{i\alpha}\ e^{-iE_i t}U^*_{i\beta}\right)\right|^2 \\ =&\ \sum_{i=1}^{3}\sum_{j=1}^{3} U_{i\alpha}U^*_{i\beta}U_{j\alpha}U^*_{j\beta}\ e^{-i(E_i-E_j)t}. \end{aligned} \tag{2.14}$$

We can then use the fact that the neutrino is ultra-relativistic to make a couple of substitutions. Firstly, as we are in the limit $p \gg m$, taking the leading order terms from a Taylor expansion yields:

$$\begin{aligned} E =&\ \sqrt{p^2+m^2} \\ =&\ p\sqrt{1+\left(\frac{m}{p}\right)^2} \\ \simeq&\ p+\frac{m^2}{2p}. \end{aligned} \tag{2.15}$$

The energy difference from equation 2.14 can therefore be written:

$$\begin{aligned} E_i - E_j =&\ \frac{m_i^2}{2p} - \frac{m_j^2}{2p} \\ =&\ \frac{m_i^2 - m_j^2}{2E} \end{aligned} \tag{2.16}$$

where E_i is the energy of the i^{th} mass eigenstate m_i and E is the energy of the produced neutrino. Secondly, $v = c$ so in natural units $t = L$ where L is the distance travelled by the neutrino. We can then express the transition probability as

$$\begin{aligned} P_{\alpha\to\beta}(t) &= \sum_{i=1}^{3}\sum_{j=1}^{3} U_{i\alpha}U^*_{i\beta}U_{j\alpha}U^*_{j\beta}\; e^{-\frac{i(m_i^2-m_j^2)L}{2E}} \\ &= \sum_{i=1}^{3}\sum_{j=1}^{3} U_{i\alpha}U^*_{i\beta}U_{j\alpha}U^*_{j\beta}\; e^{-\frac{i\Delta m_{ij}^2 L}{2E}} \\ &= \sum_{i=1}^{3}\sum_{j=1}^{3} U_{i\alpha}U^*_{i\beta}U_{j\alpha}U^*_{j\beta}\; \left[1 - 2\sin^2\frac{\Delta m_{ij}^2 L}{4E} + i\sin\frac{\Delta m_{ij}^2 L}{2E}\right]. \end{aligned} \tag{2.17}$$

Which can eventually be written as

$$\begin{aligned} P_{\alpha\to\beta}(t) = \delta_{\alpha\beta} &\\ &- 4\sum_{i=1}^{3}\sum_{j=1}^{3} \mathrm{Re}\left(U_{i\alpha}U^*_{i\beta}U_{j\alpha}U^*_{j\beta}\right)\sin^2\frac{\Delta m_{ij}^2 L}{4E} \\ &+ 2\sum_{i=1}^{3}\sum_{j=1}^{3} \mathrm{Im}\left(U_{i\alpha}U^*_{i\beta}U_{j\alpha}U^*_{j\beta}\right)\sin\frac{\Delta m_{ij}^2 L}{2E}. \end{aligned} \tag{2.18}$$

This representation is very useful in understanding the key aspects of neutrino oscillations. The first term encodes the no-oscillations scenario, and can easily be recovered if we set $\Delta m_{ij}^2 = 0$. Hence, any experimental verification of oscillations is, of itself, evidence for at least one non-zero neutrino mass. The other terms represent features of oscillations. Firstly, we can see that the mass squared difference, Δm_{ij}^2 encodes the frequency of these oscillations, a larger value creating quicker oscillations and a smaller value creating slower oscillations. The scale of the oscillations are set by the coefficients $4\sum_{i=1}^{3}\sum_{j=1}^{3}\mathrm{Re}\left(U_{i\alpha}U^*_{i\beta}U_{j\alpha}U^*_{j\beta}\right)$ and $2\sum_{i=1}^{3}\sum_{j=1}^{3}\mathrm{Im}\left(U_{i\alpha}U^*_{i\beta}U_{j\alpha}U^*_{j\beta}\right)$ which are made up of elements of the mixing matrix between the two eigen bases. The 'experimental variation' can therefore be injected via the combination of L and E which control which point of the oscillation probability wave is being probed.

The only mathematical difference in the representation for antineutrinos would be the exchange of $U_{ab} \leftrightarrow U^*_{ab}$. Clearly, this only impacts the final term, and therefore any differences in the behaviour between neutrinos and antineutrinos - known as CP-violation

- is encoded in this term. It is also worth noting that the term $\frac{\Delta m^2_{ij} L}{4E}$ can be written as $1.27\ \frac{\Delta m^2_{ij}[\mathrm{eV}^2] L[\mathrm{km}]}{E[\mathrm{GeV}]}$ to accept friendlier units.

Neutrino oscillation experiments measure the appearance or disappearance of a certain flavour of neutrino. This is done by using a source of neutrinos with a relatively well known initial flavour composition and measuring the change at a certain distance (e.g. solar or atmospheric neutrino oscillation measurements) or by measuring the flavour composition at two different points with near and far detectors (e.g. most accelerator neutrino experiments). With enough statistics, the probability of oscillation can be determined by the group behaviour and, using equation 2.18, the parameters that govern neutrino oscillations can then be extracted.

If we assume that neutrinos are Dirac, i.e. their antineutrinos are distinct particles, then the most common parametrisation of the three-flavour PMNS matrix is:

$$U = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \cdot \begin{pmatrix} c_{13} & 0 & s_{13}\, e^{-i\delta_{CP}} \\ 0 & 1 & 0 \\ -s_{13}\, e^{i\delta_{CP}} & 0 & c_{13} \end{pmatrix} \cdot \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{2.19}$$

where $c_{ij} = \cos\theta_{ij}$ and $s_{ij} = \sin\theta_{ij}$. This representation encodes the mixing in terms of 3 angles $(\theta_{12}, \theta_{13}, \theta_{23})$ and one CP-violating phase (δ_{CP}). As the name suggests $\delta_{CP} = 0\,(\text{or}\,\pi^{\pm})$ means no CP-violation, identical behaviour between neutrinos and antineutrinos, and any other value indicates some level of difference.

2.2.5 Matter Effects

The formalism adopted in the previous section assumes that the neutrinos are propagating in a vacuum. This is often not the case, and the interactions of neutrinos with any matter they pass through adds a further term to the propagation Hamiltonian. This term can modify the expected amount of mixing, due to the asymmetry between the interactions of electron neutrinos and those of muon or tau neutrinos [61]. Electron neutrinos have an extra contribution due to the existence of a charged-current electron scattering process, alongside the neutral-current process available to all flavours (see figure 2.5).

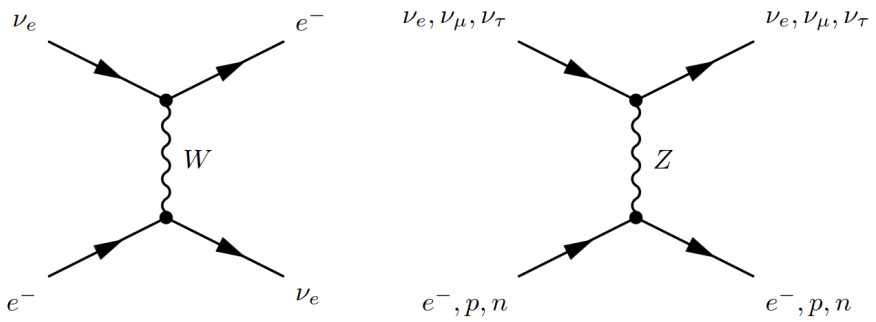


Figure 2.5: Feynman diagrams indicating the coherent forward elastic scattering that can occur between neutrinos and the fermions that constitute matter. Figure from [61].

Wolfenstein showed that this asymmetry creates a modified oscillation probability that depends on the electron density of the matter that has been traversed over the course of the experimental distance L [62]. Mikheev and Smirnov later demonstrated that this effect could become resonant in regions where the density slowly decreases, this resonance is known as the MSW effect [63].

2.2.6 Current Oscillation Parameters

A substantial body of work from a large number of experiments over the last two decades has resulted in most of the oscillation parameters (PMNS parameters and mass splittings) being known to within a few percent [64]. The combination of θ_{12} and Δm^2_{21} are often referred to as the solar parameters due to the fact they present in the electron disappearance signal measured in the solar neutrino flux. They have both been constrained well by a combination of measurements from the SNO [65], Super-Kamiokande [66] and KamLAND [67] experiments. θ_{13}, despite not forming the key contribution to either the solar or atmospheric anomalies, is now the best known of the three mixing angles. The first indications that it was non-zero came from the T2K [68] and MINOS [69] experiments in 2011. The most precise measurements now result from the Daya Bay [70] and RENO [71] experiments, both of which measure the disappearance of electron antineutrinos from a reactor source using an inverse beta decay method. The discovery of substantially non-

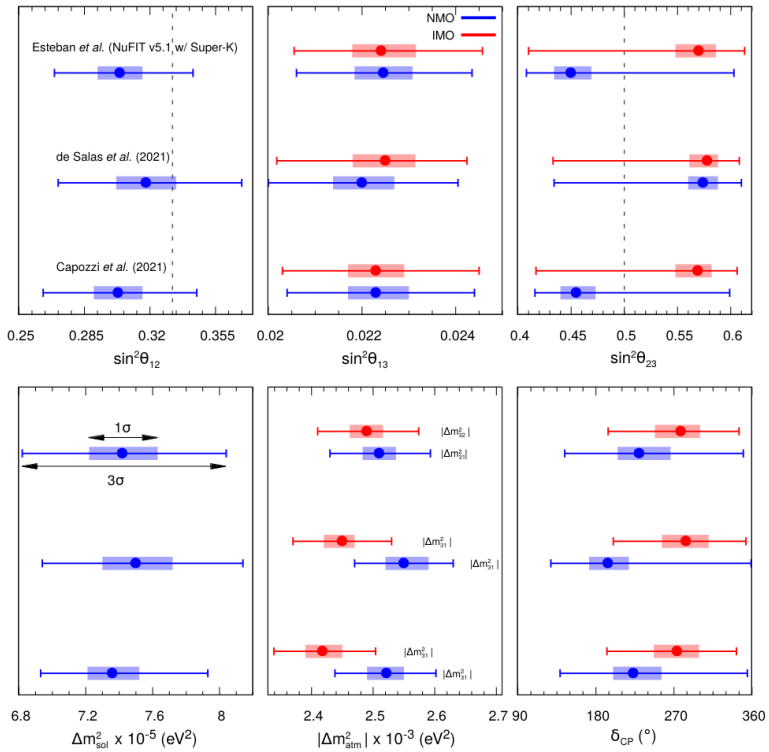


Figure 2.6: The results of three recent global fits of the 3ν oscillation parameters. Figure from [75] using fits [76, 77, 78].

zero θ_{13} allows the possibility to probe for CP-violation in the neutrino sector. Due to the coupling between θ_{13} and δ_{CP} in equation 2.19 a zero or very small θ_{13} would make δ_{CP} impossible to measure. The 'atmospheric' mixing angle, θ_{23} has been widely measured by atmospheric and accelerator experiments such as Super-Kamiokande [72], T2K [73] and NOνA [74], although it is not yet clear in which 'octant' the value lies (see section 2.2.6.3). Finally, the mass splitting $|\Delta m^2_{32}|$ is also well measured due to the fact it can be accessed by reactor, atmospheric and accelerator experiments alike. The strongest constraints on this value come from Daya Bay [70], NOνA [74] and T2K [73].

Figure 2.6 shows a plot from [75] that nicely summarises the results of three recent global fits to oscillation data in the three neutrino paradigm [76, 77, 78]. The results are presented for both mass hierarchies (see section 2.2.6.1).

The next generation of neutrino oscillation experiments, primarily DUNE and T2HK -

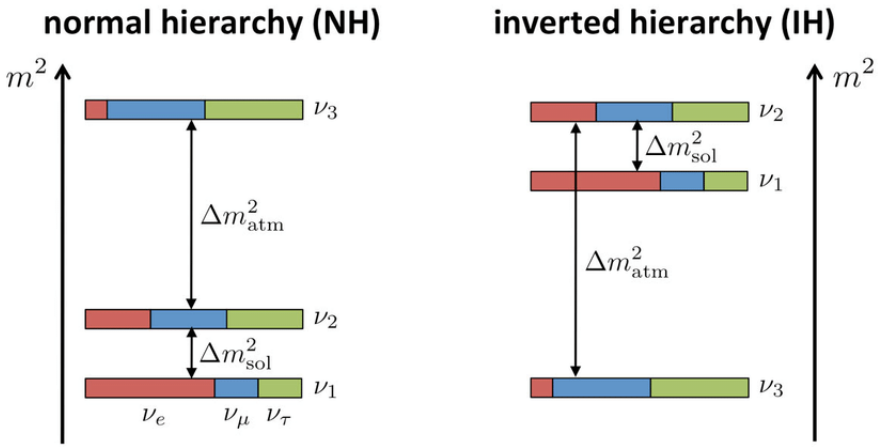


Figure 2.7: The two possible scenarios for the order of the neutrino mass states, normal and inverted ordering. Each bar also demonstrates the flavour composition of each mass state. Figure from [79].

the accelerator successors to NOνA and T2K - and JUNO - a moderate baseline reactor successor to Daya Bay - will make more precise measurements of the oscillation parameters. They will specifically target the three key areas in which current measurements have not been able to render well known values: the mass hierarchy, the value of δ_{CP} and the octant of θ_{23}.

2.2.6.1 Mass Hierarchy

Whilst initial oscillation results indicated conclusively that neutrinos did have non-zero masses, further details of these masses are difficult to determine. The oscillation effects are only sensitive to the squared mass difference, and not the individual masses themselves, whilst only the final term of equation 2.18 is sensitive to the sign of Δm^2, in the second term it is absorbed by the $\sin^2$.

Matter effects *are* dependent on the sign of the splitting, and therefore offer a handle with which oscillation experiments with sufficiently long baselines can determine this sign. This has been achieved for the solar mass splitting, Δm^2_{21}, thanks to the MSW effect in the

solar medium, but has yet to be determined for the atmospheric splitting Δm^2_{32}. This leaves an unresolved question known as the mass ordering problem, whereby it is currently unknown as to whether the mass of the third eigenstate is lighter or heavier than the first and second. The former scenario is described as the normal hierarchy and the latter as the inverted hierarchy and is illustrated in figure 2.7. Whilst all three global fits in figure 2.6 favour the normal ordering scenario, they do so only with limited significance.

2.2.6.2 δ_{CP}

Conservations of parity (P) and charge conjugation (C) were, in the early days of the Standard Model, expected to be absolute. However, experiments by Chien-Shiung Wu in the 1950s, showed that in fact both were violated in beta decay processes [80]. The idea that the combination of the two, CP, could be conserved was also shown to be incorrect in 1964 by James Cronin and Val Logsdon Fitch when they observed the rare decay of neutral kaons into a pair of charged pions [81]. CP-violation is also one of the Sakharov conditions for the creation of the universe's apparent matter-antimatter asymmetry [82].

Although CP-violation has been observed in the hadronic sector, it has not been observed in sufficient quantity to satisfy Sakharov's condition. Neutrino oscillations potentially offer a mechanism by which CP-violation could occur in the leptonic sector. Thus far, oscillation experiments have not been able to conclusively determine the value of δ_{CP} or rule out the CP-conserving scenarios of $\delta_{CP} = 0$ or $\pi^{\pm}$. The determination of the CP-violating phase is coupled to the mass ordering scenario, both of which are accessible only via the third term of equation 2.18.

In recent years the T2K experiment has published results that exclude CP-conservation in the neutrino sector at the 90% confidence level [83, 84]. These indications have not been confirmed in similar searches by the NOνA experiment [74] although neither results have the required precision for this tension to be significant.

2.2.6.3 θ_{23} Octant

The value of θ_{23}, the atmospheric mixing angle, is known to be close to the maximal value of $\frac{\pi}{4}$ but current experiments have been unable to lift the degeneracy caused by the fact that it appears in the muon neutrino survival probability ($P(\nu_\mu \longrightarrow \nu_\mu)$) as a $\sin^2 2\theta_{23}$ term. This produces a symmetrical degeneracy around the maximal and whether it is just above or just below is known as whether it sits in the higher or lower octant respectively. A value of exactly $\frac{\pi}{4}$ would indicate a symmetry between the μ and τ flavours which itself can be linked to maximal CP-violation [85, 86]. This is the only point on which there is significant disagreement between the fitting groups in figure 2.6, two favouring lower octant in the normal mass ordering scenario whilst the third favours the higher octant in both mass ordering scenarios. Measurements of electron neutrino appearance in future long baseline oscillation experiments will be able to lift this degeneracy due to the dependence of $\mathrm{P}(\nu_\mu \rightarrow \nu_e)$ on $\sin^2 \theta_{23}$.

2.2.7 Look a little ~~closer~~ nearer

It is clear from the previous discussion that increasingly precise measurements of the oscillation parameters in the 3-neutrino paradigm have developed a clear, consistent picture of neutrino oscillations as an extension (by virtue of neutrino mass) to the Standard Model. However, as is becoming the theme of neutrino physics, this is not the end of the story. A series of results across different types of experiments in the 1990s and 2000s pointed towards an inconsistency at low L/E values.

- Experiments studying electron antineutrino flux from nuclear reactors at baselines $\mathcal{O}(10-100\,\mathrm{m})$ saw a deficit relative to the prediction of nuclear reactor models. This is known as the reactor antineutrino anomaly [87].
- Gallium radiochemical experiments, such as GALLEX and SAGE, which were aimed at measuring the low-energy solar electron neutrino flux used radioactive sources at very short baselines for calibration purposes. The deficit they saw is known as the Gallium anomaly [88].

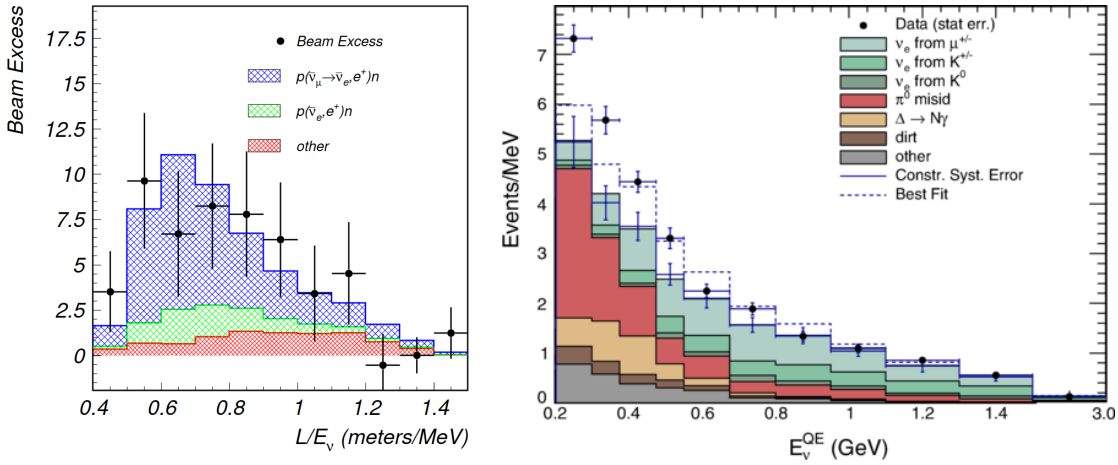


Figure 2.8: The excess of electron neutrino appearance reported by the LSND (left [89]) and MiniBooNE (right [90]) experiments. The LSND plot reports the number of data excess events, following the subtraction of the Monte Carlo - hence the negative value at higher L/E. The shapes of the different Monte Carlo contributions are shown to guide possible explanations of the excess. The MiniBooNE plot instead reports the unsubtracted data points which clearly demonstrate an excess over the Monte Carlo distribution in the lower energy region.

- Two accelerator based neutrino experiments, LSND [89] and MiniBooNE [90] using differing technologies saw the appearance of electron (anti)neutrinos over similar L/E in beams of muon (anti)neutrinos. Figure 2.8 shows the distributions of these excesses.

These results all indicated electron neutrino appearance or disappearance that could not be explained by 'standard' 3-neutrino mixing. The most popular proposed solution for these results was the existence of a fourth 'sterile' neutrino. First proposed by Bruno Pontecorvo right back in 1967 [60], these sterile neutrinos would couple only to the gravitational interaction yielding them impossible to directly detect via the weak interactions used to detect their active counterparts. This would also render them consistent with the measurements of the Z-boson width discussed in section 2.1.4.

The 3+1 model, in which a single sterile species accompanies the 3 traditional neutrino

flavours, yields an extra mass state ν_4, and thus mass splitting Δm^2_{41}, as well as an extension to the PMNS matrix. If the fourth mass state were large enough that $\Delta m^2_{41} \simeq \Delta m^2_{42} \simeq \Delta m^2_{43} \simeq 1\,\text{eV}^2$ then the resulting fast oscillations could explain the discrepancies appearing at such small baselines [91, 92].

In recent years the landscape has evolved further. Improved models of the reactor neutrino flux have all but removed the reactor anomaly [93], although some tension still remains around the origin of the 5 MeV bump observed across many reactor experiments [94]. Experiments such as DANSS [95], NEOS and RENO [96], PROSPECT [97] and STEREO [98] have all published results that disfavour sterile neutrinos and exclude significant areas of the phase space. However, these results are in tension with the Neutrino-4 reactor experiment which published results (shown in figure 2.9) indicating a 2.7σ preference for sterile neutrino oscillations with $\Delta m^2_{41} = 7.3 \pm 1.17\,\text{eV}^2$ and $\sin^2 2\theta = 0.36 \pm 0.12$ [99]. The methodology used to produce these results has met with some skepticism from the community but further results from upgraded versions of DANSS, Neutrino-4 and PROSPECT should test these claims with improved precision [100].

In the regime of the Gallium anomaly the BEST experiment has published results [101, 102] indicating a strengthening of the discrepancy with a setup dedicated to exploring the SAGE and GALLEX results. The BEST results favour a mass splitting of $\Delta m^2_{4x} > 1\,\text{eV}^2$ which is at significant tension with the apparent lack of reactor anomaly but would be consistent with the Neutrino-4 results.

Finally, the MicroBooNE LArTPC has performed searches for the low energy excess (LEE) observed by MiniBooNE. Neither the electron [103] nor photon [104] channels saw compatibility with the LEE. Their first multi-channel search for evidence of 3+1 oscillations found no evidence for the existence of a sterile neutrino, although this was limited by possible degeneracy between ν_e appearance and disappearance [105]. These results only comprise the first half of the MicroBooNE dataset, future searches will be performed with the full dataset and with techniques designed to break the aforementioned degeneracy.

Further studies have been performed by long baseline experiments such as MINOS(+) [106], NOνA [107, 108] and IceCube [109]. They can probe the existence of a sterile neutrino

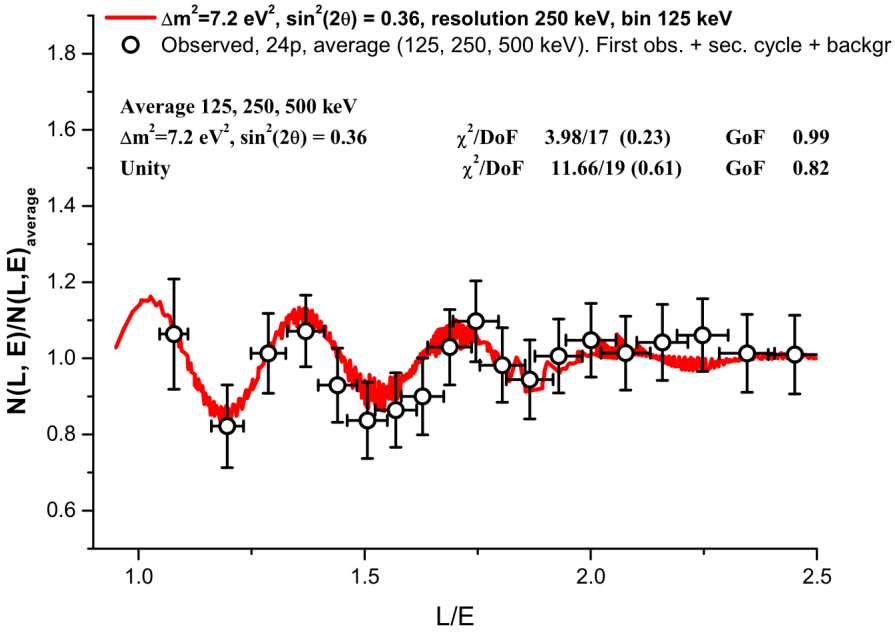


Figure 2.9: The results of the Neutrino-4 search for sterile neutrinos showing the ratio of events at a certain L/E to the average of those with the same energy over all available baselines. Plotted in red is their best fit result of $\Delta m^2_{41} = 7.3 \pm 1.17\,\mathrm{eV}^2$ with $\sin^2 2\theta = 0.36 \pm 0.12$. Figure from [99].

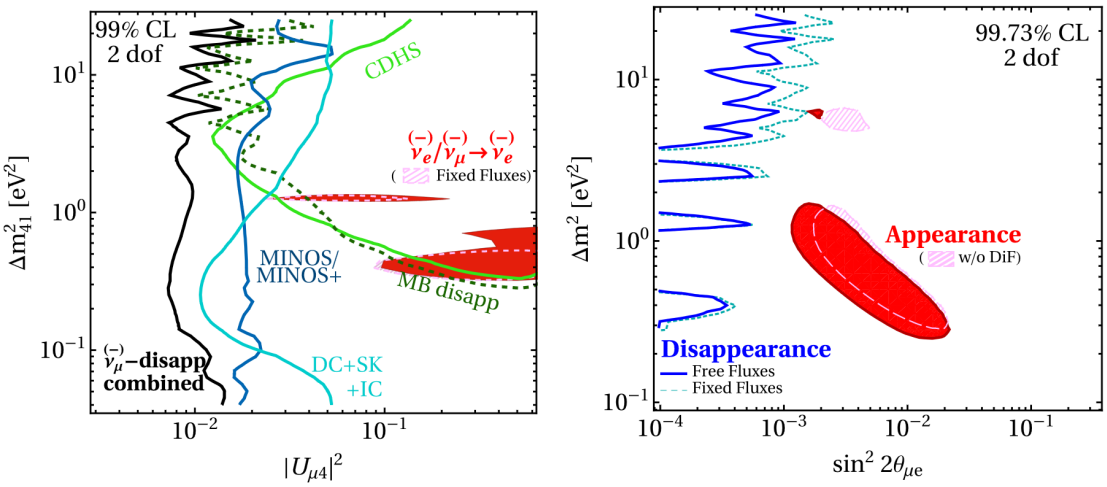


Figure 2.10: Global fits of short-baseline ν_μ disappearance and ν_e (dis)appearance in both $\Delta m^2_{41} - |U_{\mu 4}|^2$ and $\Delta m^2_{41} - \sin^2 2\theta_{\mu e}$ spaces. There is a clear tension between the allowed regions of appearance results - in red - and the exclusion curves set by disappearance results. Figure from [110].

via the muon disappearance signature which should complement the electron appearance and disappearance signatures observed in the above anomalies. None of these experiments found evidence for such a sterile induced signature.

These competing results have led to a somewhat muddled picture. Despite the reduction of the reactor anomaly, the Gallium anomaly and LSND and MiniBooNE anomalies still require explanation, and the MicroBooNE results further complicate that. The 3+1 scenario suffers from significant tension between the disappearance and appearance results as can already be seen in figure 2.10 which pre-dates both the BEST and MicroBooNE results. The most up-to-date picture of this phase space is shown in figure 2.11 in which all three recent results (BEST, MicroBooNE and Neutrino-4) are shown - alongside the original LSND appearance results.

Clearly, the 3+1 scenario faces challenges in explaining the current landscape of short-baseline neutrino physics, but the anomalies still stand and require a solution. Other extensions to the Standard Model have been postulated to explain some or all of these results, including the addition of further sterile neutrinos [111], neutrino coupling to new

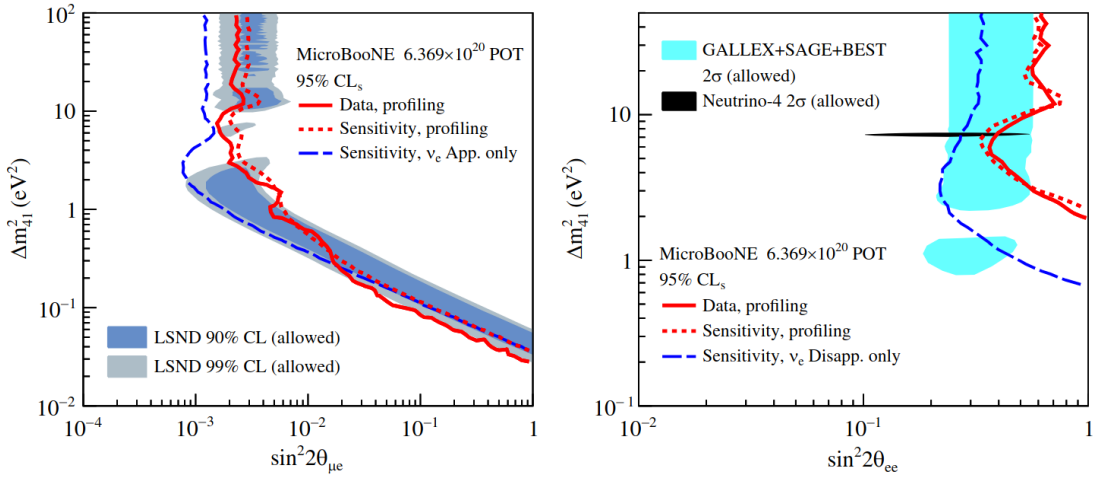


Figure 2.11: Global fits of short-baseline ν_e appearance and disappearance in $\Delta m^2_{41} - \sin^2 2\theta_{\mu e}$ and $\Delta m^2_{41} - \sin^2 2\theta_{ee}$ spaces respectively, following the first MicroBooNE search for sterile neutrinos. The appearance space shows MicroBooNE excluding a large portion of the allowed LSND region but leaving a substantial section around $\Delta m^2_{41} \sim 1\,\mathrm{eV}^2$. The disappearance space shows that the combined Gallium anomaly still covers a large allowed region and does overlap with the small Neutrino-4 region. Figure from [105].

boson states [112], non-standard neutrino interactions [113] or heavy sterile decay scenarios (heavy neutral leptons) [114]. It is in this context that the Short-Baseline Neutrino program at Fermilab has been designed and constructed, with the unique ability to simultaneously search for ν_μ disappearance, ν_e appearance and ν_e disappearance at multiple baselines in the same beam (see chapter 4).

Chapter 3

Neutrino Interactions

How do we study neutrinos? The previous chapter outlined a wealth of fascinating experiments that have been, or will be, carried out to determine the properties of this most elusive of particles. The common theme uniting them all is they study 'the result of neutrino interactions'. As neutral particles, we can only study them via the particles that they produce when they interact weakly with the material in our detectors. This makes understanding neutrino interaction cross sections; the rate at which they interact, the particles they produce and the kinematics of these particles, vital to the performance of such experiments, especially as we move into the 'precision' era of neutrino physics.

Modelling of neutrino interactions contributes to precision oscillation measurements via both the background subtraction and energy reconstruction. Even with functionally identical near and far detectors, the flux and detector acceptance differences mean total cancellation of interaction systematic uncertainties is impossible. At the energies of relevance to accelerator neutrinos, there are a number of neutrino interaction generators designed to model these interactions, most commonly GENIE [115], NuWro [116], NEUT [117] and GiBUU [118]. Due to the complex nature of these interactions and the limited data available, these generators disagree significantly in their predictions. Making improved measurements of neutrino interaction cross sections and comparing them to such generators is critical to reducing the systematic uncertainties of future experiments.

Neutrinos are both electrically neutral and colourless and as such they do not experience the effects of the electromagnetic or strong forces, they couple only to the $W^{\pm}$ and Z bosons of the weak interaction. The $W^{\pm}$ and Z bosons mediate the charged (CC) and neutral (NC) currents respectively. In a charged-current interaction a conversion occurs between the neutrino and its charged leptonic counterpart or vice versa, in a neutral-current interaction the neutrino remains as a neutrino. As described in the previous chapter, the coupling of each neutrino to its respective charged lepton is the manner by which neutrinos were first observed, and the way in which neutrino flavour can be tagged in oscillation experiments.

Before we dive into the details of the different neutrino interaction modes, it is worth defining the term cross section. Cross sections provide particle physicists with a measure of the probability of an interaction. The cross section is the effective area within which the interaction can take place. It is given by the following formula:

$$\sigma = \frac{N}{\Phi \; n_t}, \tag{3.1}$$

where N is the number of interactions, Φ is the time- and energy-integrated neutrino flux required to produce those interactions and n_t is the number of targets in the material exposed to the neutrino flux.

It is often useful to express or measure cross sections in terms of their dependence on different initial state, interaction or final state variables. This allows further subtleties of the cross section to be understood, especially when making comparisons between data and theoretical models. This is known as measuring the differential cross section.

3.1 Neutrino-Nuclear Interaction Modes

The wavelength of any wave affects its resolution. This simple fact is played out in everyday cameras, telescopes making observations of astronomical phenomena and the use of radio waves for human communications. As Louis de Broglie pointed out in 1924 [119], a direct result of wave-particle duality is a calculable 'wavelength' for any particle. This wavelength

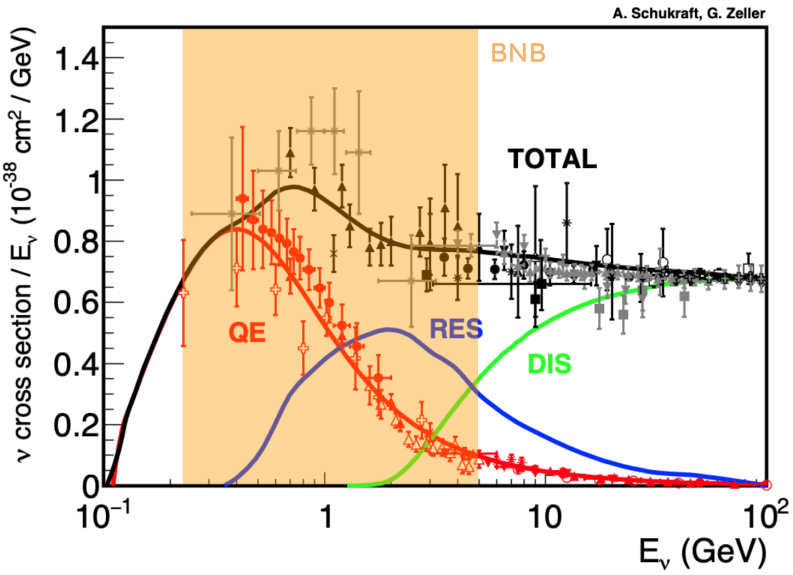


Figure 3.1: The total energy-normalised neutrino charged-current cross section per nucleon. The black and coloured lines show the total cross section and QE, Resonant and DIS contributions respectively. These predictions were created using the NUANCE generator and are compared to a series of measurements from a variety of neutrino experiments. The range of BNB neutrino energies are superimposed for reference. Adapted from [120, 121].

reduces as the particle's momenta increases, thereby allowing it to resolve smaller structure. This has a direct impact on neutrino-nuclear interactions where the neutrino's energy affects what level of nuclear structure it resolves; interacting with the nucleus as a whole, the individual nucleons or the nucleons' constituent quarks.

At the range of neutrino energies to which SBND is exposed in the Fermilab Booster Neutrino Beam (BNB) there are a number of different interaction modes that contribute significantly to the total event rate. The total neutrino charged-current cross section in the sub-GeV to 100 GeV regime is illustrated in figure 3.1. It is immediately clear from this plot that the energy regime of the BNB lies in a transition region where contributions from quasi-elastic (QE), resonant (RES) and deep inelastic scattering (DIS) processes are all significant and change appreciably in scale over the relevant energy range. The following sections will describe these different modes of interaction.

3.1.1 Quasi-Elastic and 2p2h

The dominant interaction mode for SBND, and other experiments that operate in the ∼1 GeV region, is that of charged-current quasi-elastic scattering. The neutrino scatters off a single nucleon converting that nucleon in the process:

$$\begin{aligned} \nu_l \; n &\longrightarrow l^- \; p \\ \bar{\nu}_l \; p &\longrightarrow l^+ \; n. \end{aligned} \tag{3.2}$$

The cross section for such a process can be described theoretically using the Llewellyn-Smith formalism [122]. This description describes the differential quasi-elastic cross section in terms of the vector, axial vector and pseudoscalar nucleon form factors. These form factors describe the distribution of charge within the nucleon. The pseudoscalar form factor can be neglected as its effect is heavily mass suppressed due to its cross section contribution being scaled by a factor of $\frac{1}{M}$ where M is the nucleon mass [123]. The vector form factor is accessible via electromagnetic electron scattering and can be utilised thanks to the conserved vector current (CVC) hypothesis. The axial vector form factor is therefore left to be determined via weak neutrino scattering. Traditionally, this form factor has been expressed in a dipole form:

$$F_A(Q^2) = \frac{F_A(0)}{\left(1 + \frac{Q^2}{M_A^2}\right)^2} \tag{3.3}$$

where Q^2 is the four-momentum transfer of the interaction and M_A is the 'axial mass' used to parameterise the scaling. $F_A(0)$, or g_A as it is often noted, can be determined from neutron β decay [124] leaving M_A as the final free parameter. A number of bubble chamber experiments in the 1970s and 1980s made measurements of the quasi-elastic charged-current interactions using hydrogen or deuterium. This equated to making the measurement on free or essentially free protons. The results from these experiments agreed well and gave a world average value for the axial mass of $M_A = 1.026 \pm 0.021$ GeV [125]. The neutrino mode quasi-elastic events give a particularly clean experimental signature, especially in imaging detectors, where both an outgoing lepton and proton should be identifiable from a shared vertex.

Most 21$^{\text{st}}$-century neutrino experiments use targets consisting of heavier target nuclei, due to their larger density and thus event rates. Predominantly this has consisted of

carbon (within organic scintillators), oxygen (within water) or argon. Unlike hydrogen or deuterium the component nucleons cannot still be treated as free independent particles. Generally, this is modelled using the impulse approximation, whereby the nucleons are treated as independent and then combined incoherently [126]. The treatment of nuclear effects in this combination is typically handled using a relativistic Fermi gas (RFG) approach [127] in which each nucleon is given some momentum characterised by a combination of the binding energy and Fermi momentum of the nucleus [124]. Other approaches encompass use of a 'local Fermi gas' (LFG) in which the Fermi momentum, instead of being constant, is dependent on the radial position of the nuclei within the nucleus [128]. Spectral functions and the random phase approximation can be used to add the effects of correlated motion between nucleons at short and long distances respectively [129]. Further suppression factors are added to account for the Pauli blocking effect, in which the final nucleon state cannot correspond to any previously occupied state due to the Pauli exclusion principle.

Results in the late 2000s, in particular those of the MiniBooNE experiment (see figure 3.2), demonstrated a significant disagreement between CCQE data on nuclear targets when compared to models tuned to the previous era of bubble chamber results. The MiniBooNE data showed a significant excess across their entire energy regime when compared to such model predictions, whilst contemporary results from LSND [130] and NOMAD [131] agreed well in lower and higher energy regimes respectively. One interpretation of such results is to tune the axial mass to match the newer nuclear data, yielding values around $M_A = 1.3\,\text{GeV}$, however, this approach introduces an equivalent tension with the NOMAD data. No combination of the more sophisticated extensions to the relativistic Fermi gas model could address the discrepancy [132].

An alternative proposal was presented by Martini et al. [134] in which they account for interactions involving multi-nucleon excitations, predominantly the 2-particle-2-hole (2p2h) state in which the neutrino interacts with a correlated pair of nucleons. This can result in multi-nucleon emission events which, as a result of Pauli suppression, undetectable neutrons or final state interactions (see section 3.2), can appear alongside the 'true' quasi-elastic events in experimental results. Modern experiments now usually report their results

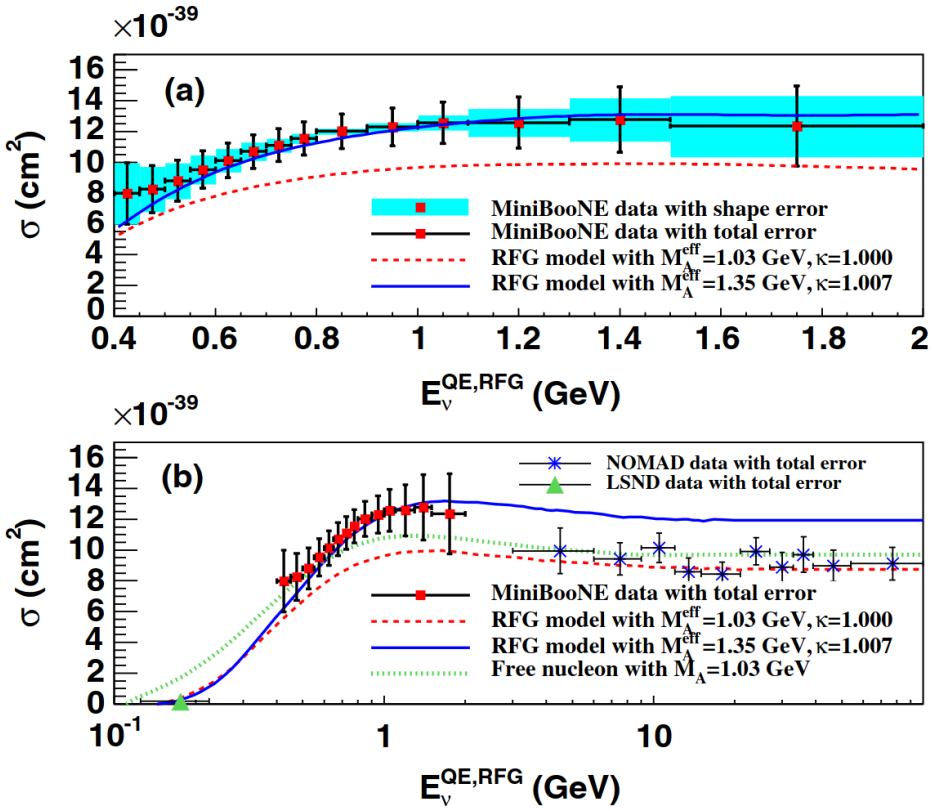


Figure 3.2: The MiniBooNE measurement of the CCQE cross section on ^{12}C as a function of neutrino energy alongside further contemporary data from LSND and NOMAD with a prediction provided by the NUANCE generator utilising a relativistic Fermi gas model. Figure and information from [133].

in terms of the detectable final state (e.g. CC $\nu_\mu\, 0\pi$) in order to more clearly account for the events they are measuring. Models including the 2p2h contributions alongside the sophisticated Fermi gas QE approaches were able to explain the MiniBooNE results, whilst preserving a value of $M_A \simeq 1\,\text{GeV}$ [135].

3.1.2 Pion Production - Resonant and Coherent

Moving beyond the elastic-like interaction types we find events in which particles other than the interacting lepton and nucleon are emitted. Due to their relatively light mass, the first particles for which this starts to occur are pi-mesons or pions. In figure 3.1 the process that quickly becomes dominant following the quasi-elastic peak is that of resonance events. In these interactions the nucleon is excited by the neutrino into a resonant state, most commonly $\Delta(1232)$. The nucleon resonant state then promptly decays back to a ground state nucleon, emitting a pion in the process. Example Feynman diagrams for charged and neutral-current resonant pion production events are shown in figure 3.3, whilst the full list of possible resonant pion production processes off free nucleons [120] are:

$$\begin{aligned}
\nu_l\, p &\longrightarrow l^-\, p\, \pi^+ & \bar{\nu}_l\, p &\longrightarrow l^+\, p\, \pi^- \\
\nu_l\, n &\longrightarrow l^-\, p\, \pi^0 & \bar{\nu}_l\, p &\longrightarrow l^+\, n\, \pi^0 \\
\nu_l\, n &\longrightarrow l^-\, n\, \pi^+ & \bar{\nu}_l\, n &\longrightarrow l^+\, n\, \pi^- \\
\nu_l\, p &\longrightarrow \nu_l\, p\, \pi^0 & \bar{\nu}_l\, p &\longrightarrow \bar{\nu}_l\, p\, \pi^0 \\
\nu_l\, p &\longrightarrow \nu_l\, n\, \pi^+ & \bar{\nu}_l\, p &\longrightarrow \bar{\nu}_l\, n\, \pi^+ \\
\nu_l\, n &\longrightarrow \nu_l\, n\, \pi^0 & \bar{\nu}_l\, n &\longrightarrow \bar{\nu}_l\, n\, \pi^0 \\
\nu_l\, n &\longrightarrow \nu_l\, p\, \pi^- & \bar{\nu}_l\, n &\longrightarrow \bar{\nu}_l\, p\, \pi^-
\end{aligned} \tag{3.4}$$

where the first three are charged-current interactions, and the last four neutral-current.

Although pion production is the most dominant mechanism in the resonance channel, the excited states, particularly at resonances beyond $\Delta(1232)$, can also decay to final states involving multiple pions, heavier mesons such as kaons, or even radiatively via the emission of a photon. Resonance production has typically been understood via the Rein-Sehgal [136], and subsequently Berger-Sehgal [137], models which account for a series of resonances up to

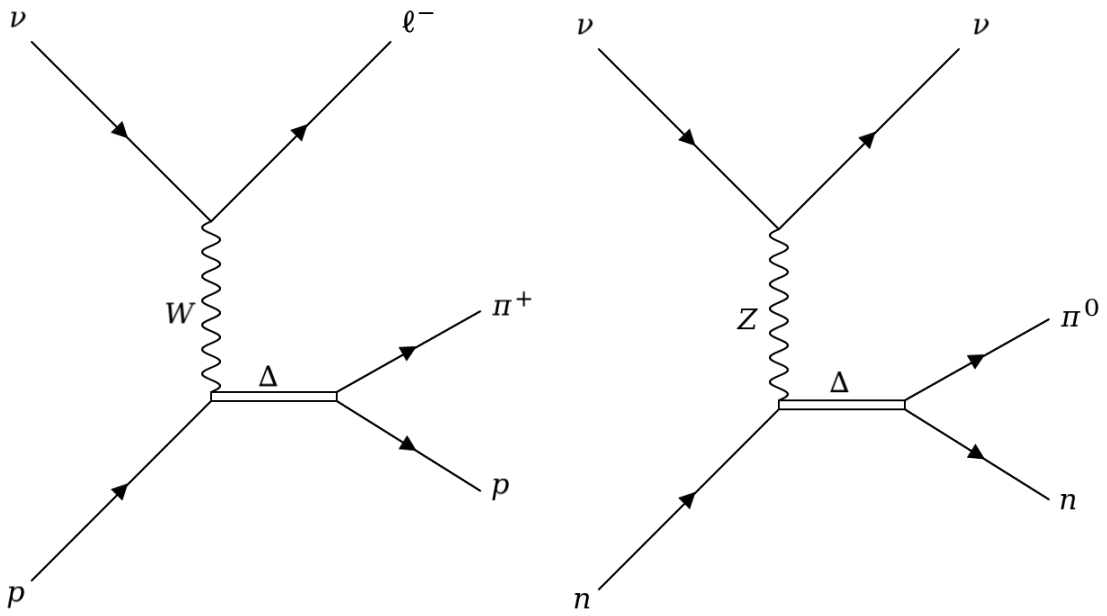


Figure 3.3: Example Feynman diagrams for resonant pion production in neutrino-nucleon interactions under charged (left) and neutral (right) currents.

about 2 GeV, using a harmonic oscillator approach. The latter model is purely an extension of the former in which an accountance is made for the mass of the lepton.

Pion production is further complicated by the existence of a coherent process in which the neutrino scatters off the entire nucleus producing a single pion, for example:

$$\begin{aligned} \nu_l\ A &\longrightarrow \nu_l\ A\ \pi^0 \qquad \bar{\nu}_l\ A \longrightarrow \bar{\nu}_l\ A\ \pi^0 \\ \nu_l\ A &\longrightarrow l^-\ A\ \pi^+ \qquad \bar{\nu}_l\ A \longrightarrow l^+\ A\ \pi^-. \end{aligned} \tag{3.5}$$

Due to the small momentum transfer associated with this interaction the pion is typically very forward going, aligned with the original neutrino. As with resonant interactions, the focus often lies on pion production, although other final states such as ρ-mesons or photons are also possible. The modelling of such processes was developed again by Rein and Sehgal [138] making use of the partially conserved axial current hypothesis (PCAC) [139] in which the differential cross section can be related to that of pion-nucleus elastic scattering at vanishing Q^2. A prescription involving the axial mass is then used to scale the cross section for non-zero values of Q^2. A microscopic approach is also possible, although only

valid in the region of the $\Delta(1232)$ resonance [120].

Initial measurements at present accelerator energies ($E_\nu \lesssim 2\,\text{GeV}$) demonstrated good agreement with the modelling of the neutral-current coherent interactions (producing a π^0), whilst searches for charged-current coherent interactions (producing a $\pi^\pm$) only yielded upper bounds, suggesting an overprediction [140, 141]. This was particularly evident in the SciBooNE experiment who measured the ratio of the two processes in the same beam. Their measurement of $\frac{\sigma_{CC\,Coh}}{\sigma_{NC\,Coh}} = 0.14^{+0.30}_{-0.28}$ [142] was much lower than predicted by either theoretical approach, and gave no evidence for non-zero charged-current coherent interactions. This tension has reduced as modern experiments with increased statistics have been able to make measurements of the charged-current rate in closer agreement with the model predictions [141]. Most recently, the T2K experiment improved their measurement of CC ν_μ induced coherent pion production alongside a first measurement of the CC $\bar{\nu}_\mu$ variety [143].

3.1.3 DIS

Towards the higher end of the neutrino energy spectrum of the booster neutrino beam, a further process becomes accessible, that of deep inelastic scattering (DIS). Unlike the previous processes in which the neutrino has interacted with the entire nucleus or individual nucleons, DIS takes place once the four-momentum transfer is large enough for the neutrino to resolve the constituent quarks that make up a nucleon.

DIS interactions are less theoretically uncertain than some of the previously assessed processes. This is primarily due to the fact that the interaction does not depend as heavily on complex nuclear dynamics and instead is parameterised by the nucleon structure functions that describe the parton model of hadronic states [120, 144]. DIS interactions are typically modelled via the Bodek-Yang model [145] with the evolution of the resulting hadronic state (hadronisation) via the Andreopoulos-Gallagher-Kehayias-Yang (AGKY) model [146]. The latter is necessary due to the fact that by interacting at a quark level, there is a resulting fragmentation of the nucleon. Quarks cannot exist outside of bound

states and so a showering (or hadronisation) effect occurs producing a series of baryons and mesons.

3.2 Final State Interactions

As was discussed in section 3.1.1, the complex nuclear environment of heavy nuclei can significantly impact the primary neutrino interaction. Further impact is felt from the nuclear medium via final state interactions (FSI). The various interaction processes outlined above all result in a final state lepton alongside a hadronic final state of varying complexity, from a single proton created in a CCQE interaction to a complex shower of hadrons resulting from a high energy DIS event. These hadrons are produced within a dense nuclear medium and are therefore subject to a strong nuclear potential. The chances of final state hadrons (primarily protons, neutrons and pions) undergoing a further reinteraction with a nucleon before escaping the nucleus are relatively high (can be as large as 80% for pions in argon [147]).

The impact of FSI is clear, the hadrons that escape the nucleus and, subject to detector thresholds and efficiencies, are therefore observable, may not be the ones produced in the primary neutrino interaction. The FSI can alter both the particle type and its kinematics. Four types of interaction are typically considered - absorption, elastic scattering, charge exchange and pion production - and are nicely summarised in figure 3.4.

The modelling of FSI differs significantly between different generators primarily because the full quantum mechanical description of the nucleus is too complex to be achieved [147]. Many use an intranuclear cascade model in which each hadron is stepped through the nuclear medium with the mean free path used to assess the probability of interaction [115]. The effects of nucleon-nucleon correlations also have to be accounted for.

Recent measurements by the MINERνA [149, 150, 151], T2K [152, 153] and MicroBooNE [154] collaborations have studied how the use of transverse kinematic imbalance (TKI) variables can offer sensitivity to nuclear effects such as FSI. These variables

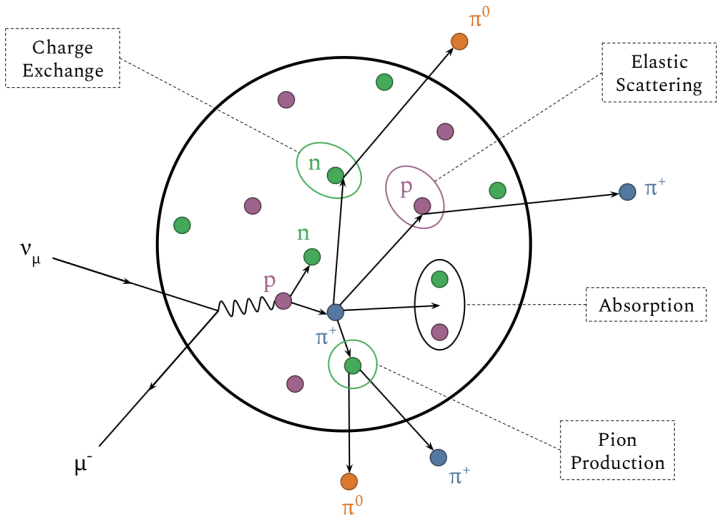


Figure 3.4: A schematic illustrating the final state interactions that hadrons can undergo whilst propagating through the nuclear medium. Figure from [148].

express the magnitude (δp_T) and direction ($\delta\alpha_T$) of any 'missing' momentum in the plane perpendicular to the incoming neutrino direction [155, 156, 157]. In the absence of FSI the angle this missing momentum makes with respect to the outgoing lepton, $\delta\alpha_T$, should be uniform, created only by the isotropic Fermi motion of the initial nucleon. Figure 3.5 shows the MicroBooNE measurement of the CC $1p0\pi$ cross section with respect to these TKI variables, clearly demonstrating that the data is not uniform in $\delta\alpha_T$. It also indicates that current FSI models describe the data well in some areas of the phase space but show some disagreement in other areas, particularly at high $\delta\alpha_T$. Measurements such as this, and further generalisation to full kinematic imbalance [157, 149, 151, 153, 158] offer significant input to better understand the impact of nuclear effects on neutrino-nuclear scattering and validate generator models accordingly.

3.3 NC π^0

The production of neutral pions via neutral-current interactions (or NC π^0) is possible via resonant, coherent and deep inelastic scattering processes. Its experimental signature is

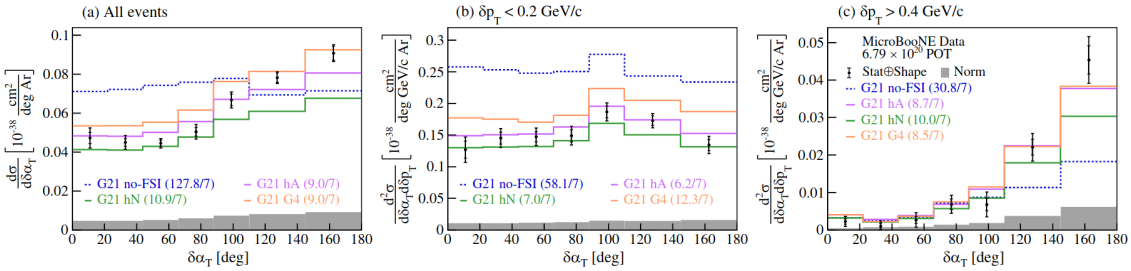


Figure 3.5: The CC $1p0\pi$ cross section as function of TKI variables in both single and double differential space as measured by the MicroBooNE experiment. Figure from [154].

characterised by the immediate decay of the neutral pion into a pair of photons. It is a channel of particular interest to modern experiments for a number of reasons. Firstly, it is a large background in ν_e appearance searches due to the potential similarities between a single electron electromagnetic shower in a CC ν_e event and a pair of photon electromagnetic showers in a NC π^0 event. This is particularly true for higher energy π^0 decays, where the opening angle between the two photons is reduced, leading to overlapping signatures and in detectors in which resolving two photons or separating photons from electrons is challenging. Secondly, pion production predictions rely on contributions from a number of processes and therefore making precision measurements of these channels is critical to reducing the interaction systematic uncertainties for future experiments. Finally, π^0 decays have a distinctive reconstructable invariant mass which can be a useful calibration tool for assessing energy reconstruction.

The first experimental observation of the production of pions, both charged and neutral, in neutrino neutral-current interactions was made in 1974 using the Argonne National Laboratory's 12 foot hydrogen and deuterium bubble chamber [167]. A total of 14 events ($7\ \nu_\mu p \longrightarrow \nu_\mu n\ \pi^+$ candidates and $7\ \nu_\mu p \longrightarrow \nu_\mu p\ \pi^0$ candidates) were observed with an expected background of 2.49 ± 0.73. This was followed by a series of observations and measurements in the 1970s and 1980s, reflective of a wider program of neutrino interaction measurements that followed on from the discoveries of the electron and muon neutrinos and the weak neutral-current. These results are summarised in table 3.1 and were predominantly expressed as ratios with respect to charged-current processes due to

Experiment	Year	Target	$\langle E_\nu \rangle$ (GeV)	Measurement	Value
BNL SC	1977 [159]	C / Al	Peak: 1-2	$\sigma(\nu_\mu \,\mathrm{NC}\, 1\pi^0)$ / $2\sigma(\nu_\mu \,\mathrm{CC}\, 1\pi^0)$	0.17 ± 0.04
Gargamelle	1977 [160]	CF_3Br	Peak: 2	$\sigma(\nu_\mu \,\mathrm{NC}\, 1\pi^0)$ / $2\sigma(\nu_\mu \,\mathrm{CC}\, 1\pi^0)$	0.45 ± 0.08
	1984 [161]			$\sigma(\nu_\mu \,\mathrm{NC}\, 1\pi^0 \,\mathrm{Coh})$	$(31 \pm 20) \times 10^{-40}\,\mathrm{cm}^2/\mathrm{nucleus}$
				$\sigma(\bar{\nu}_\mu \,\mathrm{NC}\, 1\pi^0 \,\mathrm{Coh})$	$(45 \pm 24) \times 10^{-40}\,\mathrm{cm}^2/\mathrm{nucleus}$
ANL BC	1981 [162]	Deuterium	Peak: 0.5	$\sigma(\nu_\mu \,\mathrm{NC}\, 1\pi^0) \,/\, \sigma(\nu_\mu \,\mathrm{CC}\, 1\pi^+)$	0.09 ± 0.05
Aachen-Padova	1983 [163]	Al	2	$\sigma(\nu_\mu \,\mathrm{NC}\, 1\pi^0 \,\mathrm{Coh})$	$(29 \pm 10) \times 10^{-40}\,\mathrm{cm}^2/\mathrm{Al}$
				$\sigma(\bar{\nu}_\mu \,\mathrm{NC}\, 1\pi^0 \,\mathrm{Coh})$	$(25 \pm 7) \times 10^{-40}\,\mathrm{cm}^2/\mathrm{Al}$
CHARM	1985 [164]	Marble ($CaCO_3$)	31	$\sigma(\nu_\mu \,\mathrm{NC}\, 1\pi^0 \,\mathrm{Coh})$	$(96 \pm 42) \times 10^{-40}\,\mathrm{cm}^2/\mathrm{nucleus}$
			24	$\sigma(\bar{\nu}_\mu \,\mathrm{NC}\, 1\pi^0 \,\mathrm{Coh})$	$(79 \pm 26) \times 10^{-40}\,\mathrm{cm}^2/\mathrm{nucleus}$
SKAT	1986 [165]	CF_3Br	7	$\sigma(\nu_\mu \,\mathrm{NC}\, 1\pi^0 \,\mathrm{Coh}) \,/\, \sigma(\nu_\mu \,\mathrm{CC})$	0.34 ± 0.12
				$\sigma(\nu_\mu \,\mathrm{NC}\, 1\pi^0 \,\mathrm{Coh})$	$(52 \pm 19) \times 10^{-40}\,\mathrm{cm}^2/\mathrm{nucleus}$
FNAL 15' BC	1986 [166]	Ne-H_2	20	$\sigma(\nu_\mu \,\mathrm{NC}\, 1\pi^0 \,\mathrm{Coh}) \,/\, \sigma(\nu_\mu \,\mathrm{CC})$	$(2.0 \pm 0.4) \times 10^{-4}$

Table 3.1: Summary of early NC π^0 results on a variety of targets and energies.

poor knowledge of the neutrino flux. These measurements were restricted by very limited statistics, indicated by the significant associated uncertainties, but did form a body of evidence for the existence of such interactions. The particular focus of many of these results was on the coherent production channel, with many of the experiments reporting coherent cross sections via the observation of an excess of forward going events above the predicted resonant spectrum according to the Rein-Sehgal model [136].

Interest in more precise measurements of neutrino interaction cross sections was driven by the confirmation of neutrino oscillations around the start of the millennium. Proposals for the second generation of oscillation experiments - to better measure the PMNS parameters and neutrino mass splittings - required better modelling of neutrino interactions. These efforts have only grown as it has become clear that our knowledge of neutrino interactions is incomplete, particularly on complex nuclear targets, and that this is becoming the limiting factor for modern oscillation experiments no longer limited by small sample sizes. In the last 20 years measurements have been made of the NC π^0 cross section on water, hydrocarbons, iron and argon, and are summarised in some detail in table 3.2. Whilst it has still been useful to report some measurements as ratios, to reduce the impact of correlated uncertainties, the increase in statistics and detector precision has allowed for more measurements of integrated cross section results.

Experiment	**Year**	**Target**	$\langle E_\nu \rangle$ **(GeV)**	**Measurement**	**Value**
K2K	2004 [168]	H_2O	1.3	$\sigma(\mathrm{NC}\,1\pi^0) \,/\, \sigma(\mathrm{CC\,Incl})$	$0.064 \pm 0.001(\mathrm{stat.}) \pm 0.007(\mathrm{sys.})$
MiniBooNE	2008 [169]	CH_2	0.808	$\sigma(\mathrm{NC}\,1\pi^0\,\mathrm{Coh}) \,/\, \sigma(\mathrm{NC}\,1\pi^0\,\mathrm{Coh} + \mathrm{Res})$	$(19.5 \pm 1.1(\mathrm{stat.}) \pm 2.5(\mathrm{sys.}))\%$
	2009 [170]			$\sigma(\nu_\mu\,\mathrm{NC}\,1\pi^0)$ and differentials	$(4.76 \pm 0.05(\mathrm{stat.}) \pm 0.76(\mathrm{sys.})) \times 10^{-40}\,\mathrm{cm}^2/\mathrm{nucleon}$
			0.664	$\sigma(\bar{\nu}_\mu\,\mathrm{NC}\,1\pi^0)$	$(1.48 \pm 0.05 \pm (\mathrm{stat.}) \pm 0.23(\mathrm{sys.})) \times 10^{-40}\,\mathrm{cm}^2/\mathrm{nucleon}$
				$\sigma(\nu_\mu\,\mathrm{NC}\,1\pi^0\,\mathrm{InCoh})$	$(5.71 \pm 0.08(\mathrm{stat.}) \pm 1.45(\mathrm{sys.})) \times 10^{-40}\,\mathrm{cm}^2/\mathrm{nucleon}$
				$\sigma(\bar{\nu}_\mu\,\mathrm{NC}\,1\pi^0\,\mathrm{InCoh})$	$(1.28 \pm 0.07(\mathrm{stat}) \pm 0.35(\mathrm{sys.})) \times 10^{-40}\,\mathrm{cm}^2/\mathrm{nucleon}$
NOMAD	2009 [171]	Carbon-based ($\langle A \rangle = 12.8$)	24.8	$\sigma(\mathrm{NC}\,1\pi^0\,\mathrm{Coh})$	$72.6 \pm 8.1(\mathrm{stat.}) \pm 6.9(\mathrm{sys.}) \times 10^{-40}\,\mathrm{cm}^2/\mathrm{nucleus}$
SciBooNE	2009 [172]	C_8H_8	1	$\sigma(\mathrm{NC}\,\pi^0) \,/\, \sigma(\mathrm{CC\,Incl})$	$(7.7 \pm 0.5(\mathrm{stat.}) \pm 0.5(\mathrm{sys.})) \times 10^{-2}$
				$\sigma(\mathrm{NC}\,\pi^0\,\mathrm{Coh}) \,/\, \sigma(\mathrm{CC\,Incl})$	$(0.7 \pm 0.4) \times 10^{-2}$
	2010 [173]		0.8	$\sigma(\mathrm{NC}\,\pi^0\,\mathrm{Coh}) \,/\, \sigma(\mathrm{CC\,Incl})$	$(1.16 \pm 0.24) \times 10^{-2}$
				$\sigma(\mathrm{CC}\,\pi^+\,\mathrm{Coh}) \,/\, \sigma(\mathrm{NC}\,\pi^0\,\mathrm{Coh})$	$0.14^{+0.30}_{-0.28}$
				$\sigma(\mathrm{NC}\,1\pi^0\,\mathrm{Coh}) \,/\, \sigma(\mathrm{NC}\,1\pi^0\,\mathrm{Coh} + \mathrm{Res})$	$(17.9 \pm 4.1)\%$
ArgoNeuT	2015 [174]	Ar	9.6	$\sigma(\nu_\mu\,\mathrm{NC}\,\pi^0) \,/\, \sigma(\nu_\mu\,\mathrm{CC})$	$0.094 \pm 0.022(\mathrm{stat.}) \pm 0.015(\mathrm{sys.})$
				$\sigma(\nu_\mu\,\mathrm{NC}\,\pi^0)$	$(7.1 \pm 1.7(\mathrm{stat.}) \pm 1.3(\mathrm{sys.})) \times 10^{-40}\,\mathrm{cm}^2/\mathrm{nucleon}$
			3.6	$\sigma(\bar{\nu}_\mu\,\mathrm{NC}\,\pi^0) \,/\, \sigma(\bar{\nu}_\mu\,\mathrm{CC})$	$0.042 \pm 0.022(\mathrm{stat.}) \pm 0.008(\mathrm{sys.})$
				$\sigma(\bar{\nu}_\mu\,\mathrm{NC}\,\pi^0)$	$(0.5 \pm 0.2(\mathrm{stat.}) \pm 0.1(\mathrm{sys.})) \times 10^{-40}\,\mathrm{cm}^2/\mathrm{nucleon}$
			9.6 / 3.6	$\sigma(\overset{(-)}{\nu}_\mu\,\mathrm{NC}\,\pi^0) \,/\, \sigma(\overset{(-)}{\nu}_\mu\,\mathrm{CC})$	$0.136 \pm 0.031(\mathrm{stat.}) \pm 0.017(\mathrm{sys.})$
				$\sigma(\overset{(-)}{\nu}_\mu\,\mathrm{NC}\,\pi^0)$	$(7.6 \pm 1.7(\mathrm{stat.}) \pm 1.4(\mathrm{sys.})) \times 10^{-40}\,\mathrm{cm}^2/\mathrm{nucleon}$
MINOS	2016 [175]	Iron and Carbon ($\langle A \rangle = 48$)	4.9	$\sigma(\nu_\mu\,\mathrm{NC}\,\pi^0\,\mathrm{Coh})$	$77.6 \pm 5.0(\mathrm{stat.})^{+15.0}_{-16.8}(\mathrm{sys.}) \times 10^{-40}\,\mathrm{cm}^2/\mathrm{nucleus}$
MINERνA	2016 [176]	C_8H_8	4.5	$\sigma(\mathrm{NC}\,\pi^0\,\mathrm{Hydrogen\ Diffractive})$	$0.26 \pm 0.02(\mathrm{stat.}) \pm 0.08(\mathrm{sys.}) \times 10^{-39}\,\mathrm{cm}^2/\mathrm{CH}$
NOνA	2019 [177]	Hydrocarbon ($\langle A \rangle = 13.8$)	2.7	$\sigma(\mathrm{NC}\,\pi^0\,\mathrm{Coh})$	$13.8 \pm 0.9(\mathrm{stat.}) \pm 2.3(\mathrm{sys.}) \times 10^{-40}\,\mathrm{cm}^2/\mathrm{nucleus}$
MicroBooNE	2022 [178]	Ar	0.8	$\sigma(\mathrm{NC}\,1\pi^0\,0\pi^\pm)$	$(1.243 \pm 0.185(\mathrm{sys.}) \pm 0.076(\mathrm{stat.})) \times 10^{-38}\,\mathrm{cm}^2/\mathrm{Ar}$
				$\sigma(\mathrm{NC}\,1\pi^0\,0\pi^\pm\,0p)$	$(0.444 \pm 0.098(\mathrm{sys.}) \pm 0.047(\mathrm{stat.})) \times 10^{-38}\,\mathrm{cm}^2/\mathrm{Ar}$
				$\sigma(\mathrm{NC}\,1\pi^0\,0\pi^\pm\,1p)$	$(0.624 \pm 0.131(\mathrm{sys.}) \pm 0.075(\mathrm{stat.})) \times 10^{-38}\,\mathrm{cm}^2/\mathrm{Ar}$

Table 3.2: Summary of modern NC π^0 results on a variety of targets and energies.

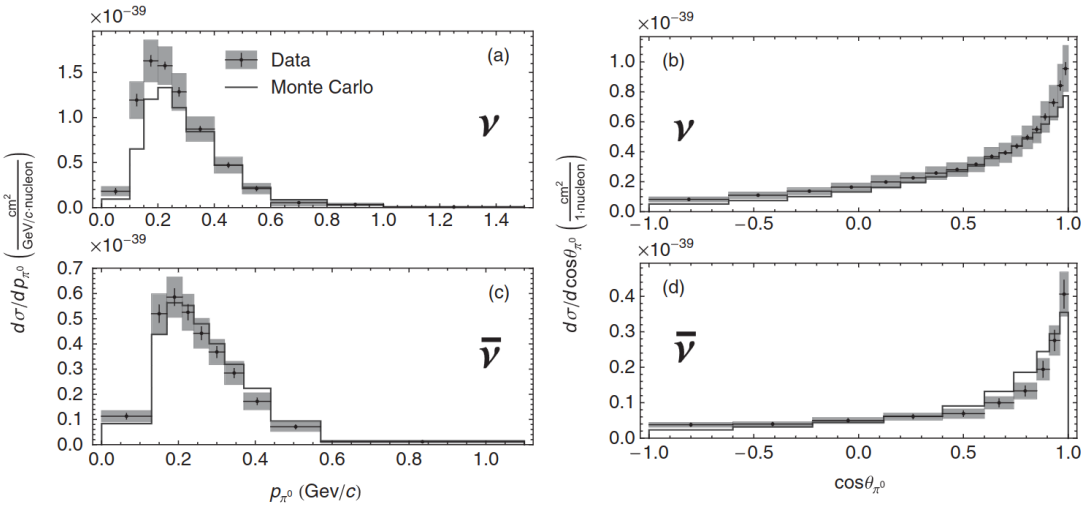


Figure 3.6: The NC $1\pi^0$ cross section measured by the MiniBooNE experiment with respect to the neutral pion's momentum and direction. Figure from [170].

The MiniBooNE experiment was the first to report a differential cross section in terms of the pion's momentum and direction, as shown in figure 3.6. Whilst their measurement indicated that the Rein-Sehgal models did not fully describe the observed cross section, the differential measurement in terms of the pion direction conclusively demonstrated a non-zero coherent contribution peaking at forward directions. The measurement of NC π^0 production was particularly important to the MiniBooNE experiment as an in-situ constraint on the largest background to their ν_e appearance results (see figure 2.8), driven by the identical appearance of electron and photon showers in the mineral oil Cherenkov detector.

The measurements of most relevance to this thesis are those of the ArgoNeuT [174] and MicroBooNE [178] collaborations. The ArgoNeuT result represented the first measurement of NC π^0 on argon and included separate measurements of the neutrino and antineutrino components. Their measurements were consistent with the predictions from both the GENIE and NuWro generators as shown in figure 3.7. In order to make such a measurement in a small ($47.5 \times 40 \times 90\,\mathrm{cm}^3$) detector, the collaboration developed a series of novel energy correction measures to account for uncontained photon showers.

The MicroBooNE experiment consisted of a much larger argon volume and collected a

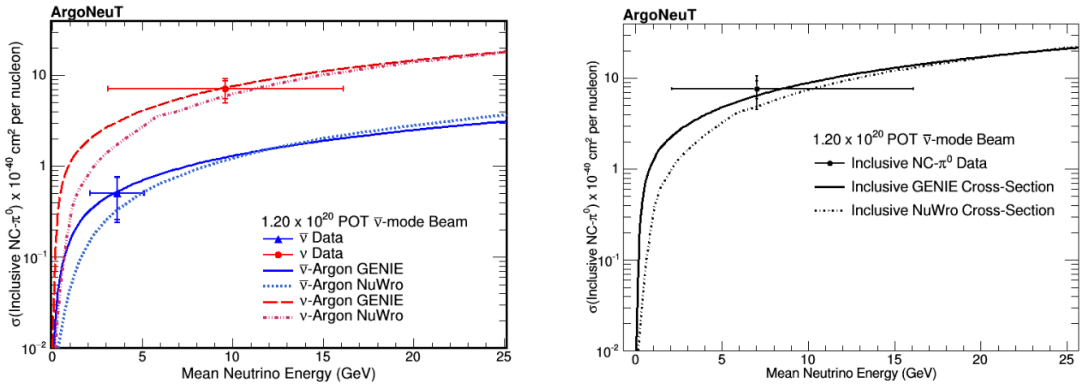


Figure 3.7: The NC $1\pi^0$ cross section measured on argon by the ArgoNeuT experiment and compared to predictions by the GENIE and NuWro generators. Figure from [174].

larger dataset of neutrino-argon interactions. Their measurement of the integrated cross section in three different channels (semi-inclusive, $0p$ and $1p$) gave further inspection of the model predictions from a range of generators (see figure 3.8). Whilst all the generators consistently over-predicted the data the results were within the expressed uncertainties. A simultaneous χ^2 test to the two exclusive channels indicated that the NEUT prediction best matched the data, but only marginally so. As the measurement that most closely matches the SBND scenario (same target and beam), it is instructive to inspect the systematic uncertainties associated with the result. Figure 3.9 shows that the statistical uncertainty was not dominant in any of the three channels (although was not negligible either) whilst the largest systematic uncertainties resulted from the modelling of the neutrino flux and interactions. In chapter 8 this thesis will assess the capabilities of the SBND experiment to make a precision measurement of this channel.

During the final preparation of this thesis the MicroBooNE collaboration published updated results in this channel, including their first differential and double-differential measurements. Their results indicate a continuation of the theme of over-prediction by the generators, especially at very forward-going angles [179].

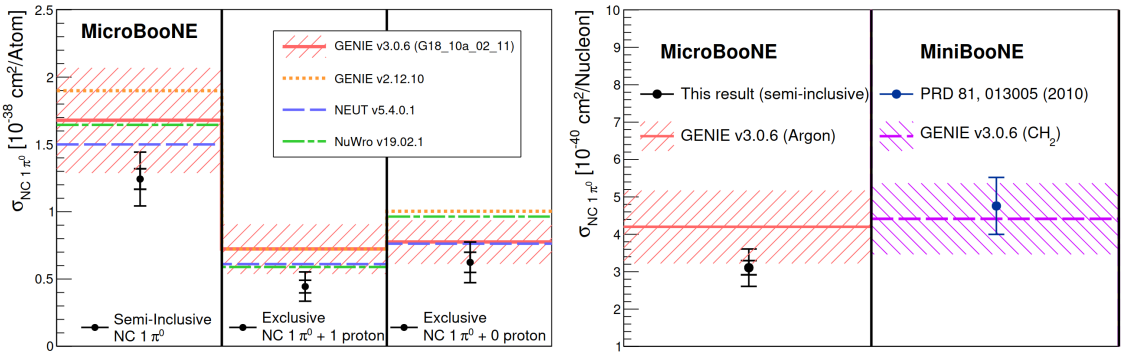


Figure 3.8: The NC $1\pi^0$ cross section measured on argon by the MicroBooNE experiment and compared to a range of generator predictions (left) and the equivalently scaled MiniBooNE result (right). Figures from [178].

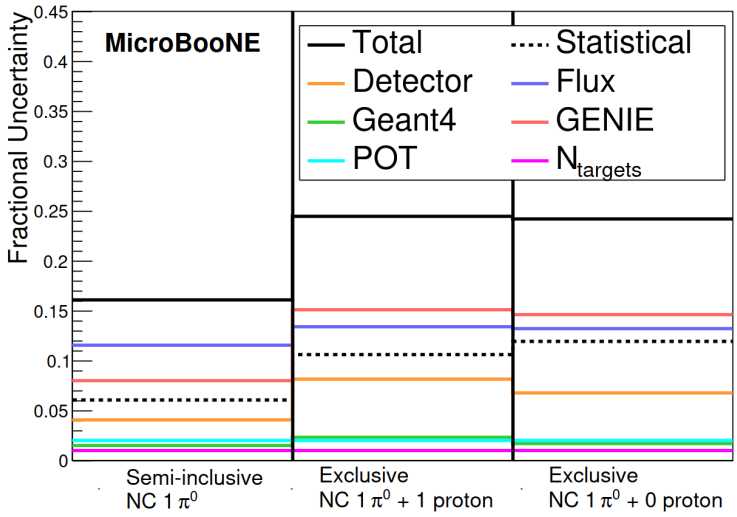


Figure 3.9: The uncertainties associated with the NC $1\pi^0$ cross section measurements presented by the MicroBooNE experiment. Figure from [178].

Chapter 4

SBND and LArTPC Technology

This chapter describes the Short-Baseline Near Detector, its hardware, physics goals and role in the Short-Baseline Neutrino program. The chapter necessarily describes work carried out over a significant period of time by members of the SBND and SBN collaborations in planning and building the detector. It also covers the general principles behind, and a brief history of, LArTPC neutrino detector technology. My personal contributions are limited to brief references in section 4.4.3 to installation and commissioning of the bottom CRT panels and the CRT## project. This work was completed alongside and relied heavily on work and support from Michelle Stancari, Lauren Yates, Erin Yandel and Vu Chi Lan Nguyen amongst others.

4.1 Short-Baseline Neutrino Program

The Short-Baseline Neutrino program (SBN) is an experiment hosted by the Fermi National Accelerator Laboratory (FNAL or Fermilab) in Batavia, Illinois. As shown in figure 4.1, it consists of a trio of LArTPCs: SBND, MicroBooNE and ICARUS, located at 110 m, 470 m and 600 m along the Booster Neutrino Beam (BNB) respectively. Together they will investigate the low energy excess (LEE) reported by the LSND and MiniBooNE

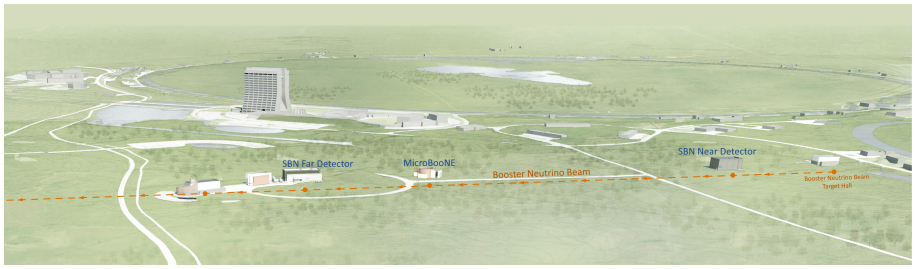


Figure 4.1: The location of the three SBN detectors along the BNB at Fermilab. Figure from [91].

experiments and search for the existence of sterile neutrinos, a possible explanation for the LEE. The use of a consistent target material and detector technology between the three detectors will serve to constrain systematics and improve the precision of the final results [180]. Alongside its role as the near detector for the SBN program, SBND has a rich standalone physics program focused on precision neutrino-argon cross section measurements and a variety of beyond the standard model searches. This is explored in detail in section 4.5.

4.2 Booster Neutrino Beam

The Fermilab accelerator complex follows a traditional method of accelerating protons; first via a 400 MeV linear accelerator followed by a series of circular accelerators of increasing energy. The first of these circular accelerators is the booster synchrotron, which accelerates protons to 8 GeV before they are extracted and, in the case of the BNB, collided with a beryllium target. The interactions of protons with the beryllium produces secondary hadrons, predominantly pions. A 1.5 T magnetic horn with a pulsed current peaking at 174 kA focuses a beam of positively charged pions, defocusing other hadronic products in the process. This is followed by a 45 m decay pipe within which the desired outcome is for the pions to decay via their dominant decay channel:

$$\pi^+ \rightarrow \mu^+ \nu_\mu \tag{4.1}$$

to produce muon neutrinos. At the end of the decay pipe is a steel and concrete beam stop

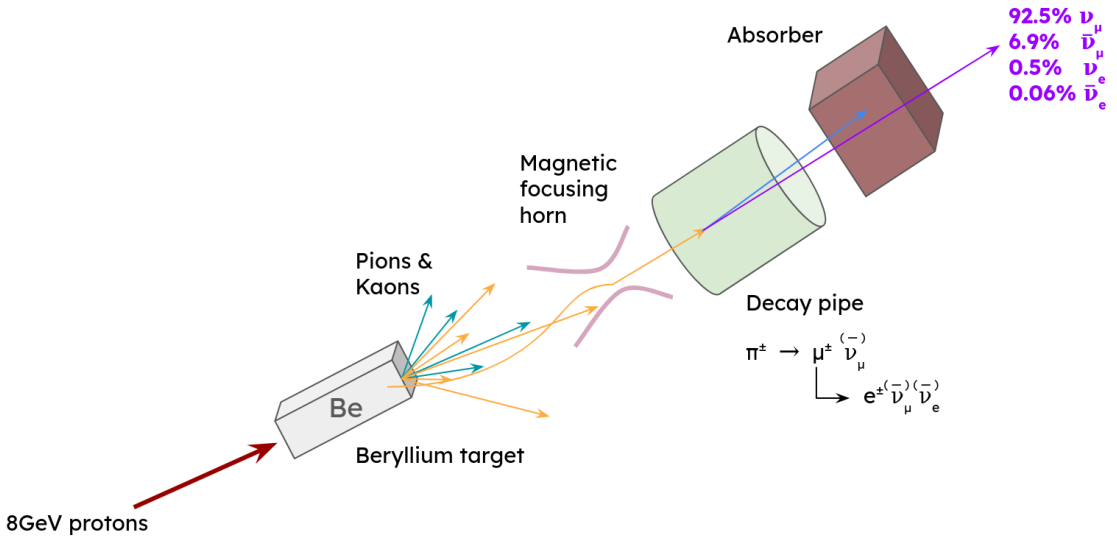


Figure 4.2: A schematic diagram showing the elements involved in the production of the BNB from the delivery of 8GeV protons from the booster synchrotron.

designed to absorb the muons and any other remaining charged particles [181]. Figure 4.2 shows a schematic layout of the elements described in the BNB production.

At this stage, the resulting beam is highly pure in muon neutrinos ($\sim$ 92.5 %). In reducing significance there is some contamination from muon antineutrinos ($\sim$ 6.9 %), electron neutrinos ($\sim$ 0.5 %) and electron antineutrinos ($\sim$ 0.06 %) produced by decays of other charged mesons (negative pions, charged kaons etc) and decay-in-flight of the muons produced in the original pion decay. The relative fluxes of all four relevant neutrino flavours are shown in figure 4.3 for the front faces of each of the SBN detectors. The beam peaks below 1 GeV with an average energy of $\sim$ 800 MeV but also has a significant tail into the multi-GeV region. The structure of the neutrino beam is driven by the original structure of the proton bunches in the booster synchrotron. Protons arrive at the target in spills of $\sim$ 1.6 μs, consisting of 81 bunches separated by $\sim$19 ns each with an intrinsic width of around 2 ns. The resulting neutrino beam arrives at the SBN detectors with the same structure. We refer to each 2 ns group of neutrinos as a bucket.

Due to the manner in which the neutrino beam is produced, the exposure is normally measured in units of protons-on-target (POT). This is calculated by using beam instrumen-

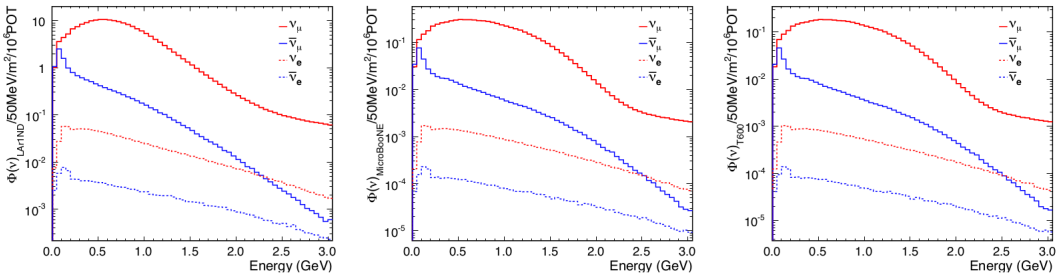


Figure 4.3: The BNB flux at the front face of the three SBN detectors: SBND (left), MicroBooNE (centre) and ICARUS (right). Figure from [180].

tation to account for the number of protons colliding with the target from each bunch. The BNB operates at around 5×10^{12} POT per spill with a maximum spill rate of 5 Hz. SBND's planned operation of 3 years should result in an exposure of around 1×10^{21} POT and it is this figure that will be used to scale the physics studies presented in this thesis.

4.3 Liquid Argon Time Projection Chambers

Liquid Argon Time Projection Chambers (LArTPCs) are often described as electronic bubble chambers. First proposed by Carlo Rubbia in 1977, they provided a new technology which, for the first time, combined the large mass of traditional 'counting' experiments like those of Cowan, Reines and Davis with the precision imaging resolution of the bubble chamber. This meant they could simultaneously deliver significant statistics whilst also providing critical topological and calorimetric information on the final state particles [182].

Liquid Argon has a number of significant benefits as a target material [182, 183]:

- its relative availability makes more affordable than other options
- its high density ($1.4\,\mathrm{gcm}^{-3}$) provides a large interaction rate in relatively modest volumes
- its high electron mobility and low electron reattachment allows for large drift lengths and electron yield given reasonable purity.

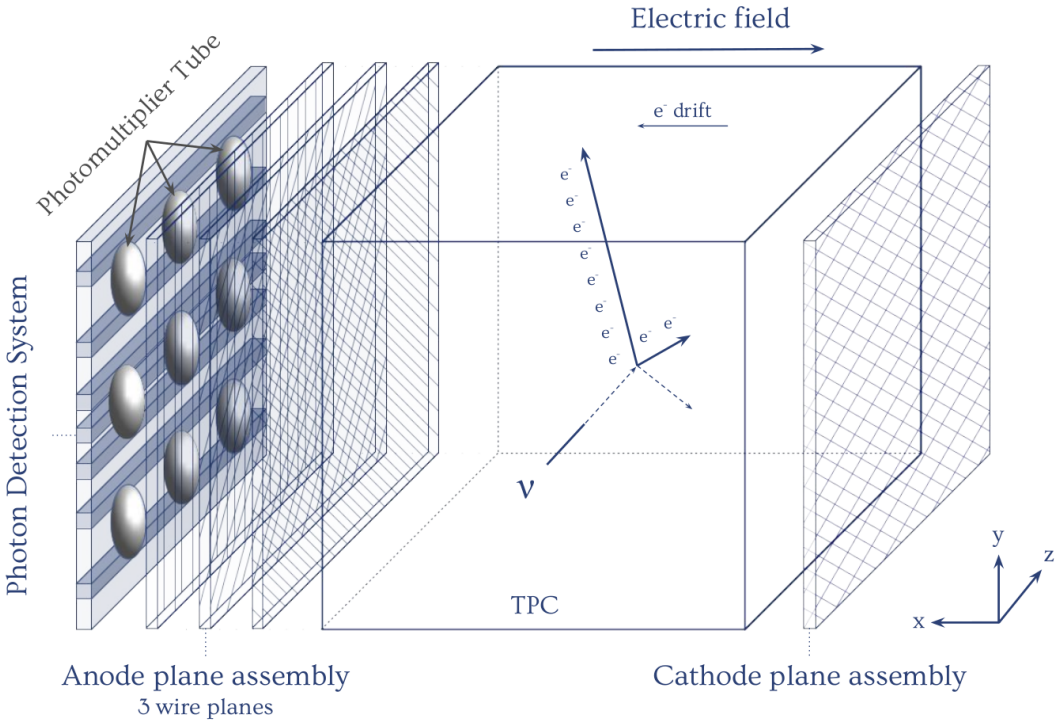


Figure 4.4: A typical LArTPC setup, consisting of an electric field established between a cathode and anode, a trio of wire readout planes at the anode and a photon detection system located behind the wire readout. The usual right handed coordinate system is shown: z being the beam direction, y the vertical and x the drift. Figure from [148].

The basic principle on which time projection chambers operate is shown in figure 4.4. The entire ionisation pattern produced by the passage of charged particles through the medium, in this case LAr, is drifted together under the influence of a strong electric field to some form of electronic readout. The electric field is created by applying a potential difference between an anode and cathode. Often a field cage is used to surround the TPC and incrementally step the voltage between the cathode and anode to maintain a uniform electric field [184]. The anode is traditionally instrumented with a series of wire planes on which current pulses are induced. The final plane is referred to as the collection plane on which a uni-polar response is recorded as the drift electrons are collected. Any preceding planes are referred to as induction planes, they are held at potentials such that they are transparent to the drift electrons ensuring their continued passage to the collection plane. As they pass by, however, they do induce a smaller bi-polar signal on the induction wires.

The signals induced on each plane form a two-dimensional projection of the ionisation charge pattern in terms of drift time and wire number. The wire planes are oriented at different angles with respect to the vertical and therefore represent three different projections but with a shared drift time coordinate. This degeneracy can be used to combine their signals into a 3D image of the event. The time recorded on the wires is naturally a combination of the ionisation time (t_0) and the x position. Hence, whilst the drift time can be used to give the relative locations of the charge depositions in the drift (x) direction, the absolute position relies on an externally provided knowledge of when the ionisation occurred. Alongside the ionisation response, a charged particle will also produce around 80,000 scintillation photons per cm at a wavelength to which the argon is transparent. This prompt scintillation signal can provide the t_0 or interaction time as well the potential for a complementary energy estimation [183, 185].

4.3.1 Signal Production

The manner in which a particle deposits energy in a medium depends upon its momentum. Highly relativistic particles will undergo radiative processes such as bremsstrahlung and pair production resulting in the formation of electromagnetic showers of particles with

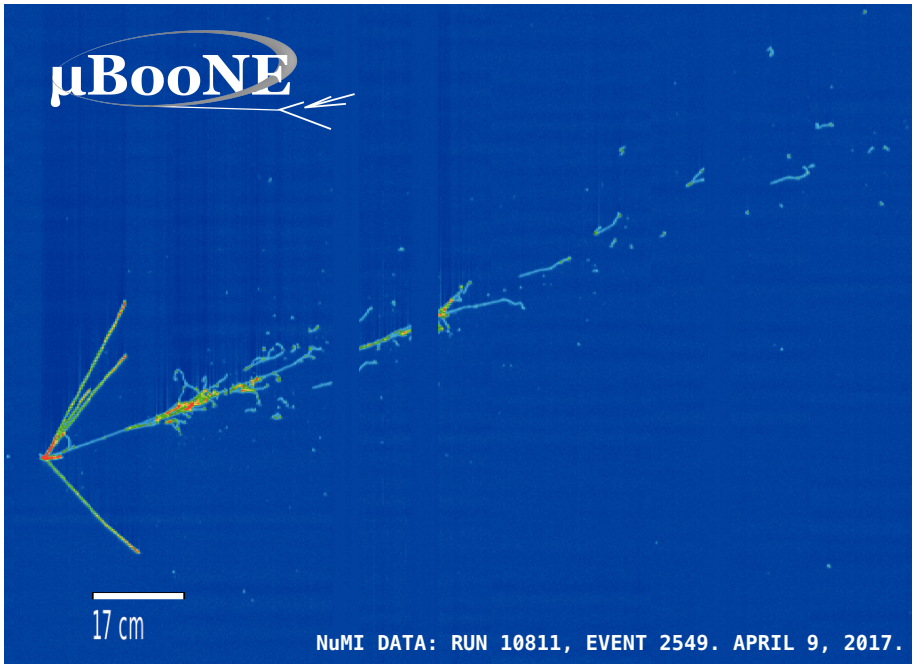


Figure 4.5: A section of an event display from the MicroBooNE detector showing a neutrino interaction from the NuMI beam resulting in both direct track-like and electromagnetic shower-like activity. Figure from [186].

decreasing energies. In the typical energy regime accessed by liquid argon neutrino experiments (including SBND) only electrons (positrons) and photons result in such an effect. The other prevalent particles: muons, charged pions and protons will primarily deposit their energy through direct ionisation and excitation of the argon atoms [32]. These two processes result in distinctively different signatures in the charge readout patterns. Figure 4.5 demonstrates this via a section of a 2017 MicroBooNE event display showing the production of a series of track-like particle signatures as well as a large electromagnetic shower most likely resulting from an electron produced in charged-current electron neutrino interaction.

The processes that result in the production of ionisation electrons and scintillation photons from the particle's energy depositions are summarised in figure 4.6. The energy transferred to the argon atom can either result in its excitation or ionisation, the latter resulting in the production of an ionisation electron and a positive argon ion. In the absence of an electric field, all ionised electrons and argon ions would undergo a process called recombination,

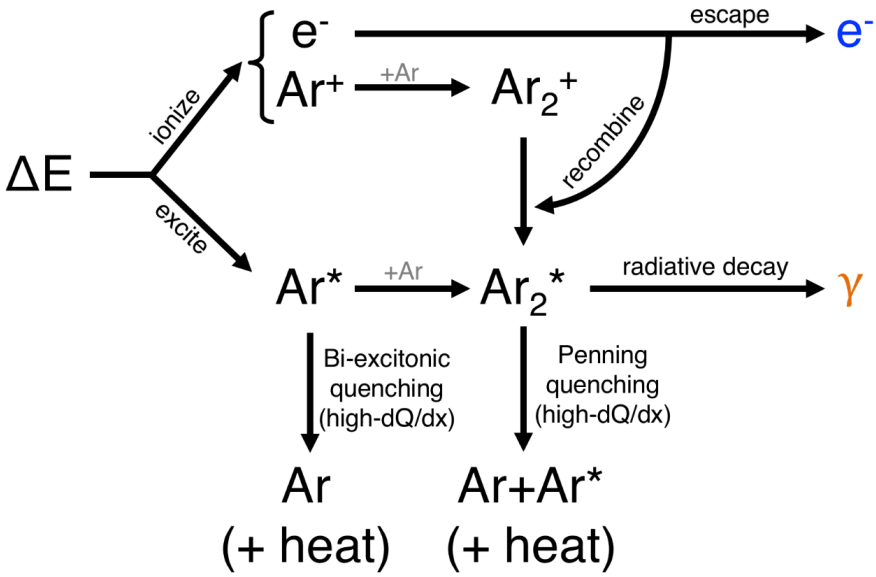


Figure 4.6: The mechanisms of production of ionisation electrons and scintillation photons in liquid argon. Figure from [185].

forming excited Argon dimers (Ar_2^*). These dimers then de-excite, resulting in the emission of a scintillation photon [187]. However, the introduction of an electric field in a LArTPC results in the 'escape' of some fraction of these electrons and their resulting drift to the charge readout planes. This, in turn, naturally reduces the scintillation light yield due to the lower rate of formation of argon dimers. This inverse relationship is known as charge-light anti-correlation. The designed nominal field strength for SBND of 500 V/cm results in approximately equal yields of charge and light as is visible in figure 4.7.

A good modelling of recombination is critical to a complete simulation and reconstruction of neutrino interactions in LArTPC detectors, particularly the calorimetric reconstruction of energy. In order to accurately convert the amount of observed ionisation charge to the original energy deposition, one must know what proportion of electrons underwent recombination. This relationship is usually expressed as:

$$\frac{\mathrm{d}E}{\mathrm{d}x} = \frac{W_{ion}}{\mathcal{R}} \frac{\mathrm{d}Q}{\mathrm{d}x} \tag{4.2}$$

where $\mathrm{d}Q/\mathrm{d}x$ (e/cm) is the measured charge deposition per unit length along the particle's

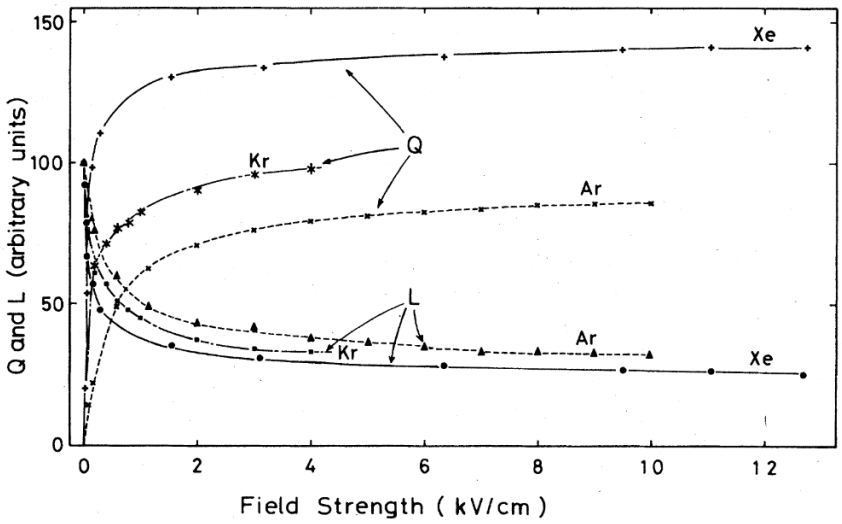


Figure 4.7: The relative charge and light yields in liquid argon (as well as krypton and xenon) for varying electric field strength. Figure from [188].

trajectory, $\mathrm{d}E/\mathrm{d}x$ (MeV/cm) is the equivalent energy deposition, $W_{ion} = 23.6\,\mathrm{eV}/e$ is the energy required to singly ionise one argon atom and $\mathcal{R}$ is the recombination factor, the proportion of electrons that survived recombination [189].

There are two main models that have been used to describe recombination in LArTPCs. The first is Birk's law:

$$\mathcal{R}_{Birks} = \frac{A}{1 + k(\mathrm{d}E/\mathrm{d}x)/\mathcal{E}} \tag{4.3}$$

where A and k ($\mathrm{kV/cm \cdot g/MeV\,cm^2}$) are fitted model parameters, $\mathrm{d}E/\mathrm{d}x$ is the stopping power of the particle in the material ($\mathrm{MeV\,cm^2/g}$) and $\mathcal{E}$ is the electric field (kV/cm). The second is the box model:

$$\mathcal{R}_{Box} = \frac{1}{\beta(\mathrm{d}E/\mathrm{d}x)} \ln\left(\alpha + \beta\frac{\mathrm{d}E}{\mathrm{d}x}\right) \tag{4.4}$$

where $\alpha = 1$ and β (cm/MeV) are fitted model parameters and $\mathrm{d}E/\mathrm{d}x$ is the particle's energy deposition per unit length (MeV/cm). Both of these models contain a significant number of assumptions and, as such, should be treated as phenomenological expressions, valid in the parameter space for which they have been verified [190]. Both approaches consider the group behaviour of electrons and ions rather than the direct recombination of

a specific electron ion pair [189].

Birk's law evaluates a cylindrical volume around the ionised trajectory [191], whilst the box model of Thomas and Imel [192] evaluates the charge density in a box-shaped volume. Birk's law is known to break down when considering highly ionising particles, whereas the box model is mathematically compatible with such particles. However, the box model diverges from the data at low values of $\mathrm{d}E/\mathrm{d}x$. The ArgoNeuT collaboration's measurement of recombination demonstrated that a compromise can be found with a modified box model [189]. The term 'modified' refers to allowing the parameter α in equation 4.4 to take values other than 1. ArgoNeuT's results suggested values of $\alpha = 0.93 \pm 0.02$ and $\beta = 0.319 \pm 0.003\,(\mathrm{MeV/cm})^{-1}$ agreed well with data across all relevant $\mathrm{d}E/\mathrm{d}x$ values. They also provided a value of $\beta' = 0.212 \pm 0.002\,\mathrm{kV/cm\,g/cm^3}\,(\mathrm{MeV/cm})^{-1}$. This value represents a more fundamental quantity from which other experiments can extract a value of β specific to their experimental configuration using $\beta = \beta'/\rho\mathcal{E}$ where ρ is the argon density and $\mathcal{E}$ is the electric field strength.

The secondary signal of scintillation light is shown in figure 4.6 to be a product of two different routes, both of which result in the formation of an excited argon dimer state, Ar_2^*. This dimer subsequently de-excites to produce two argon atoms and a 128 nm vacuum ultraviolet (VUV) scintillation photon. The dimer can exist in either a singlet or triplet state which affects the rate of de-excitation resulting in a fast ($\tau \sim 6\,\mathrm{ns}$) and slow ($\tau \sim 1.6\mu\mathrm{s}$) component to the emitted light [180]. As with the ionisation signal, the presence of impurities in the argon, specifically oxygen or nitrogen, can cause a reduction of the light yield. This occurs via a quenching process in which the dimer state's energy is absorbed without the emission of a photon [193, 194]. The density of the ionisation signal has a direct impact on recombination rates and therefore on the amount of light emitted, hence, the light yield also depends on the $\mathrm{d}E/\mathrm{d}x$ of the charged particle producing the signature [195, 187].

4.3.2 Signal Propagation

The propagation of the charge and light signals through the liquid argon volume is subject to a number of transport effects before they reach their respective readouts.

4.3.2.1 Ionisation Charge

The energy required to singly-ionise an argon atom is 23.6 eV [196] and given that a minimum ionising particle will deposit approximately 2.12 MeV/cm [32] this results in large 'clouds' of electrons drifting together from the ionisation trail. The strong electric field in the TPC causes these electrons to drift through the argon. At the intended SBND field strength of 500 V/cm the electron drift velocity will be around 1.6 mm/μs. Whilst drifting, the electrons are subject to a number of effects that can modify their propagation, primarily diffusion, attenuation by electronegative impurities and the space charge effect (SCE).

Diffusion

The drifting electron clouds undergo a stochastic diffusion process with components both longitudinal, D_L, and transverse, D_T, with respect to the drift direction. This causes the size of the electron 'cloud' to expand between the point of ionisation and their detection at the anode planes. This diffusion reduces the spatial resolution of the detector by 'blurring out' the images captured by the wires and can also be seen to bias the recorded charge deposition affecting energy scale calibration and particle identification techniques [197].

The time spread (or width) of a current pulse on a wire (σ) can be modelled as:

$$\sigma^2(t) = \sigma_0^2 + \left(\frac{2D_L}{v_d^2}\right) t \tag{4.5}$$

where σ_0 is the initial intrinsic width of the electron cloud, v_d is the drift velocity, t the drift time and D_L the longitudinal diffusion coefficient [198].

Longitudinal diffusion can be measured in LArTPCs by considering the dependence of

pulse widths on drift time [198, 199]. This can only be done with a sample of tracks for which the t_0 is well-known, usually using auxiliary detector systems or tracks that cross the entire detector volume. Transverse diffusion, i.e. that which happens in the plane parallel to the readout, cannot be measured in situ by LArTPCs due to the intrinsic resolution limit created by the wire pitch. In current LArTPC experiments the pitch has been of the order of a few mm while the scale of the transverse diffusive effect across a drift distance of 2 m at 500 V/cm would be expected to be around 1 mm. A value for the transverse diffusion coefficient can be extracted from the longitudinal coefficient using knowledge of the electric field (E):

$$\frac{D_L}{D_T} = 1 + \frac{E}{\mu}\frac{\partial \mu}{\partial E} \tag{4.6}$$

where μ is the electron mobility ($\mu = v_d \,/\, E$) [200]. The most recent measurement by the MicroBooNE collaboration gave a value of $D_L = 3.74^{+0.28}_{-0.29}\,\mathrm{cm}^2/\mathrm{s}$ [198], which agrees well with that reported earlier by ICARUS $D_L = 4.74 \pm 1.01\,\mathrm{cm}^2/\mathrm{s}$ [199], and corresponds to a value of $D_T = 5.85^{+0.62}_{-0.33}\,\mathrm{cm}^2/\mathrm{s}$ for the transverse component.

Attenuation

Electronegative impurities such as oxygen or water can result in the attenuation of drift electrons via their attachment to such molecules. This reduces the amount of charge that reaches the anode plane readouts, particularly for events close to the cathode plane which, by definition, have a longer drift time. The drift electron attenuation can be expressed as exponentially dependent on the drift time (t):

$$\frac{Q_{obs}}{Q_0} = e^{-t/\tau} \tag{4.7}$$

where Q_{obs} is the charge observed at the anode, Q_0 is the original ionisation charge and τ is a time constant referred to as the electron lifetime. Keeping the concentration of these contaminants very low is critical to the successful operation of LArTPCs. Clearly, electron lifetime is particular to each LArTPC and depends on their argon supply, filtration and purification process, and can evolve during the operation of an experiment. The argon purity can be measured and monitored in a number of ways such as purity monitors, gas analysers, laser tracks or via minimally ionising cosmic muon tracks [184, 201, 202]. Modern

LArTPCs with drift distances on the scale of a few metres require lifetimes greater than a few milliseconds in order to preserve enough charge from deposits at the cathode. The MicroBooNE experiment demonstrated a stable lifetime above 18 ms [202], far exceeding their goal lifetime of 3 ms and representing a 12 % loss of charge across the whole drift distance at 273 V/cm.

Space Charge Effect

For each ionisation electron produced so is an argon ion. These argon ions also drift under the influence of the electric field but with two key points of distinction from the electrons. Firstly, their positive charge means they drift towards the cathode and secondly, due to their mass being many orders of magnitude larger than that of the electrons, they drift at a much slower rate. This leads to a build up of argon ions in the active volume with increasing density towards the cathode. This varying charge profile leads to non-uniformities in the electric field known as the space charge effect. There are two main ways in which this affects the reconstruction and must be accounted for. The non-uniform field results in distortions of the recorded trajectories; an effect which becomes more severe nearer the cathode. Simultaneously, the dependence of recombination on the electric field means that the relation between the observed charge and deposited energy can no longer be treated as constant across the whole detector volume.

4.3.2.2 Scintillation Light

For the scintillation light there are two key effects that modify their propagation. Firstly, Rayleigh scattering lengthens the mean free path of the photons delaying their arrival and altering their direction. Secondly, scintillation light can also be absorbed by impurities in the argon in a process distinct from the quenching which prevented the initial emission of the photon. The Rayleigh scattering length for 128 nm photons in liquid argon has been measured with quite a wide range of values reported, the most recent of which being $\lambda_{Rayleigh} \simeq 99\,\text{cm}$ [203].

4.3.3 History of LArTPCs

The first demonstration of this technology in action was provided by the ICARUS collaboration which operated a 3 ton prototype at CERN in the early 1990s [204]. They were able to show that they had overcome the key technical hurdles in order to use this novel detector type for physics. Primarily, the initial and maintained purity of the LAr at a contaminant level of 0.1 ppb ensured a long electron lifetime, the wire readout planes were operated at voltages that allowed for clear signals whilst maintaining the transparency of the induction plane and low-noise preamplifiers allowed for a signal to noise ratio of 6 on the induction plane and 10 on the collection plane. This first operation also demonstrated the significant particle identification capabilities of LArTPCs via range and energy deposition measurements.

The ICARUS collaboration followed this success with the design and operation of a 500 ton detector, ICARUS-T600, the first large-scale LArTPC. This was first operated in a demonstration run of 3 months in Pavia in 2001 [205] followed later by installation at the Gran Sasso underground laboratory (LNGS) and operation in the CERN to Gran Sasso neutrino beam (CNGS). Alongside a demonstration of running a detector of this size, ICARUS also made measurements of crucial detector physics such as the recombination [190] and diffusion effects [199] as well as an electron neutrino appearance search in the phase space of the LSND anomaly [206] and a refutation of the OPERA result claiming observation of superluminal neutrinos in the same beam [207].

The USA's Liquid Argon neutrino project began with ArgoNeuT, a tabletop detector of roughly 170 litres of active LAr. ArgoNeuT operated in the NuMI beamline at Fermilab for a few months in 2009 and 2010 situated directly in front of the MINOS near detector, which itself was used to measure muon range given the small argon volume. ArgoNeuT was the first experiment to publish neutrino cross sections on argon, including inclusive charged-current ν_μ [208] and ν_e [209] and of most relevance to this thesis, neutral-current neutral pion production [174]. Subsequently, the LArIAT experiment operated in the Fermilab Test Beam Facility in the mid-2010s. LArIAT consisted of a small LArTPC re-using the ArgoNeuT cryostat, with a series of auxiliary detector systems monitoring the beamline.

Placing such a detector in a testbeam allowed for large samples of pions, muons, electrons, protons and kaons to be collected and analysed to significantly deepen the understanding of LArTPC detector responses to different particle types and momenta in a controlled test beam environment [210]. Publications have already been made on low-energy electron calorimetry [185] and the π^--Ar scattering cross section [211].

Following the success of ArgoNeuT, the MicroBooNE detector was operated in the Booster Neutrino Beam between 2015 and 2020 as the first of the SBN program detectors. To date, MicroBooNE has already published a vast series of neutrino cross section results, BSM searches and the first set of LEE results, with no indication of the MiniBooNE / LSND anomaly in either electron-like [103] or photon-like channels [104]. Finally, the ICARUS T600 detector, which was refurbished and upgraded at CERN after its initial operation at LNGS, was transported to Fermilab to form the far detector for the SBN program. ICARUS arrived at Fermilab in 2017 and in the time since has been installed and commissioned, and began taking physics data in 2022 [212].

Meanwhile in Europe, the Deep Underground Neutrino Experiment (DUNE) collaboration have run a series of prototypes for the DUNE Far Detector modules at CERN's neutrino platform. Phase one of this program involved the operation of a horizontal-drift single-phase detector as well as a vertical-drift dual-phase detector in a test beam facility. Phase two will test upgraded horizontal and vertical drift designs both in single liquid phase. The design, construction and operation of these detectors has, and will continue to, inform the preparations for the DUNE Far Detector modules, which will be, by a margin, the largest LArTPCs constructed, at an argon mass of approximately 10 kton per module [213].

4.4 Short-Baseline Near Detector

The Short-Baseline Near Detector (SBND) will be the final element of the SBN program to commence operations. A purpose built detector with the benefits of the experience of previous large-scale LArTPCs, SBND will be situated at 110m along the BNB. Alongside the core TPC, SBND will consist of two major auxiliary detector systems: the photon

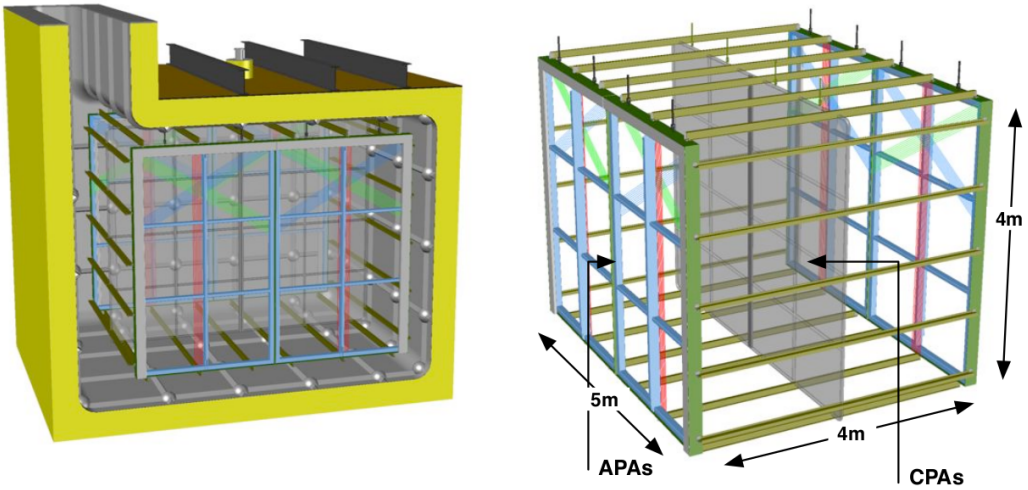


Figure 4.8: The design of the SBND cryostat and TPCs. Indicative arrangements of sections of the wire planes are shown in red (collection plane), blue and green (induction planes). Figure from [180].

detection system (PDS) and the cosmic ray taggers (CRT).

4.4.1 TPC

The SBND is a 112 ton, 4m × 4m × 5m detector comprising of 2 TPCs with a shared central cathode and a 2m drift length on either side. Retaining small drift lengths in larger detectors via this kind of modularisation prevents diffusion and attenuation from becoming more significantly problematic, as well as reducing the cathode voltage required to attain the nominal drift field. Figure 4.8 demonstrates the design of the TPCs and their location within the cryostat.

The anode planes situated on either side of the detector contain three planes of wires, two induction planes, oriented at ±60° with respect to the vertical (U and V views) and a collection plane oriented vertically (Y view). Each plane is separated by a distance of 3 mm, as are the wires within each plane, and to ensure transparency the planes will be held at biases of -200 V (first induction), 0 V (second induction) and 500 V (collection) respectively. A total of 11,264 150 μm copper-beryllium wires will be present across the

four anode plane assemblies. The geometric spacing and orientation of the anode plane wires is deliberately consistent with MicroBooNE and ICARUS detectors ensuring that the three detectors' performances are as similar as possible. For example, the wire spacing (or pitch) directly impacts the energy threshold for charged particle detection. Another consideration is that, given three wire planes, the choice of $0, \pm 60°$ gives the most uniform reconstruction performance across θ_{yz}.

Technically each anode is actually made up of a pair of anode plane assemblies. This helps to keep the frame size more manageable and ensure consistent wire tension across the planes. The wire tension is critical during the cooling process in order to avoid wire sagging and maintain the flatness of the wire planes [214]. The induction planes for the two APAs on either side are then connected via jumper cables, whilst biased electrodes ensure that electrons heading for the 'dead region' are diverted to the nearest active wires. This effect can be corrected for in the reconstruction of hit positions via knowledge of the biases used.

The current pulses are shaped, amplified and digitised by front end motherboards located at the top and side edges of the APAs. A 16-channel Application-Specific Integrated Circuit (ASIC) performs the shaping and amplification before passing the signals into a 2 MHz analogue to digital converter (ADC). The signals from 8 ASICs (128 channels) are then multi-plexed before leaving the cryostat. Having this waveform processing take place in the cold LAr ensures a better signal to noise ratio than can be achieved if the signals remain analogue into the warm. This comes from both a reduction of thermal noise as well as a shorter cable length resulting in lower total capacitance [180, 215]. Hardware and software filters will then be employed to further remove the remaining noise sources which are predominantly due to effects within the readout wires themselves.

Like the anodes, the cathode is formed of two panels aligned end-to-end. Each cathode plane assembly (CPA) is constructed from a set of sub-panels each consisting of a pair of mesh screens holding a reflective foil between them. The foil is coated with tetra-phenyl-butadiene (TPB) and forms a passive part of the photon detection system described in section 4.4.2. The primary utility of the cathode is to provide the high voltage surface

required to establish the drift field. The cathode will be operated at -100 kV, ensuring an electric field of 500 V/cm across the 2 m drift. A series of copper coated field cage strips surround the TPC volume parallel to the anode and cathode planes and are biased in stepped increments of 3 kV in order to maintain a uniform field across the whole drift length.

4.4.2 PDS

The PDS is designed to measure the component of the deposited energy released as light. It was noted earlier that the scintillation light produced in argon has a wavelength of 128 nm (with a width of $\sim$ 10 nm) sitting in the VUV spectrum. Even the most optimised VUV photodetectors have a reasonably low quantum efficiency of around 15% [216] and so it has been typical to coat photodetectors with a wavelength shifting material that absorbs and re-emits the light in the visible spectrum (VIS) around the peak detection efficiency region.

SBND's photon detection system (PDS) will comprise a total of 120 8" Hamamatsu photomultiplier tubes (PMTs) and 192 novel X-ARAPUCA light trap devices using silicon photomultipliers (SiPMs) [217, 218]. The X-ARAPUCA devices are formed of a reflective box with use of a dichroic filter. This first converts the photons to a wavelength which the filter is transparent to and then, once inside, to a wavelength which the filter is reflective to. This combination traps the photon within the box until it is detected by the SiPM. 96 of the PMTs will be coated with tetra-phenyl-butadiene (TPB), a wavelength shifter with an emitted wavelength of 430 nm, and the other 24 left uncoated. Similarly, half of the X-ARAPUCAs are designed to be sensitive only to VUV whilst the other half are only sensitive to visible light. The photodetectors are arranged in 24 PDS boxes situated behind the anode wire planes, figure 4.9 shows how the 5 PMTs and 8 X-ARAPUCAs in each box are arranged.

The final part of the SBND PDS is a passive element. The cathode panels are reflective and coated in the same wavelength shifting TPB as the PMTs. This provides SBND with

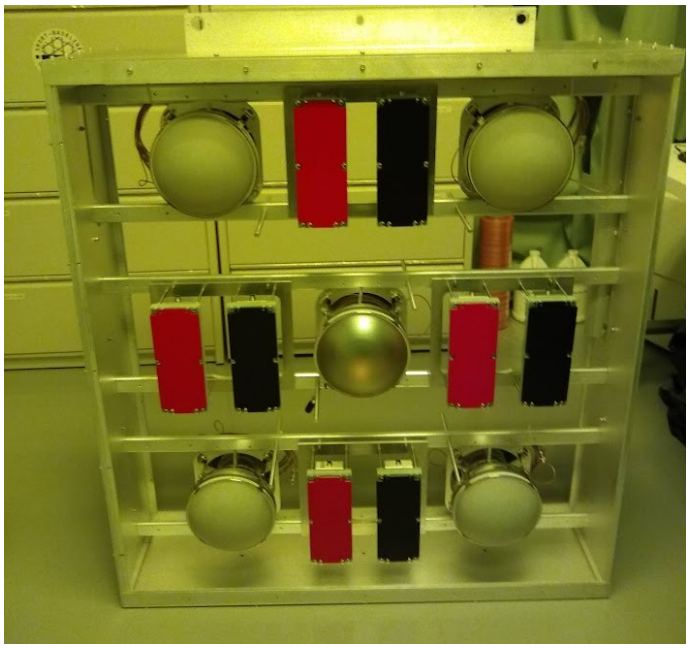


Figure 4.9: An example of one SBND PDS box containing 5 photomultiplier tubes and 8 X-ARAPUCA light trap devices. Note the visual distinction of the 4 TPB coated PMTs compared to the uncoated central PMT as well as the black and red covers indicating VUV and VIS sensitive X-ARAPUCAs respectively. Picture from [219].

a number of benefits. Firstly, by making the cathode reflective, the light yield for particles near the cathode is substantially increased. Secondly, the Rayleigh scattering length for the wavelength shifted light is longer, further increasing the yield of this component. Finally, by making use of the fact that some of the photodetectors are sensitive to the unshifted VUV light and some to the shifted VIS light, a determination of the drift position can be made with the PDS alone, something previous LArTPC experiments have not been able to do. This will be done by measuring the ratio of the two light components and extracting the drift position via a calibration curve developed using highly vertical cosmic muons selected with the CRT system.

One of the key features of the PDS is the nanosecond resolution it provides. As was mentioned earlier LArTPCs are 'slow' detectors. It will take SBND $\sim$1.3 ms to read out one full drift length (2 m). A few consequences of this are poor time resolution from the TPC, ambiguity about the x location of elements of the reconstructed charge readout, and the passage of a number of cosmic ray muons through the detector during the readout. In comparison, the scintillation light takes nanoseconds to reach the PDS. By matching the shape of a flash pattern to a distinct area of charge in the TPC readout, a t_0 (or interaction time) can be assigned to that area of charge. This helps to locate the charge in the x direction and exclude activity known to occur outside of the beam window. For this reason the PDS (specifically the PMTs) will also be used to trigger the detector readout when a flash of light is seen during a small window in which the neutrino beam is known to be passing through the building.

4.4.3 CRT

SBND's second auxiliary detector elements are the Cosmic Ray Taggers (CRTs). The CRT comprises 7 walls of taggers around the outside of the cryostat. As can be seen in figure 4.10 there is one wall on each side with a pair of walls on the top of the cryostat to form a 'telescope' arrangement. Each wall comprises of two layers of plastic scintillator arranged at 90° to one another, each layer consisting of a number of modules. Each module, designed and fabricated by the Bern group, comprises 16 optically isolated strips of 1 cm scintillator

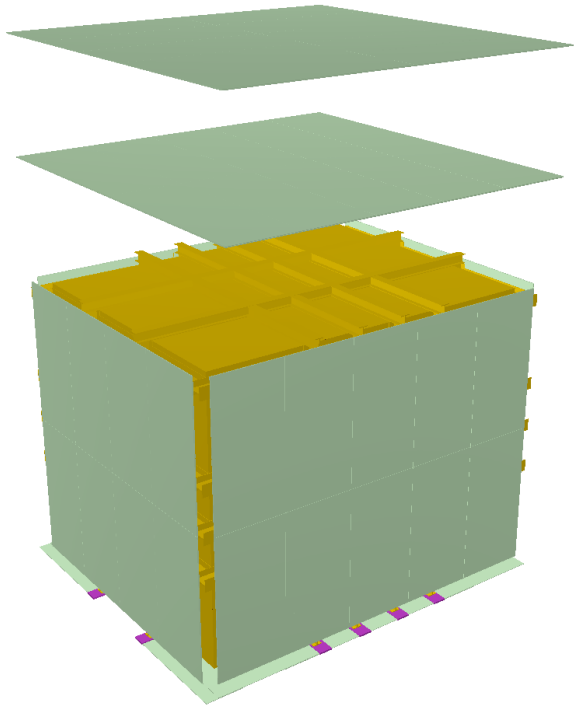


Figure 4.10: The location of the 7 SBND CRT walls around the cryostat.

and a pair of wavelength shifting fibres. The fibres run the length of the sides of the strip and are attached to a readout SiPM at one end and coated with reflective aluminium at the other. Other than the modules used in the CRTþ (under the cryostat) and those used to 'patch' the gap in north wall where the protego valve enters the cryostat, each strip is ∼11.2 cm wide and either 3.6 m or 4.5 m long depending on their location.

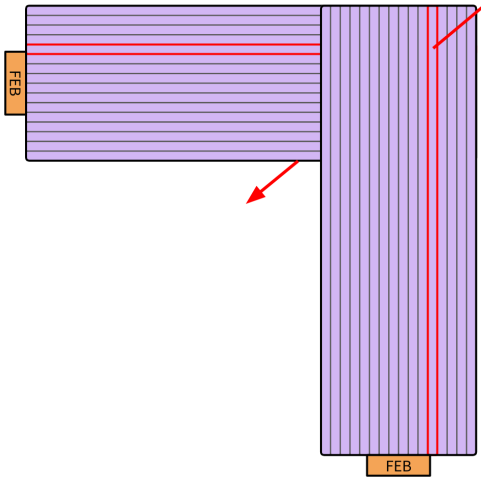
The perpendicular arrangement of the wall's layers provides the ability to reconstruct three-dimensional positions of energy depositions as well as reducing radiogenic backgrounds by requiring a four-fold SiPM coincidence. Figure 4.11 shows how the modules and the corresponding trigger cabling is arranged in order to enforce this. The chains of trigger cables are setup such that one loop provides the trigger *output* for one orientation and the trigger *input* for the second orientation. The other loop provides the inverse combination. The readout board is then configured to only produce a readout for that module if it recieves a signal originating from one of the perpendicular modules within a 160ns coincidence window. Like the PDS, the CRT system provides nanosecond resolution and the primary

goal of the system is to reject cosmic activity in the TPC by matching it to hits or tracks in the CRTs. By showing that the activity occurred outside the beam window, or that it passed fully through the detector, it provides clear evidence it is not a result of a neutrino interaction in the argon.

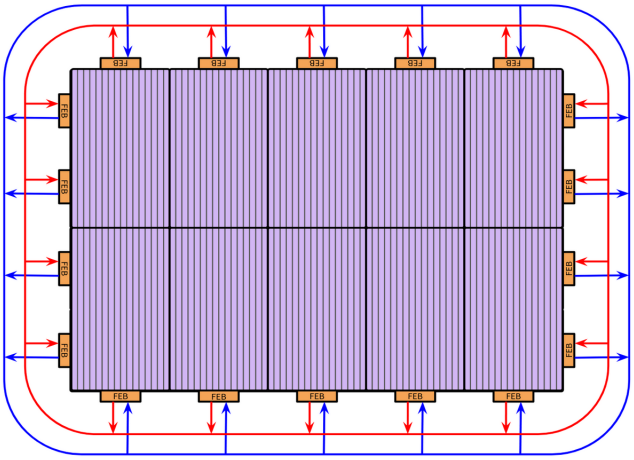
Whilst based at Fermilab I installed the front-end readout electronics, trigger, timing and data cabling and began the commissioning of the bottom wall of CRT modules. The bottom CRT, or CRT♭, has some key differences from the other walls, primarily due to its location under the cryostat. The modules are arranged to fill the gaps between the cryostat supports, and in many areas they are therefore a single layer, without the benefits of the perpendicular overlap, or no coverage at all. The CRT♭ also comprises a combination of Bern modules (smaller than those installed on the other walls) and a smaller number of modules repurposed from the MINOS experiment. Figure 4.12 shows the arrangement of the modules and cabling loops (in August 2023) and myself performing installation work. Further details are given in [220].

Before the installation of the SBND TPC in its cryostat, we operated a test stand for the CRT system at the SBN-ND building, referred to as the CRT## project (pronounced CRT sharps). Two temporary frames were attached to the upstream and downstream walls of the cryostat. Each consisted of two vertical and two horizontal small Bern CRT modules arranged like a # symbol, which were operated through the full SBND Data Acquisition (DAQ) system via the servers on the building mezzanine. The project aimed to inform CRT commissioning plans, develop the DAQ system, demonstrate DAQ synchronisation across multiple sub-systems and prepare a CRT crossing muon trigger. One of the key results using CRT## was the demonstration of the BNB bucket substructure using the CRTs independently and then in conjunction with the SPEC TDC (Simple PCIe FMC Carrier Time to Digital Converter) [221].

Figure 4.13 shows the results of this demonstration. Both plots show the time of three-dimensional CRT hits constructed from signals in one vertical and one horizontal strip. Only hits consistent with the BNB beam arrival at the building are plotted, a modulus of 18.94 ns is then applied, which represents the separation between the buckets within the

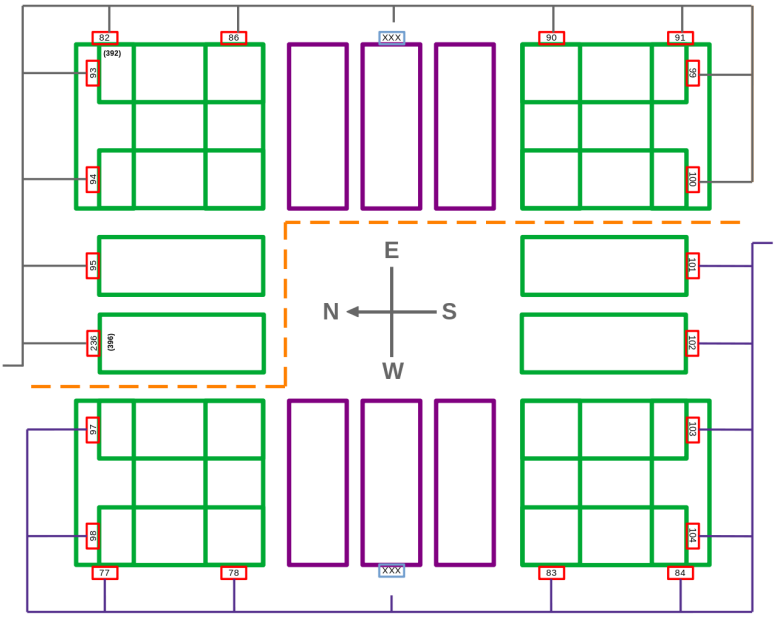


(a) The principle of detection in the CRT walls. Each module consists of a number of strips and the perpendicularly overlapping layers alow for constraints in both directions within the plane of the wall.



(b) An example of the arrangement of CRT modules within a wall. The horizontal and vertical modules are connected via two chains in order to send and receive trigger signals. They are arranged to ensure that the modules only readout when a module in the perpendicular layer also sees a signal of interest.

Figure 4.11: Two schematics illustrating the principles behind the perpendicular arrangement of modules, and the corresponding trigger cabling, in the CRT walls.



(a) The arrangement of the CRTþ modules and cabling routes in August 2023. The green shows the 20 Bern modules and the purple the 6 MINOS modules. The red shows the front-end readout boards connected by daisy chained cables for data, timing and triggering, separated into east and west chains.

(b) Me installing cables in the CRTþ setup, February 2023.

Figure 4.12: CRTþ (the CRT modules underneath the SBND cryostat).

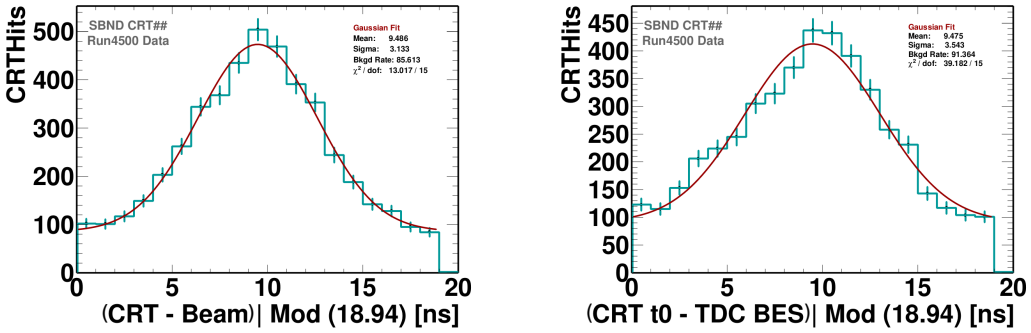


Figure 4.13: Demonstration of the BNB bucket substructure using the CRT## setup individually (left) and with the SPEC TDC system (right) with data taken in autumn 2022. Fits of a gaussian plus a constant (to represent the constant cosmic background) have been performed to each distribution.

beam. The use of a modulus is motivated by the relatively low statistics of the CRT## test stand runs, the full CRT system should be able to resolve the 81 buckets independently with just a few weeks of data taking. The clear peaked shape of the data within this region indicates immediately the ability of the CRT system to use its $\mathcal{O}(ns)$ timing resolution to identify individual buckets within the wider beam peak.

The difference between the two plots is the piece of timing information used to construct them. The plot on the left uses one of the two internal clocks (T_1) on the CRT front-end boards (FEBs). This clock was cabled to receive the beam arrival signals sent to SBN-ND $\sim 334\,\mu$s before the neutrinos pass through the building. The time value used was the number of ticks since the board last received such a signal. The plot on the right uses the FEB's other clock (T_0). This clock is instead reset by the pulse-per-second (PPS) signal delivered to all of the detector sub-systems by the White Rabbit timing system, in order to establish synchronised nanosecond precision [221]. The same beam arrival signal used for the left hand plot was delivered to the White Rabbit SPEC TDC, recording its time with respect to the PPS to picosecond precision. This value could then be subtracted, once cable delays were accounted for, from the CRT T_0 time to produce the right hand plot.

The fact that both plots were able to clearly show the beam substructure to similar

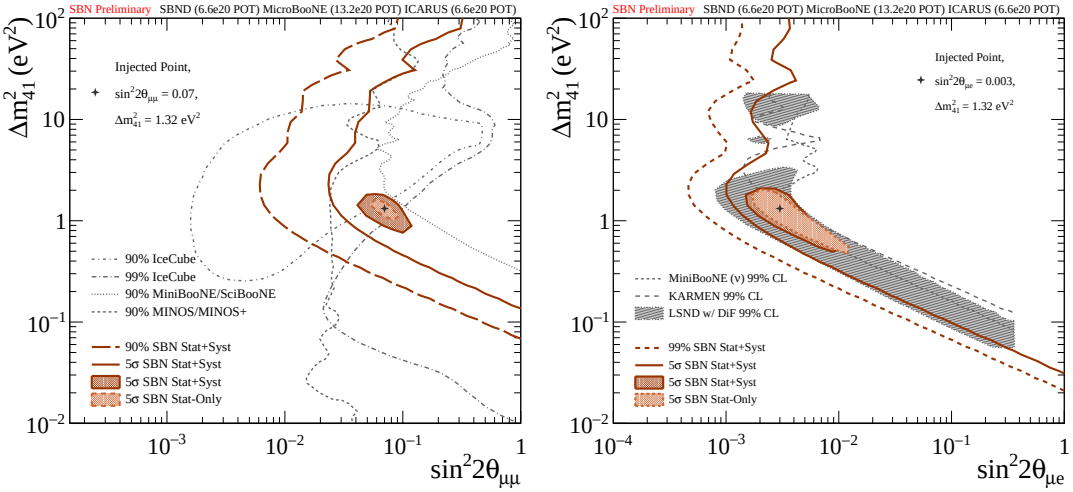
precision indicated the successful use of the White Rabbit SPEC TDC to reference multiple systems together with nanosecond precision, a use that will be particularly critical for 'off-beam' runs in which accelerator beam signals cannot be used. These plots do not represent the final precision of the CRT system. Further work is planned to correct for clock drift effects on the individual FEBs which will improve the resolution even further.

4.5 SBND Physics Program

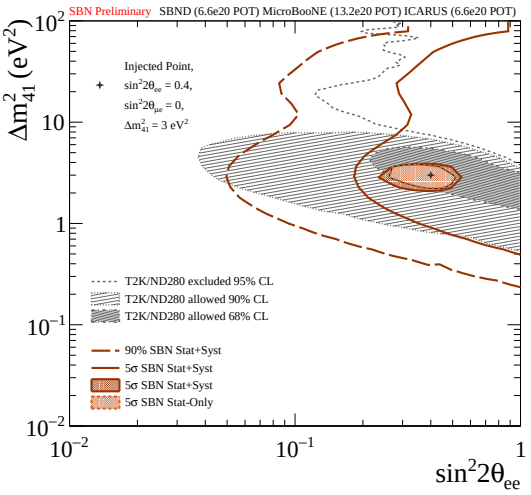
The SBND Physics Program consists of three main pillars: neutrino oscillations, neutrino-argon cross sections and beyond the standard model (BSM) searches [222, 91].

The primary goal of the whole SBN program is to conclusively address the LSND and MiniBooNE low energy excess and search for eV-scale sterile neutrinos. The locations of the SBN detectors are optimised to be maximally sensitive to the oscillatory effects of such a sterile neutrino, and will be able to simultaneously measure ν_e appearance, ν_e disappearance and ν_μ disappearance. Current world data shows significant tension between ν_e appearance and ν_μ disappearance, so the ability of SBN to constrain both in the same experiment will help to eliminate such tension. Figure 4.14 shows a recent analysis of the SBN program's sensitivity in the parameter spaces of each of the three possible channels. The MiniBooNE and LSND allowed regions are almost entirely covered by the SBN sensitivity to the 3+1 sterile neutrino scenario at 5σ, demonstrating SBN's ability to make a conclusive determination on the existence of such particles. SBND will play a crucial role in these analyses as the program's near detector; making measurements of the unoscillated flux to significantly reduce systematic uncertainties and probing the sterile neutrino parameter space at very short baselines.

SBND will also pursue a rich single-detector physics program. As a result of its size and position very close to the BNB source, it will be able to record an order of magnitude more neutrino-argon interactions than any previous experiment. This will amount to around 2 million CC ν_μ and 15,000 CC ν_e events for each of the 3 years of planned operations. The energy distribution and exclusive breakdowns of these events are shown in figure 4.15. This



(a) ν_μ disappearance (b) ν_e appearance



(c) ν_e disappearance

Figure 4.14: Recent SBN sensitivities to the three different oscillation channels in a standard 3+1 sterile neutrino scenario. Each plot shows relevant world limits, the SBN sensitivity contours and an injected point to demonstrate discovery potential. The SBN contours show coverage at high significance for the key region in each of the three channels - a combined measurement of these channels will allow SBN to investigate the tension between the allowed and excluded regions based on previous results.

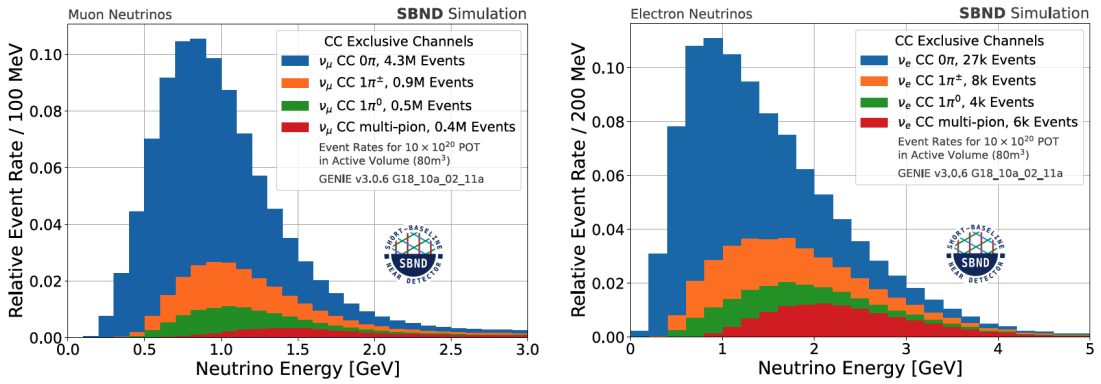


Figure 4.15: Expected event rates for CC ν_μ (left) and CC ν_e (right) interactions in the SBND active volume for an exposure of 1×10^{21} POT.

will allow for precision neutrino cross section measurements in a wide variety of exclusive final states.

SBND's proximity to the BNB target also gives it increased sensitivity to a wide variety of proposed BSM scenarios, most of which produce a signature such as an electron-positron pair which could explain the electron-like excess reported by MiniBooNE. These include dark neutrinos [223, 224], heavy neutral leptons [225] and Higgs portal scalars [226]. Another benefit of the location so close to the target is that, unlike most neutrino beam experiments, SBND will be able to sample a non-negligible range of off-axis angles. This concept, referred to as PRISM, following nuPRISM [227], allows SBND to view different neutrino energy distributions as visualised in figure 4.16 and can be utilised in all three key aspects of the SBND physics program.

4.6 SBND Current Status

The SBND construction and installation was finished on 1st December 2023 and at the time of writing the collaboration is commissioning the detector ready to take physics quality data later in 2024.

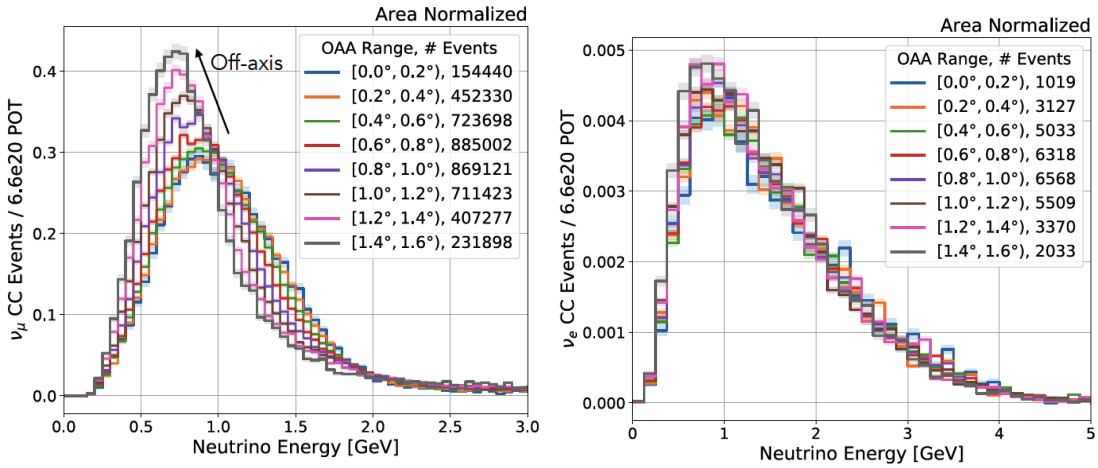


Figure 4.16: The energy distribution of CC ν_μ (left) and CC ν_e (right) interactions in the different SBND PRISM off-axis angle bins.

Figure 4.17: The SBND TPC being installed inside its cryostat at the SBN-ND building, Fermilab, 25th April 2023. Photo from [228].

Chapter 5

Simulating and Reconstructing Neutrino Events in SBND

This chapter describes the event processing chain for the Short-Baseline Near Detector, primarily the approach used to generate simulation samples and then the tools used to reconstruct both simulation and data events. This chapter describes a large body of work carried out by the liquid argon, neutrino and wider high energy physics community over the past decade or more. I personally contributed to the development of the current TPC vertex reconstruction described in section 5.2.1.3 including writing a new 'vertex refinement' algorithm for the Pandora pattern recognition software. I also wrote the CRT reconstruction detailed in section 5.2.3. The Pandora work was supported primarily by Andrew Blake and Dominic Brailsford as well as the wider Pandora development team. The CRT work benefited from the previous iteration of the reconstruction and tools written by Tom Brooks and was improved significantly by the input of Michelle Stancari, Dominic Brailsford and others in the SBND Reconstruction group.

5.1 Simulation

Ironically, given the fundamental nature of their targeted science, the world of high energy physics experiments is an incredibly complex one. Creating a prediction of what the data in a vast detector should look like, based upon a particular hypothesis, requires the use of a wide range of tools, each one simulating a step between the underlying physics event and the observed result. Such simulations are crucial to develop tools, design detector components, predict background contaminations and assess analysis performance. Figure 5.1 schematically visualises the rough steps the SBND simulation chain can be split into:

- Simulate the neutrino flux at the detector (red).
- Simulate the underlying interactions and final-state particles - a neutrino interacts, a cosmic-ray particle arrives, a BSM particle interacts or decays (orange).
- Simulate the resulting passage of particles through the argon depositing energy, decaying and reinteracting (green).
- Simulate the response of the detector to these energy depositions (purple).

These simulations make use of Monte Carlo methods which sample from probabilistic distributions to create, with large enough statistics, representative population outcomes. Event processing at SBND (both simulation and reconstruction) takes place within the LArSoft framework, a shared codebase allowing for collaborative effort and benefit across a number of LArTPC experiments, including ArgoNeuT, LArIAT, MicroBooNE, DUNE and ICARUS@FNAL [229].

5.1.1 Let's throw some neutrinos - BNB Flux Simulation

The first element of the SBND simulation chain is the BNB flux simulation. The SBN experiments make use of the detailed simulation developed by the MiniBooNE collaboration [181]. This utilises the Geant4 software [230] which provides a comprehensive suite of tools for modelling the passage of particles through matter. The geometry considers

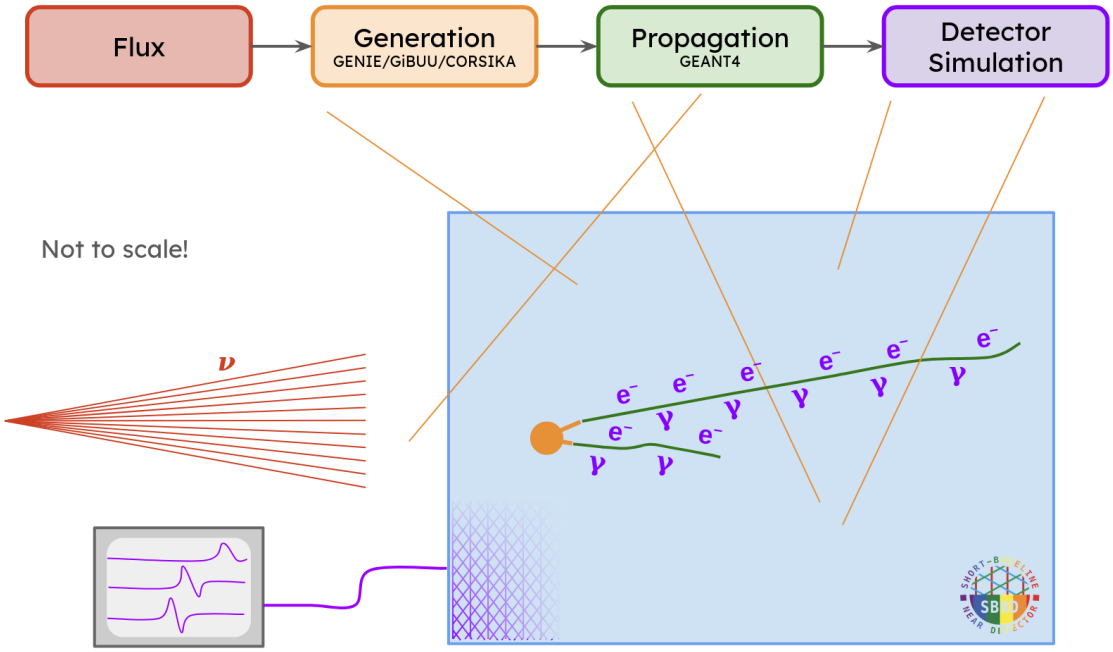


Figure 5.1: A schematic visualisation of the steps involved in the SBND simulation chain.

the final 50 m of the proton beamline, the target hall and the meson decay pipe length. The magnetic field is simulated within the horn cavity only.

The initial proton bunch is simulated as a group of totally uncorrelated particles 1 cm upstream of the target, with their positional and directional spread sampled from Gaussian distributions verified against beam monitoring devices. From this point onwards particle propagation is handled by Geant4, utilising their default cross section tables. The one exception is for the primary proton-beryllium interactions, where a dedicated model was put in place, tuned to existing data from the HARP [231] and BNL 910 [232] experiments. Once the propagation and decays of the resulting particles have been considered, the resulting neutrino flux at SBND is determined by using the kinematics of each resulting neutrino to project it to the plane in which the detector front face lies (for SBND this is 110 m from the target). The positional and kinematic distributions of these neutrinos are saved if they lie within a 10 m × 10 m box centred on the detector face. A wider 80 m × 80 m box is used when products from neutrino interactions in the 'dirt' upstream and

around the detector are to be considered.

5.1.2 Something Happened - GENIE GiBUU and CORSIKA

The next stage of the simulation is known as the generator stage. This covers simulating the primary physics occurring in the detector, either a neutrino interaction (GENIE or GiBUU) or cosmic-ray backgrounds (CORSIKA). This stage takes the input of the simulated flux, or in the case of CORSIKA a library of cosmic-ray showers at ground level, and uses Monte Carlo methods to produce each random event prediction. The key output is the list of particles produced and their initial momentum.

5.1.2.1 GENIE

GENIE (Generates Events for Neutrino Interaction Experiments) is an event generator designed primarily for use in the $\sim$100 MeV to $\sim$100 GeV range accessed by current accelerator based neutrino experiments. As described in detail in chapter 3 this is a particularly challenging region to model due to the transition from nucleon-level interactions to quark-level interactions. This challenge is heightened by a general lack of available data, as well as the fact that the experiments that have made measurements have often done so on different target nuclei (mainly carbon, hydrogen, oxygen and argon) and the scaling between these different targets introduces further uncertainty.

GENIE utilises a combination of theoretical and empirical models tuned to experimental data across a wide energy range to create a patchwork quilt covering the various types of interactions. Taking the flux provided by the BNB simulation, neutrino trajectories are extrapolated into the detector volume and the probability of an interaction is assessed en route using the total cross section. The type of interaction and the final state kinematics are provided by the differential cross section models, with final state interactions also accounted for. Some samples discussed in this thesis are simulated using the larger 'rockbox' geometry, in which a volume upstream of the detector hall is also considered, as well as a wider area around the detector itself. This allows the simulation to account for neutrino interactions

in the rock or surrounding materials which produce particles, particularly muons, that enter the detector and leave a visible signature.

The GENIE event generator is used to produce all the neutrino interaction Monte Carlo samples utilised in this thesis. In particular, we use GENIE v03.04.00 with a comprehensive model configuration (CMC) developed for ongoing DUNE and SBN studies known as AR23_20i_00_000. This configuration is very similar to G18_10a_02_11b [233] and consists of the following models:

- The initial nuclear state is modelled using a correlated Fermi gas. This is an extension of the local Fermi gas model in which a high momentum tail is added for nucleons above the Fermi momentum [234].
- The quasi-elastic cross section is predicted with the Valencia model [235], whilst the 2p2h contribution uses the SuSAv2 model [236].
- The resonant cross section is predicted with the Berger-Sehgal model [137].
- The deep inelastic scattering cross section is predicted with the Bodek-Yang model [145].
- The coherent pion production cross section is predicted with the Berger-Sehgal model [137].
- The hadronisation of DIS hadronic final states is predicted with the AGKY model [146].
- Final state interactions are predicted with an INTRANUKE hA model [115].

5.1.2.2 GiBUU

SBND has recently benefited from the implementation of a second, independent, neutrino interaction generator, GiBUU [237]. GiBUU is based on the Giessen Boltzmann-Uehling-Uhlenbeck model [118, 238] and is a purely theory-driven model which, unlike GENIE, is not tuned on experimental data. GiBUU has provided good agreement modelling of neutrino and electron nucleus interactions across the energy range of relevance to SBND.

Having the ability to make direct comparisons of multiple generators in SBND analyses will be critical to improving the modelling of interactions for future neutrino experiments.

5.1.2.3 CORSIKA

SBND is a surface level detector with no significant overburden and as such will receive a significant flux of cosmic-ray particles throughout data taking. The 'slow' nature of a TPC readout means that every neutrino event recorded will also contain around 5 cosmic-ray muons, whilst roughly 1 in 10 triggers will be caused by a cosmic-ray muon crossing the detector during the beam spill. Modelling the cosmic-ray particle flux is important in understanding the behaviour of such a large background but also because cosmic-ray particles will be a critical calibration tool in early running. This role is performed by the CORSIKA generator originally developed for the KASCADE experiment [239]. It provides a detailed simulation of the showers induced by high energy cosmic-ray particles and is able to consider various particles including protons, light nuclei and photons (or a mix of these types) as the primary initiators of the showers. SBND uses purely proton primaries due to better agreement with data taken by the MicroBooNE experiment. The key component for consideration by SBND are the muons produced in the initial interactions of these primaries in the upper atmosphere.

5.1.3 Where did it go? - Geant4

The next stage of the simulation transports the final-state particles through the detector medium producing a list of any further particles and depositions of energy that occur as a result. The LArSoft framework provides an interface to the Geant4 toolkit [230], to which the results of the neutrino interaction or the initial cosmic-ray particle are passed. Geant4 is an extensive tool for modelling the propagation of particles through matter. As well as high energy physics applications it is widely used across nuclear physics, radiation physics, space science and medical applications. Using a comprehensive model of the detector geometry and a set of 'physics lists' determining the modelling to be used, each particle is

stepped through the argon and surrounding materials. At each step appropriate energy is deposited and the probability of reinteraction or decay is assessed. Any products of such reinteractions or decays are then handled in the same way as the initial particles until all of the resulting particles have stopped, been absorbed or left the simulated volume (a large area around the detector and building).

5.1.4 What did we do with it? - Detector Simulation

The energy deposits produced by the Geant4 simulation are then converted into a simulated detector response - raw TPC and PDS waveforms, and CRT hit-based readouts. The first stage converts these energy deposits in the argon into charge and light, accounting for the recombination model and drift field distortions. The charge drift, under the influence of the electric field, and subsequent induction of current in the sense wires is handled by the WireCell software [240, 241]. The drift stage accounts for the attenuation of electrons due to electronegative contaminants, diffusion of the electron cloud both transverse and longitudinally with respect to the drift direction, and non-parallel transport due to field distortions primarily resulting from space charge effects. Next a 2D (averaged over wire direction) Garfield simulation is used to simulate the field response functions in the vicinity of the wire planes. This handles the electron motion from a response point 10 cm in front of the first induction plane through to their termination at the collection plane. These drift paths are independent of the individual charge cloud and are calculated in advance, accounting for the effects on 10 wires either side of the nearest wire. The resulting current on each wire, calculated via Ramo's theorem, is then subject to a simulation of the electronics response: pre-amplification, RC filtering and digitisation [240, 241].

A similar process is undertaken for the scintillation light produced at each point of energy deposition. Unlike the charge depositions, the photons propagate isotropically, unaffected by the drift field. Their path however, is affected by absorption, reflection and scattering. Individually simulating each photon's propagation via software such as Geant4 is excessively computationally demanding for each simulated event. A number of methods have been developed in order to combat this, typically these methods run the full Geant4

simulation to respond to typical energy deposits and store that information in a library which can then be utilised during standard simulation. SBND uses a semi-analytical model in which the voxelised library response is used to simulate the number of photons arriving at each photodetector alongside a parametrised arrival time distribution in order to avoid smearing out the non-negligible effects described in section 4.3.2 that serve to delay photon arrival [242].

Outside of the cryostat, the Geant4 energy depositions within the cosmic-ray tagger's scintillating strips are also considered. The energy is converted to scintillation light which is then divided between the fibres on either side of the strip according to the lateral position of the deposit within the strip. Next, the signal attenuation and time delay are accounted for using the longitudinal distance down the strip. Finally, the electronics response at the front-end board is emulated by assessing whether, and when, any pairs of SiPMs pass the threshold requirement, resulting in an internal trigger to readout a single ADC value from all 32 SiPMs at that time. The resulting deadtime for that front-end board is then enforced.

By the end of the detector simulation, simulated events have been produced in the same format as raw data coming from the real detector, from this point onwards, Monte Carlo and data can, and should, be treated in the same way.

5.2 Reconstruction

The process of interpreting the LArTPC images and auxiliary PDS and CRT data, and recovering the original final-state particles and physics quantities, is known as event reconstruction. Figure 5.2 illustrates the outline of the reconstruction chain, there are some common themes in the reconstruction pathways for all three sub-systems despite their differences.

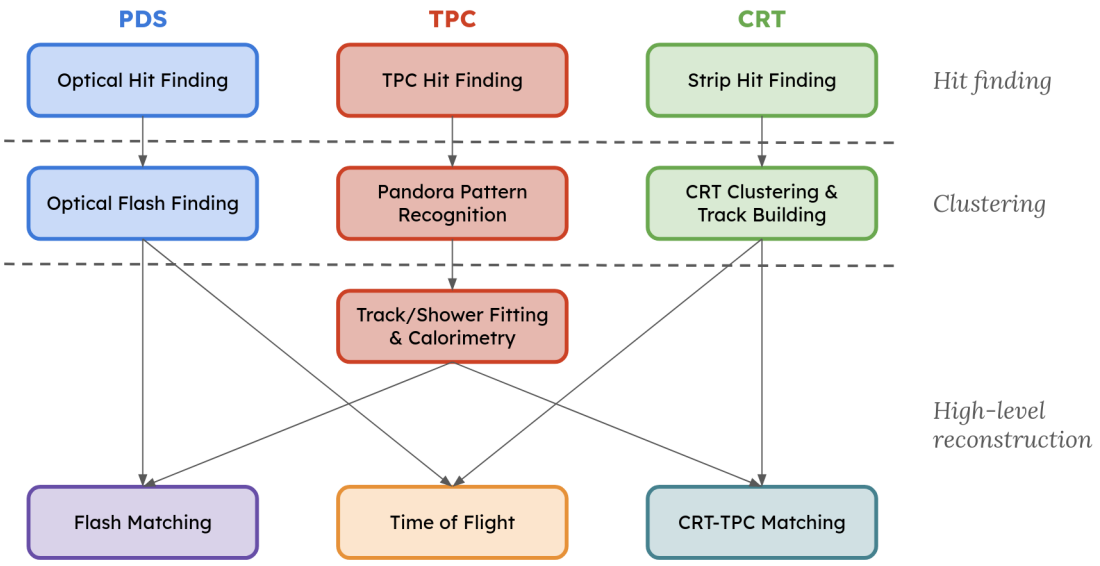


Figure 5.2: An outline of the reconstruction pathways for the three SBND subsystems, including combined reconstruction tools.

5.2.1 TPC

5.2.1.1 Signal Processing and Hit Finding

The first stage in the TPC reconstruction is the determination of the time-profile of energy deposits in each wire. This process is known as deconvolution and involves removing the effects of the field and electronics responses. The deconvolution is performed by transforming the raw waveform into the frequency domain via a Fourier transform and then applying a series of noise filters before performing an inverse Fourier transform back into the time space. SBND again utilises the WireCell toolkit to perform the deconvolution and signal processing [240, 241], allowing the deconvolution to be applied in a two-dimensional space of wires and time. In particular, the signal processing considers the impact of the 10 wires before and after the wire of interest. This improves upon the simpler one dimensional approach, where the signals are deconvolved in the time space only. Figure 5.3 shows an example of the signal processing chain from raw waveforms through to the optimised deconvolved and filtered output, which looks much more like the true ionisation pattern.

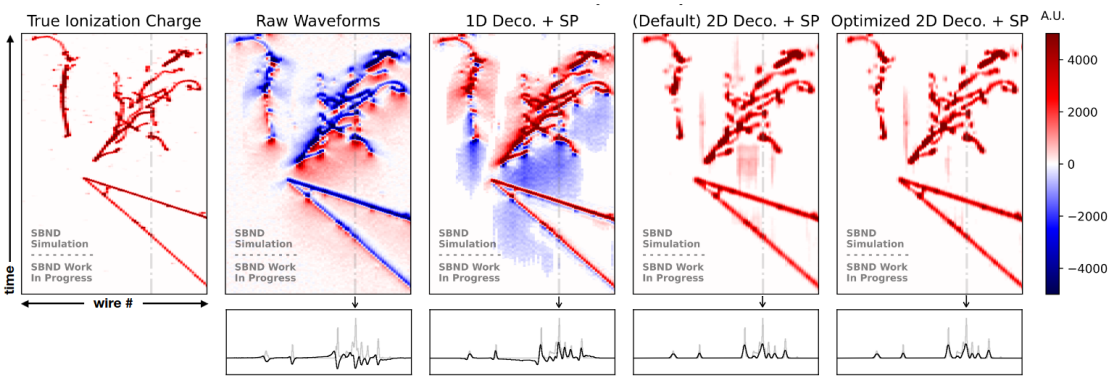


Figure 5.3: True, raw, and deconvolved U-view waveforms for a simulated neutrino event in SBND. The last three panels show three different implementations of the signal processing (SP). One-dimensional, two-dimensional and then two-dimensional with opitimised parameters. A single wire (same wire accross all panels) has been chosen to illustrate how the waveform looks at each stage (black) relative to to the true charge (grey). Figure from [243].

The one dimensional deconvolution demonstrates unphysical negative regions not present in the two-dimensional versions. The two-dimensional approach is particularly powerful for tracks parallel to a single wire, where the bipolar shape of the induction plane waveforms can cause significant destructive interference [240, 243].

Due to the slow nature of the TPC readout, the waveforms are digitised at 2 MHz for ∼1.5 ms, yielding ∼ 3000 bins for each waveform. Given the number of readout channels this would result in unwieldy raw file sizes. Region-of-interest (ROI) finding is performed to reduce this size by zero-suppressing the waveforms outside of these identified areas. These truncated waveforms are then passed to the hit-finding algorithm which identifies when a threshold level is exceeded and performs a Gaussian fit to identify the location and size of this 'hit'. In the case of more complex regions involving overlapping energy depositions, the waveform differential, erosion (rolling local minimum) and dilation (rolling local maxima) are used to determine multiple candidate hit positions. A fall-back option in the case of poor fit agreement is to enforce N evenly-spaced fixed-width Gaussians - incrementing N until the χ^2 agreement of the summed fits with the original waveform passes below a threshold [244, 245]. This is particularly necessary in high activity regions

such as around the neutrino vertex or for 'long' hits that are very perpendicular to the wire direction. Three examples of hit finding are shown in figure 5.4:

- Channel 956 contains a long deposit originating from a particle emitted from the vertex very parallel to the drift direction. The hit reconstruction has produced 4 Gaussians to cover the extent of the deposit.
- Channel 959 shows a region that looks at first glance like a single hit but can actually be separated into three hits, which is much more representative of the underlying true energy deposits. Splitting the hits in this way makes reconstructing the three particles that emanate from the vertex in a broadly forward direction much more accurate.
- Finally, channel 1001 shows a more 'traditional' situation away from the complexity of the vertex. This deposit from one of the tracks is well isolated from other activity and can be simply reconstructed with a single hit.

Each hit can therefore be characterised by its peak time, width and height. The integrated area under the pulse, calculated via the width and height, represents the scale of the energy deposition and a combination of the peak time and the wire on which it was recorded informs its location. In the plots shown in figure 5.4 the heights of the simulated energy deposits and the deconvolved waveform are matched for ease of display. For the purposes of energy reconstruction a calibration is performed to map the charge seen on the wire to the energy deposited in the detector.

5.2.1.2 Pandora Pattern Recognition

The reconstructed hits map out three two-dimensional pixelated images in time-wire space. The next step is pattern recognition, which aims to convert the sets of two-dimensional hits in each wire plane into a three-dimensional representation of the interaction vertex and final-state particle trajectories. In SBND, and in many other LArTPC experiments, this pattern recognition is performed using the Pandora software suite [246]. Pandora takes a

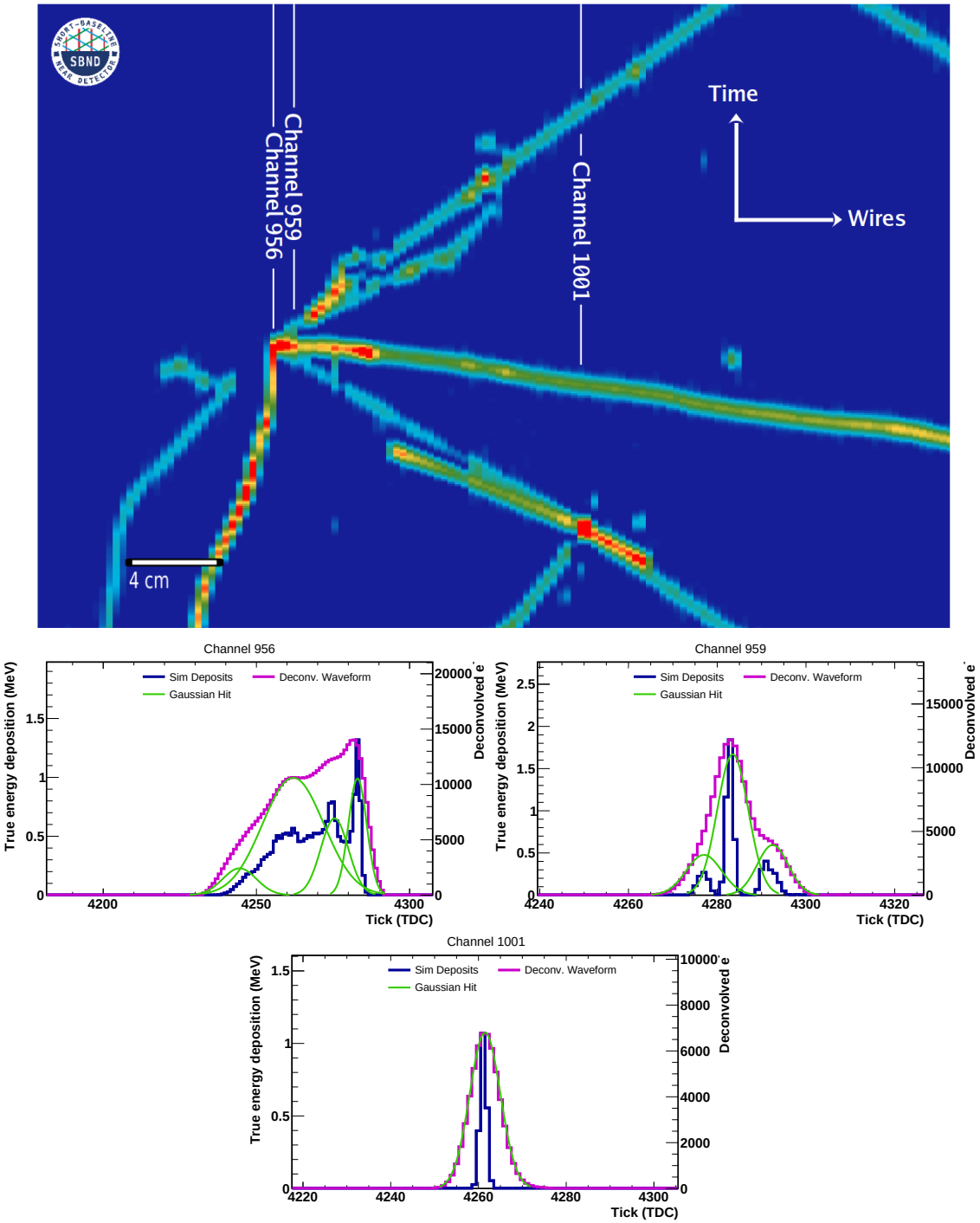


Figure 5.4: An example region of an event display from SBND Monte Carlo. Three channels are highlighted with example hit finding displayed in the plots below. Each shows the distribution of the true energy depositions in blue and the deconvolved waveform in magenta, with the Gaussian hits created from the deconvolved waveform overlayed in green.

multi-algorithm approach, where each of the over 100 algorithms available perform a very specific task and can be combined in a configuration best suited to each experiment and the requirements of a particular physics goal. The general philosophy employed by the Pandora approach is for each algorithm to be conservative in its decision making, only committing to changes that it is very confident in. In this way it aims to avoid making clustering mistakes which are then more difficult to unpick later on.

The workflow executed by the SBND Pandora configuration can be split into two passes, each containing a series of distinct stages. The first, known as the cosmic pass, is aimed at reconstructing activity from cosmic-ray particles and removing their associated hits. The second, known as the neutrino pass, targets reconstructing complex neutrino interaction final-state topologies. The cosmic pass will be referred to regularly as unambiguous removal, and the resulting removed tracks as unambiguous cosmics. It is tailored towards the assumption that all particles will leave long tracks with some small amount of electromagnetic activity surrounding them from associated delta rays and/or Michel electrons. Following this pass, each grouping, characterised by the main track, is assessed against a series of geometric metrics to determine its consistency with the cosmic-ray muon hypothesis. This includes looking at the containment, directionality and topology of the reconstructed objects, as cosmic-ray muons should primarily be downward going, enter through the top of the detector and be dominated by long straight tracks. If all the objects in an area of activity are rated as being cosmic-like then the associated hits are removed from consideration for the neutrino pass.

The second, neutrino-like pass contains more sophisticated algorithms, especially for identifying the neutrino vertex and reconstructing electromagnetic showers. Before the start of the second pass the remaining hits are broken up into distinct ‘slices’ in space and time. The ambition is that each slice is of independent origin, e.g. a cosmic-ray particle and any subsequent electromagnetic activity or the entire result of a neutrino interaction and so each slice is reconstructed completely independently from others in the event. This idea is demonstrated in figure 5.5 which shows the Pandora reconstruction of an SBND Monte Carlo event with each slice indicated in a different colour.

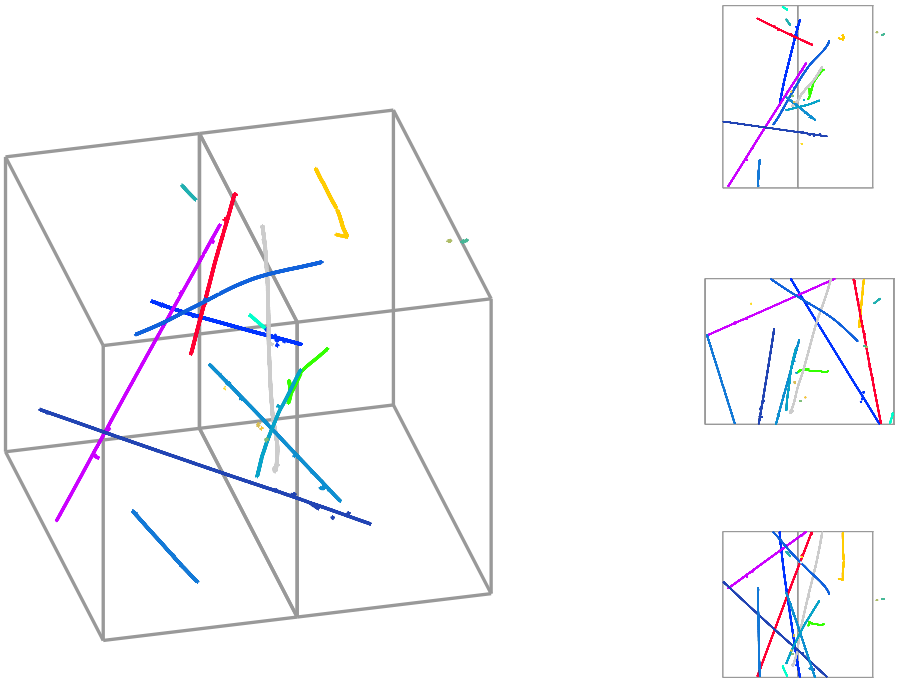


Figure 5.5: An example of a Pandora reconstructed event. Each colour represents a different slice with the two TPCs drawn in grey. The left-hand panel shows 2D projections from the top (xz), side (yz) and front (xy) in descending order.

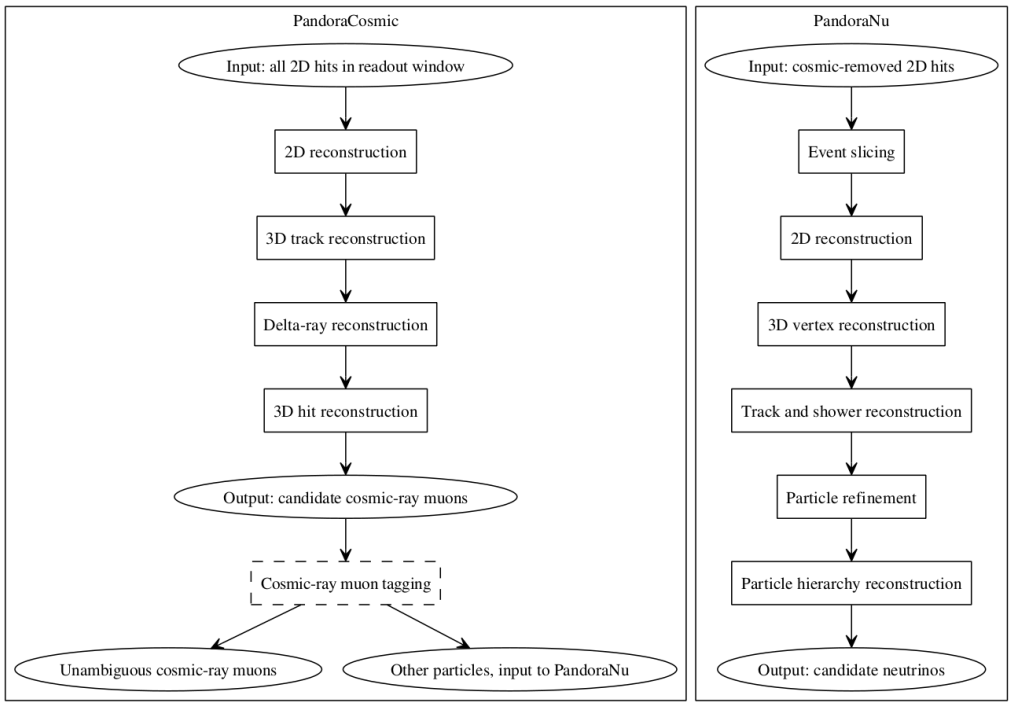


Figure 5.6: Outline of the Pandora reconstruction workflow. Figure from [247].

Both passes follow the same broad outline as represented in figure 5.6, starting with a series of two-dimensional clustering algorithms. The sets of hits from each readout plane are treated independently and identically. The initial algorithms cluster hits that form straight lines in the time-wire space. These small straight clusters are then merged by a series of subsequent algorithms that account for detector non-continuities (dead wires / designed detector gaps) and overlapping or intersecting particles. This stage also includes the ability to split clusters where clear kinks occur, or other clusters appear to originate from a point within it. Following the initial two-dimensional clustering, the second stage of the neutrino path then attempts to reconstruct the interaction vertex. Broadly this is achieved by identifying a series of potential candidate positions in each two-dimensional clustering output and then matching these into three-dimensional positions from which the most promising vertex is selected. More details on this procedure can be found in section 5.2.1.3. With the vertex now identified, Pandora is then able to use this to guide the further reconstruction, firstly by splitting any two-dimensional clusters bisected by the vertex point.

The two-dimensional clusters are now matched across the three views primarily via their shared drift coordinate. A combination of positions in any two views, for a shared x location, can be used to predict the equivalent location in the third, hence leaving a level of degeneracy that can be used to make quantitative assessments of the level of agreement between clusters in all three views. A complex set of algorithms assesses all possible permutations and employs a series of corrective tools to address a variety of ambiguities that can occur in this procedure. A single example of this would be the scenario in which two true particles create a pair of clusters in each of two views but appear back-to-back in the third view and have formed a single cluster across the entire shared drift range. In this scenario the 'overshooting' cluster will be split and the resulting clusters reconsidered in the assessment. This scenario is visualised in figure 5.7.

In the neutrino pass each of the resulting three-dimensional objects, known as particle flow objects (PFOs), are characterised as either track-like or shower-like. This characterisation considers the length, transverse extent (width), directional consistency, and proximity to the reconstructed vertex. All shower-like PFOs are at this stage dissolved and the hits are re-assessed by a series of clustering algorithms specifically targeted at the more diffuse, branching shower signatures. These algorithms build showers around a longer central spine, progressively adding branches that appear consistent with the growth of an electromagnetic shower. A similar iterative process to the three-dimensional track building is now carried out to create three-dimensional PFOs from the shower clusters, addressing a range of possible ambiguities. Any residual clusters that were not included in any track-like or shower-like PFOs are examined by a set of recovery algorithms with lower quality requirements to create a final set of PFOs. The shower-like PFOs are also passed through a set of mop-up algorithms which, as named, gather up any remaining small clusters in the fringes of the shower. This is first done in pure two dimensions and then later with the three-dimensional object projected into the view of each cluster.

Another assessment of the PFOs is now made to determine their characterisation as track or shower-like, this time by a more sophisticated algorithm utilising a boosted decision tree (BDT), a machine learning technique described further in the following section [248, 249, 250]. This BDT is trained to identify muon, proton and charged pion PFOs as tracks and

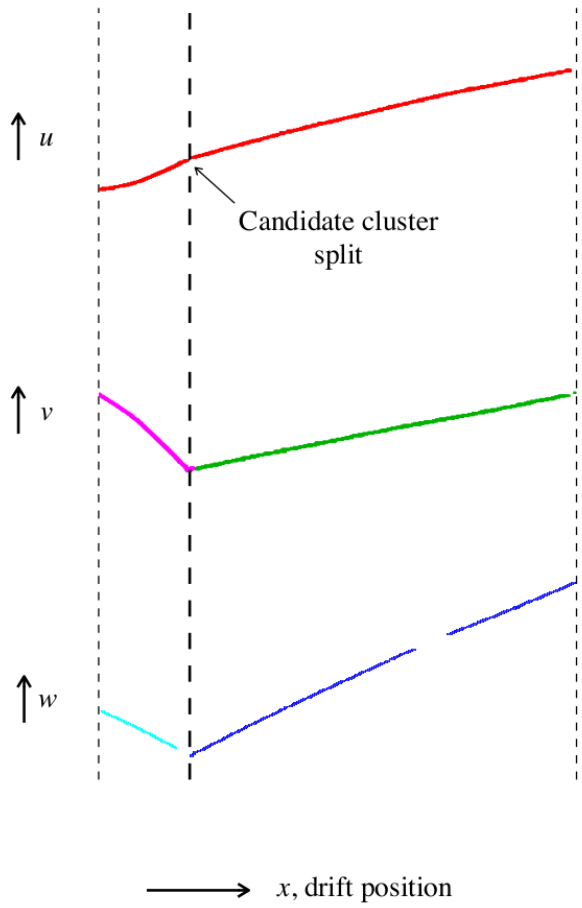


Figure 5.7: An example of an ambiguous topology presented during Pandora's two-dimensional to three-dimensional matching. The ambiguity results from the overshooting of the U view cluster. Figure from [247].

electron and photon PFOs as showers. The output score is known as Pandora's TrackScore. This characterisation is used to determine the method by which three-dimensional hits or 'space points' are created. A recent addition to the SBND Pandora configuration is to then run a recursive procedure in which a series of mop up algorithms aim to add any prevailing two-dimensional clusters to the PFOs. After each iteration the space point creation is re-run and the whole process repeats until no more changes are made [245]. The final stage assembles all the PFOs in a slice into a neutrino hierarchy. A PFO with no hits or clusters is created to represent the neutrino itself, any PFOs determined to be associated with the reconstructed interaction vertex are then attached as primary children of the neutrino, with further PFOs added as subsequent descendants (such as a Michel electron as the child of a primary muon track). Another BDT creates a score assessing how 'neutrino-like' the slice is to help inform analysis decisions. This is discussed in more detail in chapter 6.

The final output from Pandora consists of slices, each consisting of a hierarchy of PFOs. Each PFO is identified as track- or shower-like. The output also contains an interaction vertex, the vertices associated with each PFO and a list of three-dimensional space points indicating the PFO trajectories.

5.2.1.3 Improvements to Neutrino Vertex Reconstruction

The neutrino vertex reconstruction forms a key stage in the heart of the neutrino pass. Not only is the interaction point key to the deriving of many physics observables but its accuracy affects the rest of the pattern recognition which uses the vertex to guide decision making on further clustering, merging, splitting and hierarchy determination as well as downstream reconstruction such as calorimetry and particle identification.

The vertexing stage consists of two sections, candidate creation and vertex selection. All two-dimensional clusters that contain a minimum of 5 hits are considered in a pairwise fashion across two views at a time. If their endpoints are consistent in x then a candidate vertex is created, with a three-dimensional position, from the information provided by the two views. This process repeats for all combinations of two of the three views resulting in a list of candidate vertices. A two stage BDT is then used to determine the best candidate

from this selection. First, the candidates are grouped in regions of 10 cm with one candidate chosen to represent this region. The first BDT selects the most promising region, whilst the second then views all the candidates in the chosen region to determine which candidate will be selected as the best vertex. The following section details a pair of improvements made by the author to the vertex reconstruction - the addition of a new vertex refinement algorithm and the inclusion of a number of new variables to hone the performance of the BDT stage.

Vertex Refinement Algorithm

A new algorithm, referred to as vertex refinement, was developed to improve vertex positions based on the pointing information supplied by nearby clusters. The algorithm starts with the existing vertex position and projects it into each view. Any clusters that pass within 10 cm of the vertex are then collected and a principal component analysis (PCA) is run on each one. PCA is a technique for reducing a multi-dimensional dataset (in this case positions in two-dimensions) into a set of orthogonal direction vectors (and corresponding sizes) that encode the variation in the data along decreasingly variant axes [251]. The primary principal component is used in this scenario to find the direction of the cluster, $\vec{u_i} = (u_{x,i}, u_{y,i}, u_{z,i})$. Each cluster can then be represented by a simple line equation:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = a_i \begin{pmatrix} u_{x,i} \\ u_{y,i} \\ u_{z,i} \end{pmatrix} + \begin{pmatrix} x_{0,i} \\ y_{0,i} \\ z_{0,i} \end{pmatrix}, \tag{5.1}$$

where $(x_{0,i}, y_{0,i}, z_{0,i})$ is a particular point on the line, a_i is a running coefficient and (x, y, z) is the general point on the line. Each cluster is also assigned a weight, w_i defined as

$$w_i = \frac{1}{1 + r_i^2}, \tag{5.2}$$

where r_i is the magnitude of the distance from the original vertex position to the point of closest approach on cluster i. This is motivated by the fact that clusters near the interaction point are more likely to reflect the original direction of travel of each final-state particle. As you move further away, scattering or reinteraction effects become more significant and smear the resolution with which each cluster points to the interaction point.

Following the approach in [252, 253], the lines and weights from each cluster are used to construct a matrix equation

$$G \cdot m = d, \tag{5.3}$$

where

$$G = \begin{pmatrix} w_1 & 0 & 0 & -w_1 \cdot u_{x,1} & 0 & 0 & ... & 0 \\ 0 & w_1 & 0 & -w_1 \cdot u_{y,1} & 0 & 0 & ... & 0 \\ 0 & 0 & w_1 & -w_1 \cdot u_{z,1} & 0 & 0 & ... & 0 \\ w_2 & 0 & 0 & 0 & -w_2 \cdot u_{x,2} & 0 & ... & 0 \\ 0 & w_2 & 0 & 0 & -w_2 \cdot u_{y,2} & 0 & ... & 0 \\ 0 & 0 & w_2 & 0 & -w_2 \cdot u_{z,2} & 0 & ... & 0 \\ & & & & ... & & & \\ w_n & 0 & 0 & 0 & 0 & 0 & ... & -w_n \cdot u_{x,n} \\ 0 & w_n & 0 & 0 & 0 & 0 & ... & -w_n \cdot u_{y,n} \\ 0 & 0 & w_n & 0 & 0 & 0 & ... & -w_n \cdot u_{z,n} \end{pmatrix}, \tag{5.4}$$

$$d = \begin{pmatrix} w_1 \cdot x_{0,1} \\ w_1 \cdot y_{0,1} \\ w_1 \cdot z_{0,1} \\ w_2 \cdot x_{0,2} \\ w_2 \cdot y_{0,2} \\ w_2 \cdot z_{0,2} \\ ... \\ w_n \cdot x_{0,n} \\ w_n \cdot y_{0,n} \\ w_n \cdot z_{0,n} \end{pmatrix}, \qquad m = \begin{pmatrix} x \\ y \\ z \\ a_1 \\ a_2 \\ ... \\ a_n \end{pmatrix}.$$

This corresponds to the group of simultaneous equations representing the lines associated with each cluster. This equation is solvable for m, the point of closest approach for the lines considered, via the equation

$$m = (G^T \cdot G)^{-1} \cdot G^T \cdot d \tag{5.5}$$

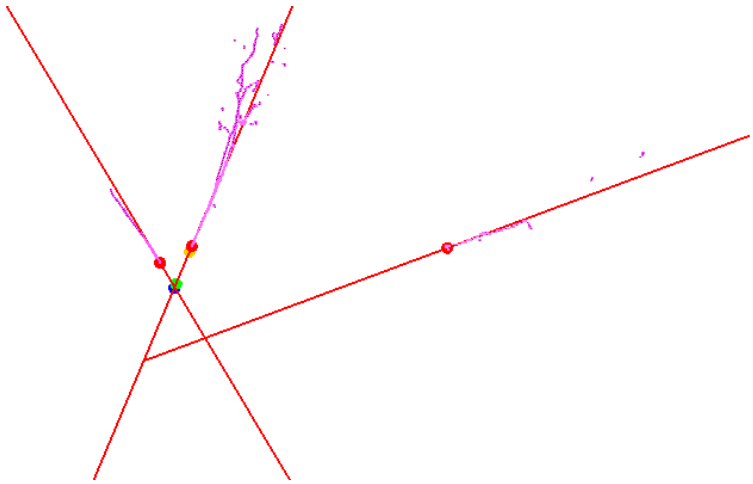


Figure 5.8: An example of the vertex refinement algorithm in action. This display shows its use in an example event display for a single view. The yellow dot represents the initial vertex candidate position, the red lines and dots represent the lines produced from the PCA of each surrounding cluster, the blue dot shows the refined vertex position and finally the green dot shows the true location of the neutrino interaction vertex projected into this view.

for the case where $G^T \cdot G$ is not singular. A more detailed description and an example for $n = 2$ can be found in appendix A. For $n > 2$ non-parallel lines in 2 dimensions or $n > 1$ non-parallel lines in three dimensions there will not generally be a single point of intersection, so this technique yields a least squares minimisation, finding the point at which the sum of the squared distances of closest approach to each weighted line is smallest. The algorithm was designed to be applicable to two-dimensional or three-dimensional scenarios. The use case detailed here for SBND is two-dimensional only, applied once per vertex in each view. In this scenario the y-components are all zero, the x-components are the shared drift coordinate and the z-components represent the wire coordinate for that view (either u, v or w).

Figure 5.8 illustrates the vertex refinement for an example event in a single view, demonstrating an improvement in the resulting vertex position relative to the true interaction point.

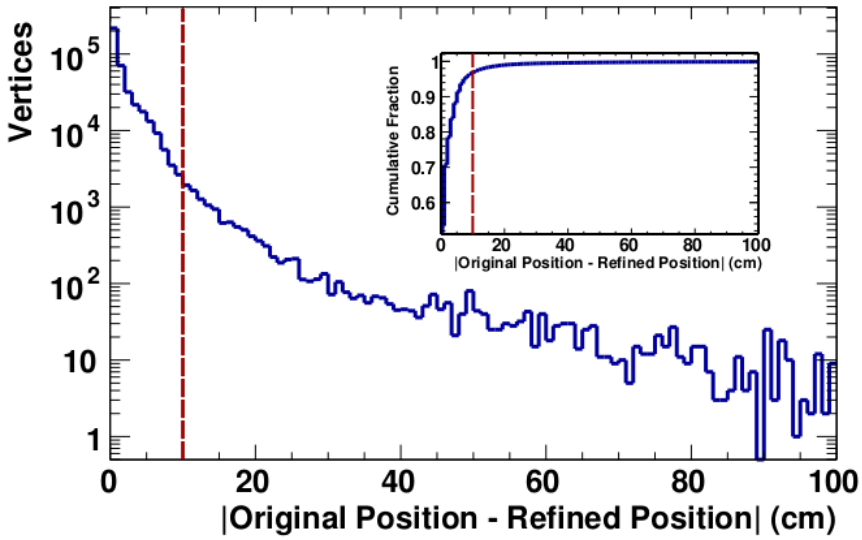


Figure 5.9: The distance between the original and refined position following the application of the vertex refinement algorithm to a vertex candidate in a single wire plane view. The cut value of 10 cm is indicated via the red dashed line. The inset plot shows the cumulative distribution, indicating that over 96 % of instances pass the cut.

Once this process is complete for all three planes the three 'refined' two-dimensional vertices are combined using one of a series of Pandora functions that use combinations of 2 views or all 3 views to form a three-dimensional vertex. The choice of function is determined based on which one returns the result with the best χ^2 agreement when projecting the result back into each view. Finally, the deviation of the new vertex position is checked to be within 10 cm of the original vertex position in order to avoid introducing new failure modes. As figure 5.9 shows, over 96 % of scenarios with at least 2 clusters do not move the vertex any further anyway. As reflected in the algorithm's name, its goal is to take good vertex reconstruction and make it very good, not to fix more fundamental failure modes. This algorithm is applied to all candidate vertices before passing its output to the selection BDTs. The performance improvements as a result of adopting this algorithm are shown in figure 5.10. This also shows the added performance bonus when the weighting aspect of the algorithm is applied.

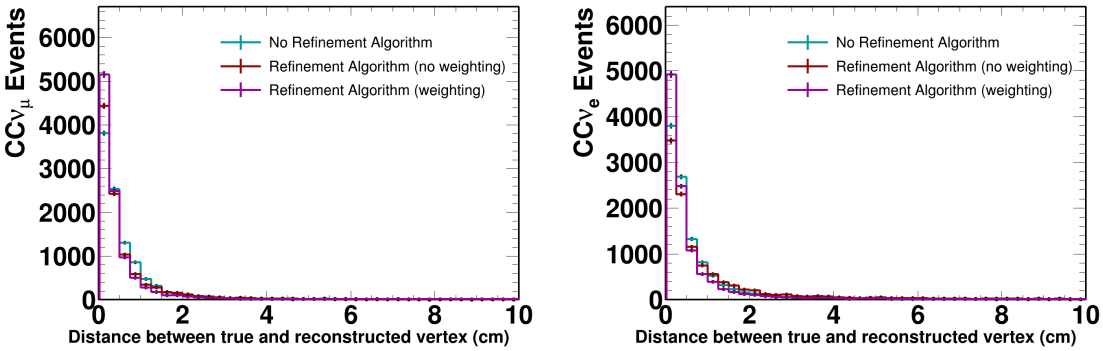


Figure 5.10: The improvement in vertex reconstruction for CC ν_μ (left) and CC ν_e (right) events as a result of the use of the vertex refinement algorithm only. The three curves show the performance without the algorithm (cyan), with the algorithm but without weighting the lines (red) and then the full algorithm with the weighting included (magenta).

Vertex Selection BDTs

A BDT is a machine learning technique based on the simple concept of a decision tree. In a single decision tree a condition is posed at each node, if passed the left branch is taken, if failed the right. All routes eventually end up at a leaf i.e. an outcome to the overall question. Figure 5.11 shows an incredibly simplified version of this idea, in the context of deciding what type of particle is represented by a reconstructed PFO. Each purple node asks a question about the PFO, the results of which lead to 5 leaves representing the 5 possible particle ID outcomes. This tree will clearly have some inefficiencies. What happens if the photon pair produces very quickly and appears to originate at the vertex? What happens if a very short proton track at the vertex gets merged with a long muon track making the start of the track appear to have a very high $\mathrm{d}E/\mathrm{d}x$? The leaves will not be pure in their desired outcome.

The decision tree can be 'trained' with a sample with known labels, i.e. a set of reconstructed Monte Carlo particles for which particle ID is known from truth. Maximising a figure of merit such as the purity of the leaves categorisations gives a metric for attaining the best possible tree. One approach to improve the results further involves using an

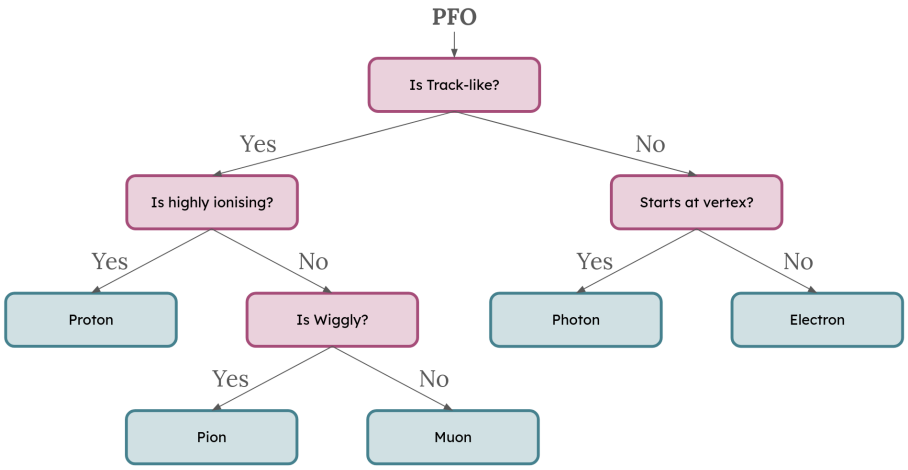


Figure 5.11: An overly simplistic example of a decision tree designed to identify particle types from a reconstructed particle object (PFO).

ensemble of different trees, each using different cuts, perhaps on different variables, to arrive at their classification. The final output can then be constructed by averaging the results of the entire ensemble. Unlike the example given here, these ensembles are almost exclusively designed to answer binary classification problems with a 'signal' and 'background'. As such, the averaged output is a score representing how 'signal-like' the input was. Boosting is a particular technique for training and combining trees that uses a high proportion of mis-classified events from the previous tree to train the subsequent tree. BDTs are used widely across high energy physics, as they allow a larger number of weak classifiers, which on their own are very impure discriminators, to be combined in a way that efficiently uses their complementarity to produce a much stronger discriminant (the output score) [248, 249, 250].

As outlined above, the final stage of vertex reconstruction is the selection of a single point from the list of (refined) candidates. A series of variables describing two of the candidates each provide some weak discrimination between them and the BDT produces a score which reflects which of the two is favoured. A winner-stays-on approach is then used where each vertex is compared to the current 'best' either replacing it or being discarded until all candidates have been compared and the resulting 'winner' is selected. The variables used in the original iteration of this tool were:

- Slice Level Variables - *These inform the BDT of which of the vertex characteristics could be important in the event rather than distinguish between the vertices themselves.*

    - Total number of vertex candidates in the slice.
    - Total number of two-dimensional clusters in the slice.
    - Total number of hits in the slice.
    - Total energy contained by the slice's hits.
    - The 'span' of the slice's clusters are calculated in each view and then the average spans in x and z (wire coordinate) are used to construct the area (product) and longitudinality (z span divided by the sum of the x and z spans) variables. This should indicate the rough shape of the slice's footprint, longer for CC ν_μ events and squarer for CC ν_e events.
    - The 'showeryness' of the slice. This is defined as the fraction of the hits in the slice attached to clusters deemed at this point of the reconstruction to be shower-like.

- Vertex Level Variables - *These variables are vertex specific and help determine which candidate is more likely to be the neutrino interaction point.*

    - The energy kick variable assesses the transverse energy of each cluster with respect to its distance from the candidate. This utilises the fact that primary particles should be both more energetic and aimed at the vertex.
    - The beam deweighting variable assesses how 'upstream' the candidate is within the slice, given the known knowledge of the beam direction. The true vertex position is likely to be one of the most upstream points available with the general momentum of the interaction boosted forwards in the beam direction.
    - A series of asymmetry variables constructed in a method visualised in figure 5.12. First, an energy weighted direction of the clusters is created (indicated by the black line), then all the hits are projected onto that axis (red arrows) and finally the energy is summed on each side of the vertex candidate

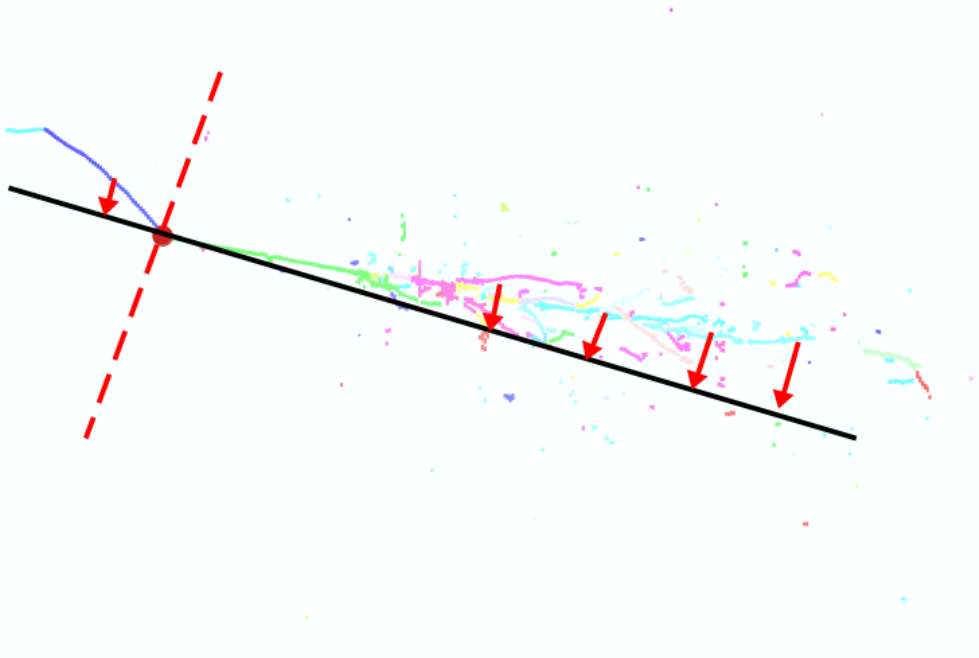


Figure 5.12: A schematic to show the projection technique used to create the asymmetry class of variables in the Pandora vertex selection BDT. The black line indicates the energy weighted average direction of the clusters, the red arrows indicate the projection of all hits onto that axis and finally the red dot indicates the vertex candidate. Each variety of the asymmetry variable sums a metric on either side of the vertex candidate.

(red dot). Three versions of this variable are created: global (all clusters), local (only clusters that have very similar directions) and shower (only clusters labelled as shower-like).

- A hit angle variable that finds the weighted sum of the ratios of each hit's radial angular position in circular coordinates in each plane. This variable intends to indicate whether the vertex position is consistent with a generally momentum balanced forward direction of a neutrino interaction. This variable is used for the second stage BDT only, once the region has been identified.

Four variables were added to the SBND configuration to improve the performance at both

stages of the selection:

- Vertex Level Variables:
  - The sum of the energy of the nearest hit to the candidate position in each of the three views. This is targeted at finding examples of high activity 'shared' hits at the vertex, with contributions from multiple primary neutrino children.
  - Another variant of the asymmetry variables described above. This time the energy depositions are normalised by the length they span on the projected axis. This is a proxy for the $\mathrm{d}E/\mathrm{d}x$ and aims to indicate examples where a candidate sits on the divide between back-to-back tracks of different ionisation strengths (e.g. primary muons and protons).
- Shared Variables - *A new group of variables that describe the relationship between the two candidates being compared.*
  - The separation between the two vertex candidates.
  - The number of hits on the axis between the two candidates normalised to their separation (as shown in figure 5.13). This is targeted at identifying whether the candidates being compared sit at either end of a clear track or not.

Distributions of the new variables are shown in figure 5.14. The vertex specific variables show strong discriminatory power between the correct and incorrect vertex. They are plotted for the instance of the region finding stage BDT in which all vertex specific variables show better performance. The shared variables are plotted for the second stage selection BDT instance. As expected they do not show any explicit discrimination between the two vertices, as a shared variable they cannot. It is worth revisiting the fact that this BDT operates by comparing a pair of vertices. The training data is labelled into two categories - one where the candidate vertex in the first slot is the closest candidate to the true position ("correct vertex") and one where the candidate vertex in the first slot is not the closest candidate to the true position ("incorrect vertex"). The shared variables are constructed using both candidates, hence there is no difference in distribution for the two

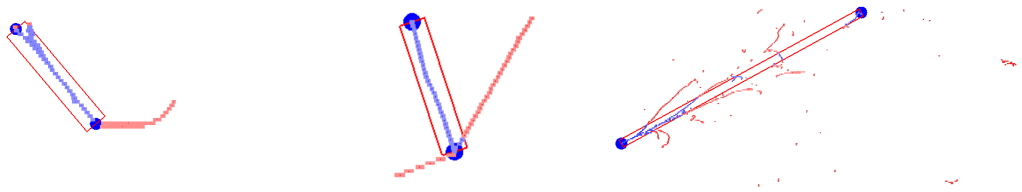


Figure 5.13: A trio of examples of the construction of the shared vertex variables for the Pandora vertex selection BDTs. The blue dots show the two vertex candidates being compared, and the red box the selection region around them. The separation variable is purely the distance between the two candidates, the length of the box. The axis hits variable is constructed by counting the hits that fall within the selection region (the blue hits) and then normalising by the separation.

labelled categories, the same is true for the event level variables. The role of these variables therefore is to inform the BDT of the broader context of the two vertices being considered, and as such influence which vertex variables are given priority in the determination. The output distribution for the second stage BDT is shown in figure 5.15 and demonstrates an uplift of 5 % in the validation accuracy with all four new variables included. The difference between the middle and bottom panes (an uplift of $\sim 2\%$) show the impact of the shared variables only.

Performance Improvements

As a result of the introduction of the vertex refinement algorithm, the added variables in the selection BDTs, and some adjustments to the hyperparameters used in the training of the BDTs, significant improvements were seen in the quality of the vertex reconstruction. Figure 5.16 shows the distance between the true and reconstructed vertex for both CC ν_μ and CC ν_e events before and after the improvements implemented. The proportion of CC ν_μ interactions for which the vertex is reconstructed within 0.5 cm of the true position increases by over 14 % from 54.6% to 68.9%, with over 80 % now reconstructed with 1 cm. The performance for the CC ν_e case start from a lower base due to the fact that this topology naturally provides more challenges to vertexing. For example, it is easy to misidentify the start of the primary electron depositions as a separate particle track before the first few

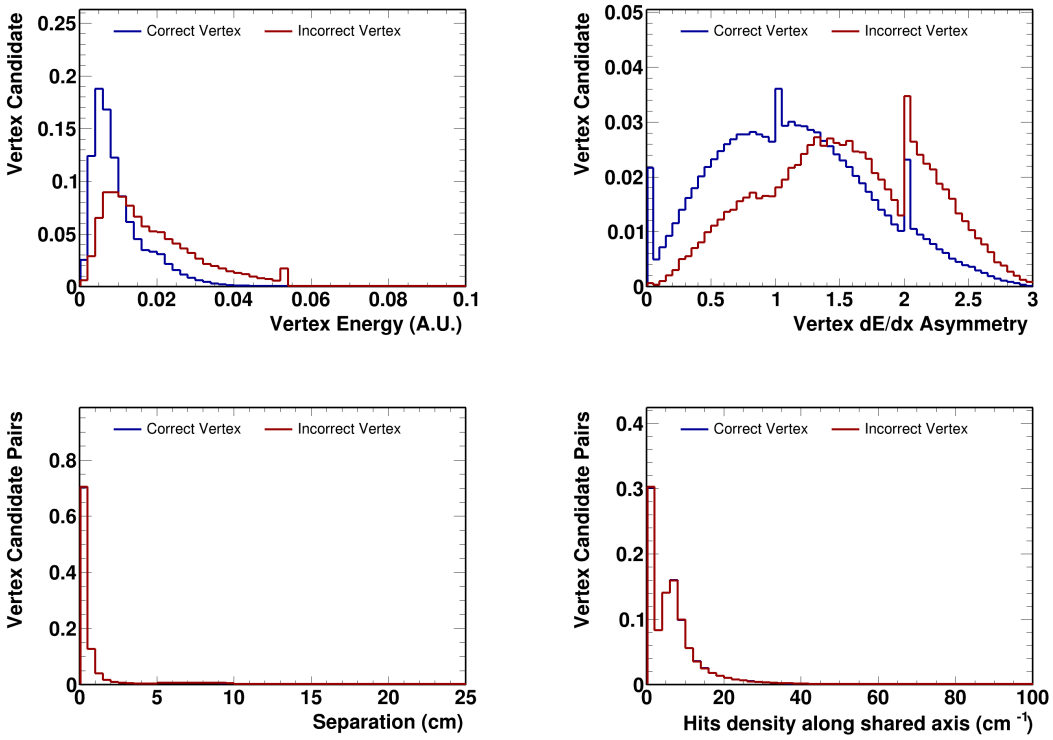


Figure 5.14: The four variables added to the SBND configuration of the Pandora vertex selection BDT. The two on the top row are plotted for the region finding BDT whilst the two on the bottom row are plotted for the final selection BDT. The BDT operates by comparing a pair of vertices. The training data is labelled into two categories - one where the candidate vertex in the first slot is the closest candidate to the true position ("correct vertex") and one where the candidate vertex in the first slot is not the closest candidate to the true position ("incorrect vertex"). The shared variables (second row) are constructed using both candidates, hence there is no difference in distribution for the two labelled categories.

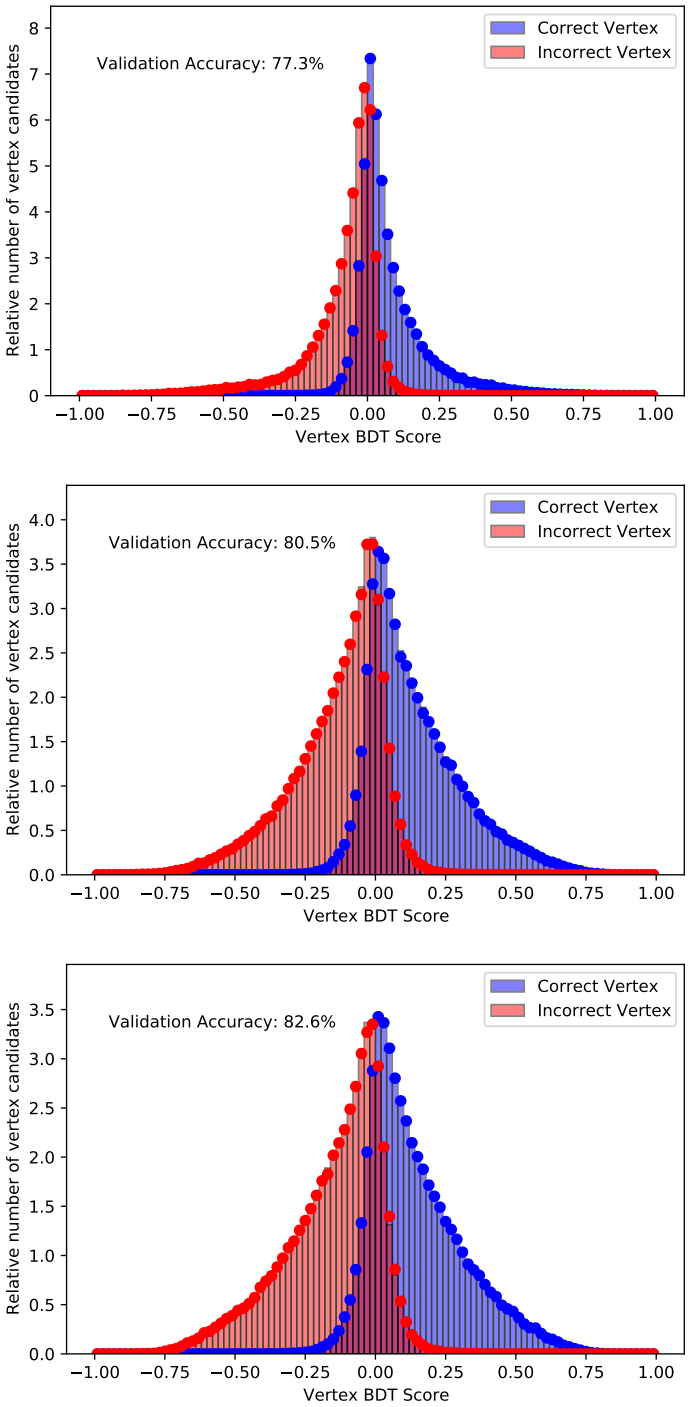


Figure 5.15: The second stage vertex selection BDT before any changes (top), following the addition of the two new vertex level variables (centre) and then finally following the addition of the two new shared variables as well (bottom). The validation accuracy refers to the proportion of correct inferences made for each pairwise comparison.

electromagnetic cascade processes create the distinctive shower shape. In this scenario the vertex could be misplaced at the point at which this transition occurs. However, similar scale improvements are shown for the CC ν_e case from the slightly lower baseline. When broken down in more detail it was clear that the refinement algorithm helped to finesse the location providing the broad region was correctly identified whilst the BDT improvements served to make larger improvements to a limited set of pathologies.

5.2.1.4 High Level TPC Reconstruction

The output of the Pandora pattern recognition consists of a series of slices, each representing a three-dimensional particle hierarchy made up of vertices, 'particles' (PFOs) and space points. The reconstructed PFOs are then fed through a series of tools designed to further characterise the particles, producing variables useful to physics analysis. The space points that make up each particle are used to fit the reconstructed track and shower hypotheses. Traditionally, only the PFOs labelled as track-like by Pandora would undergo the track fitting, whilst only the shower-like PFOs would undergo the shower fitting. However, SBND recently began to apply both characterisations to all PFOs, allowing each analysis to determine the most useful output individually.

The track reconstruction uses sliding linear fits to create a fitted trajectory along the length of the track, providing both a smoothed location and a direction at each point. This information is subsequently required for the calculation of the $\mathrm{d}E/\mathrm{d}x$ of the track. For each point along the fitted track the effective pitch ($\mathrm{d}x$) can be calculated in each plane, as can the charge deposited on the wire ($\mathrm{d}Q$) accounting for the conversion from ADC counts and attenuation effects. Together this yields a $\mathrm{d}Q/\mathrm{d}x$ value per point per plane, which can be converted into the $\mathrm{d}E/\mathrm{d}x$ by accounting for recombination using the modified box model with values measured by the ArgoNeuT experiment [189]. The variation in $\mathrm{d}E/\mathrm{d}x$ can be used to look for Bragg peaks in stopping tracks and the total calorimetric energy can be calculated by summing the product of the $\mathrm{d}E/\mathrm{d}x$ and the pitch at each point. Further tools estimate the momentum of the particle via range or multiple coulomb scattering, where the former is only appropriate for particles that stop within the

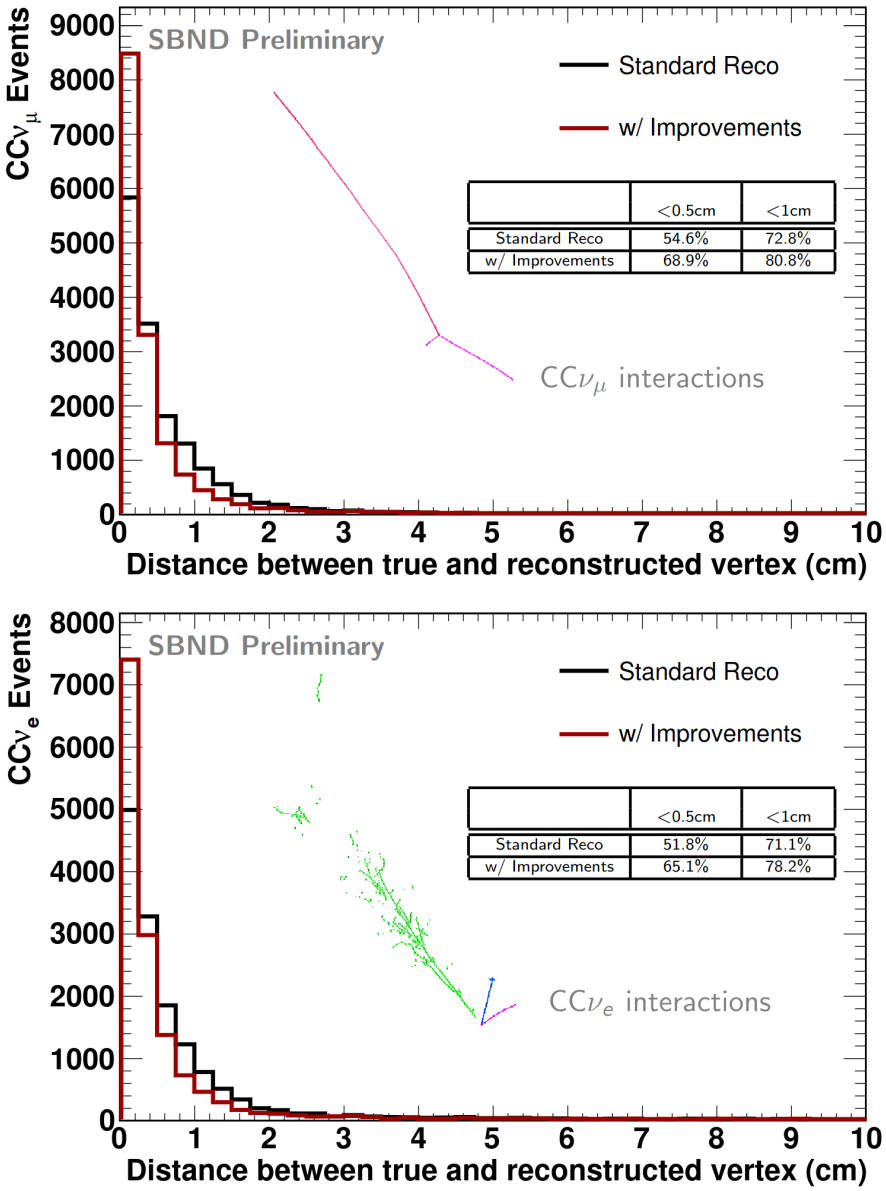


Figure 5.16: The improvement in vertex reconstruction for CC ν_μ (top) and CC ν_e (bottom) events following the combination of the inclusion of the refinement algorithm and the updates to the boosted decision tree The black curves show the original vertex reconstruction whilst the dark red shows the combined updates.

detector.

The shower reconstruction is more complex due to the underlying complexity of electromagnetic showers. The key metrics extracted by the shower fit are the start point, length, direction, opening angle, energy and $\mathrm{d}E/\mathrm{d}x$. The opening angle and length encode the size and shape of the cone containing the shower's activity. The $\mathrm{d}E/\mathrm{d}x$ is a more difficult concept to define for a shower type topology than for tracks. The algorithm attempts to identify the initial 'track stub' before the shower develops into a large number of particles and calculates the energy deposition along just this length.

5.2.2 PDS

The PDS reconstruction proceeds in a similar fashion for both the PMTs and the X-ARAPUCAs and consists of two stages: hit finding and flash construction. A hit represents a single photodetector (PD) above threshold, whilst a flash is intended to represent the total light created by a single interaction in the TPC. The raw waveforms produced by both types of PDs are put through a deconvolution process to remove the effects of the detector responses. For the PMTs an initial smoothing step is applied to mitigate the effects of noise in their high frequency sampling. Both types are then put through a Gaussian filter which further eliminates noise and removes the bipolar nature of the signal.

The hit reconstruction algorithm operates by looking for peaks in the deconvolved waveform that pass a threshold based on a combination of the deconvolved single electron response and the baseline root mean squared value. The number of photoelectrons that arrive at a PD is represented by the integral of the hit and the arrival time is defined by the time at which the waveform passes 15 % of the maximum peak. This helps to reduce the time biases introduced by smearing in the deconvolution process and the likelihood of multi-photon arrival particularly in the prompt portion of the light.

Optical hits with similar arrival times are then combined across multiple PDs to form flash objects. Flashes are only created when enough PDs see a certain number of photoelectrons in a defined time interval. Its start time represents the interaction time and will later be

used to help resolve the drift time ambiguity in the charge reconstruction and help reject cosmic activity outside of the beam arrival window. SBND Monte Carlo studies, performed using real detector efficiencies and light yields from teststand results, have demonstrated the ability of the PDS to achieve 2 ns timing resolution and thus resolve the substructure (81 buckets) of the BNB neutrino spill. To make this distribution resolvable requires correction factors for both the neutrino time of flight and then the photon time of flight in the detector. Both can be calculated using the PDS only, the former thanks to the high granularity in the yz-plane whilst the latter utilises the separate sensitivity to both VIS and VUV light and the reflective cathode foils in order to resolve the interaction position in x [254].

5.2.3 CRT

The CRT reconstruction workflow was re-developed from the ground up by the author to provide more flexibility and better handling of high multiplicity scenarios. Similarly to the TPC and PDS workflows the CRT reconstruction begins with a 'hit-finding' procedure. This is followed by clustering of these hits within each tagger wall to form three-dimensional objects and finally the linking of such objects into tracks between multiple walls.

5.2.3.1 Strip Hit Reconstruction

The CRT readout is significantly different to that of the TPC and PDS in that it does not comprise of full waveforms. Instead, the front-end-boards (FEBs) provide a single ADC value for each of the 32 channels. The board amplifies and shapes the signal from each SiPM and provides a threshold trigger with a 50 ns coincidence requirement between the two SiPMs that service each strip. If this requirement is met then all 32 channels are sampled and the values held in bias until the board's CPU makes a decision to readout. The shaping and sampling time are configured such that the rising edges of the channels that triggered the readout will be at their peak value at the time they are being sampled. The readout decision can, depending on the hardware setup, be conditional on the arrival

of a similar trigger signal from another module arranged perpendicularly to the module in question. Readouts can also be triggered by the arrival of a reset signal to either of the board's two nanosecond clock counters. The type of readout is indicated by a 'flags' variable which also encodes the 'state' (well defined or not) of each of the clocks at the time of that readout. Once the CPU initiates a readout, the 32 channels are passed from the holding circuit through an analogue-to-digital converter (ADC) to produce a single 12-bit value per channel.

A single FEB readout therefore consists of the flags indicator, the time according to each of the clocks and 32 ADC values. Further metadata is also stored relating to: the time since the last readout, information on the group of readouts removed from the board's buffer in each 'poll' cycle and information from the server receiving the CRT readouts. The first job the reconstruction must perform is to reduce the 32 values in each readout down to only those strips of interest. This is done via a hit finding algorithm that evaluates each FEB readout looking for concurrent pairs of channels with values at least 60 ADC above their pedestal. The general principles of identifying strips of interest and then using the light balance to reconstruct a lateral position is shown in figure 5.17. The lateral position within the strip, d, is reconstructed using the parameterisation

$$d = \frac{w}{2} \cdot \left(1 + \tanh\left(\log\left(\frac{\mathrm{ADC}_2}{\mathrm{ADC}_1}\right)\right)\right), \tag{5.6}$$

where w is the strip width (typically 11.2 cm) and ADC_1 and ADC_2 represent the ADC counts recorded for the two channels. The hyperbolic tangent function has a range of (-1,1) which constrains the position within the physical bounds of the strip whilst the logarithmic ratio best represents the light balance between the two channels. A number of approaches were tested for the lateral position reconstruction before selecting the one used in equation 5.6. The performance of three such approaches are shown in figure 5.18. A nominal error on the position is constructed using an order-2 polynomial with coefficients extrapolated from a fit to a Monte Carlo study of the resolution of this approach. The reconstructed object also contains three time fields, the two clock values (with a provision included for applying cable delays) and a global time provided by the server, as well as a

pair of booleans to indicate whether or not the ADC value for each channel had hit the saturation point of 4095 ADC. This is critical information in the reconstruction as it limits both the calorimetric and positional reconstruction.

The left hand plot in figure 5.19 indicates the typical number of strip hits expected in a single recorded event in which CORSIKA was used to simulate cosmic-ray particles throughout the detector readout. The right hand plot shows the performance of the positional reconstruction via the difference between the true and reconstructed positions which corresponds to a positional resolution of 2.5 cm (including the 1 cm strip depth within which the reconstruction has no handle to resolve). Although this kind of resolution appears coarse when compared to a fine grain detector such as the TPC, this plot demonstrates that the light balance method creates a finer resolution than would be achieved by purely using the strip centre.

5.2.3.2 Clustering and Space Point Building

The next stage of the reconstruction aims to correlate strip hits in space and time and reconstruct three-dimensional space points that pinpoint the locations of the activity in the CRT taggers. A clustering algorithm groups any strip hits in the same tagger that fall within a 50 ns window of each other. A splitting analysis is performed which evaluates whether or not the cluster's hits overlap in distinct regions, any clearly distinct areas of activity will be split off from the cluster. This happens rarely, but not vanishingly so. Roughly 3 % of events will have a pair of cosmic-ray muons crossing a single tagger within 50 ns of each other. There will also be other scenarios that will create multiple areas of activity in a single time window, such as multiple particles from a neutrino interaction escaping the TPC, or hadronic particles such as protons or pions created from muon hard scatters in the vicinity. Figure 5.20 shows a couple of event displays with different clusters of strip hits indicated across multiple taggers.

A 'back-tracking' algorithm was written to identify the simulated energy deposits that contributed to each simulated FEB readout, and thus eventually to the cluster. This allowed for the following definitions of completeness and purity to be used in assessing the

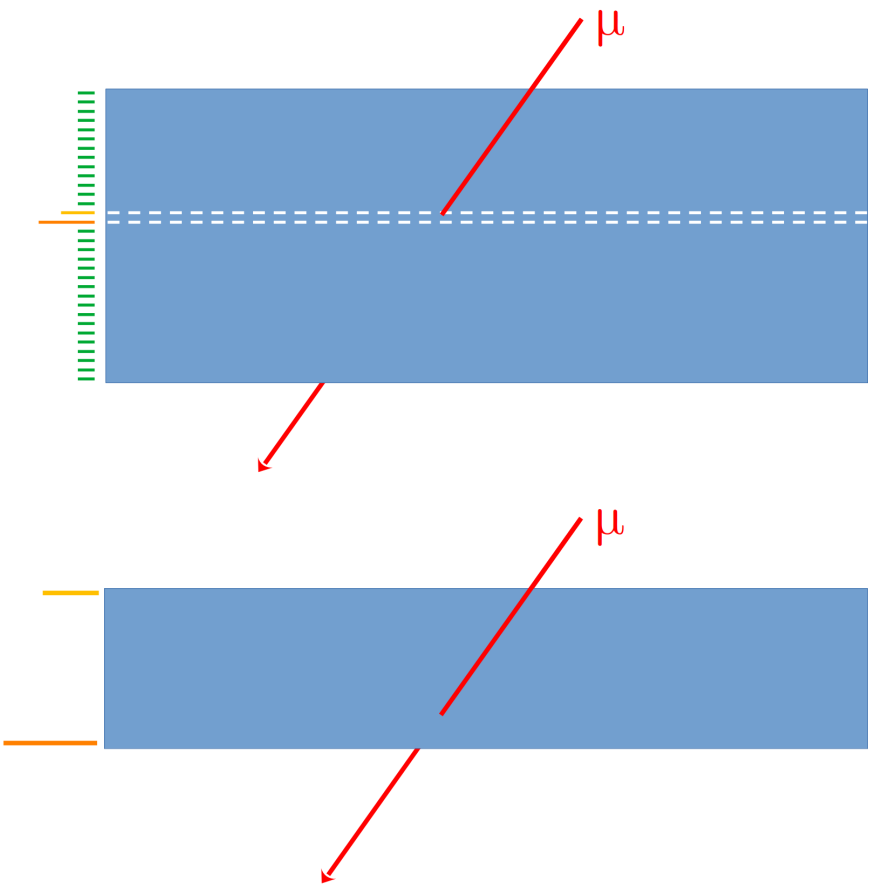


Figure 5.17: A schematic illustration of the principles behind the reconstruction of CRT strip hits. The upper panel shows the ADC readouts from an entire module. A muon crosses a single strip, illustrated by the white dashed lines, and its ADC counts are larger, indicated by the hotter colours. The lower panel shows a zoomed version of the single strip. It is clear that the relative position of the muon across the strip is creating a higher ADC count from the lower channel and a lower ADC count from the upper channel. It is this 'light balance' that will be used to reconstruct this lateral position of the muon.

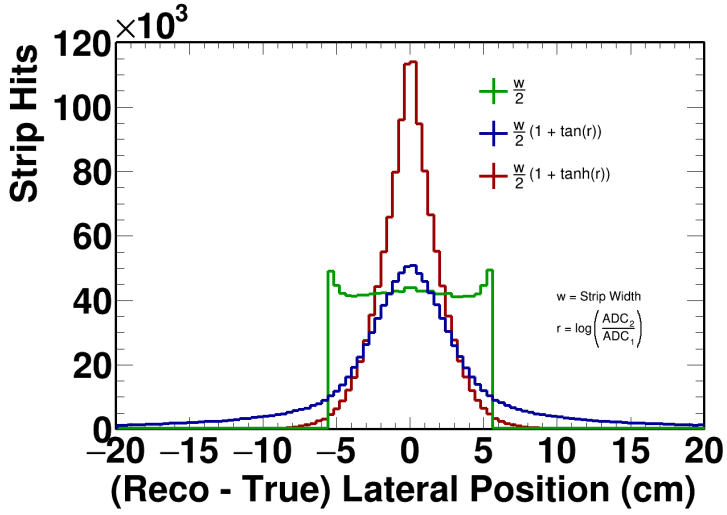


Figure 5.18: The performance of three different methods for reconstructing the lateral position within the strip: taking the strip centre (green) or using the logarithm of the light balance in the tangent (blue) and hyperbolic tangent (red) functions.

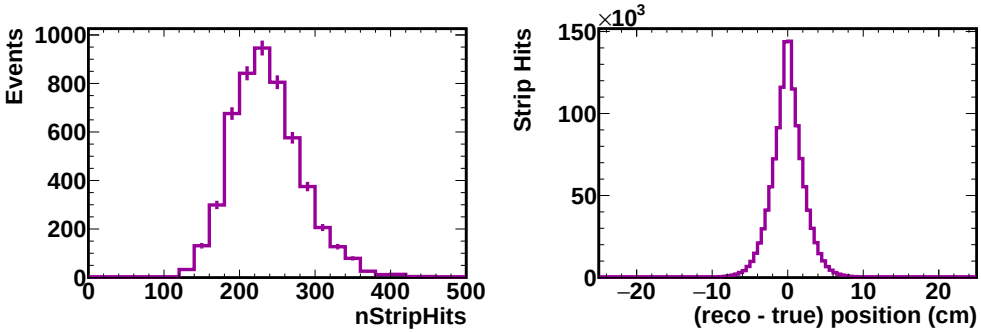


Figure 5.19: The left plot shows the number of strip hits in a typical simulation of a BNB event with CORSIKA cosmic-ray particle simulation throughout the readout. The right hand plot shows the accuracy of the positional reconstruction within the strip.

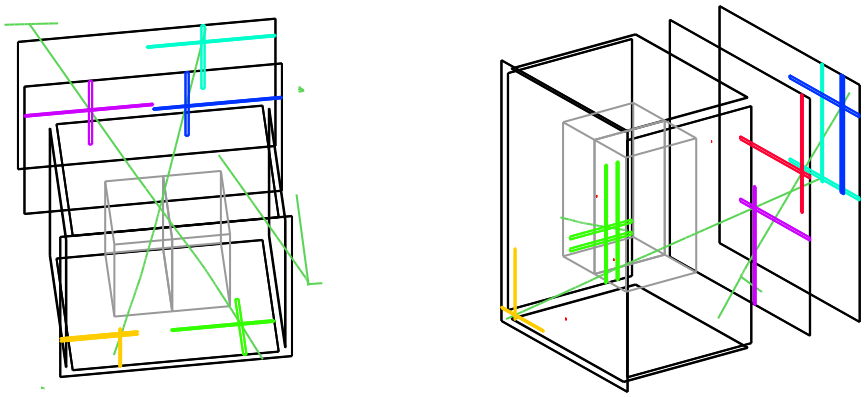


Figure 5.20: Two event displays, each from a short time snapshot of simulated cosmic-ray particle flux, demonstrating the clustering of strip hits in the CRT tagger walls. The grey boxes in the centre indicate the TPC, the black lines indicate the 7 CRT tagger walls, the green lines represent true particle trajectories and the clusters of strip hits are then represented in a range of colours, one colour per cluster.

clustering performance:

$$\text{Completeness} = \frac{\text{Total true energy from the leading particle accounted for in the cluster}}{\text{Total true energy deposited by the leading particle across the whole tagger wall}} \tag{5.7}$$

and

$$\text{Purity} = \frac{\text{Total true energy from the leading particle accounted for in the cluster}}{\text{Total true energy from any particles accounted for in the cluster}}. \tag{5.8}$$

Figure 5.21 shows both the completeness and purity of the clusters produced using the above algorithm. They indicate that the performance of the clustering is very good, over 90 % of clusters have both a completeness and purity of above 95 %. It also shows plots that indicate the general make up of the clusters. The majority of clusters have 2 strip hits, but there are significant numbers with 1, 3 or 4 as well. The composition variable indicates which coordinate directions can be constrained by the strip hits in each cluster. Each strip natively constrains two coordinates (via its depth and width) whilst the third (length) is only constrained to an order of metres. The combinations of differently oriented strips within one cluster can,

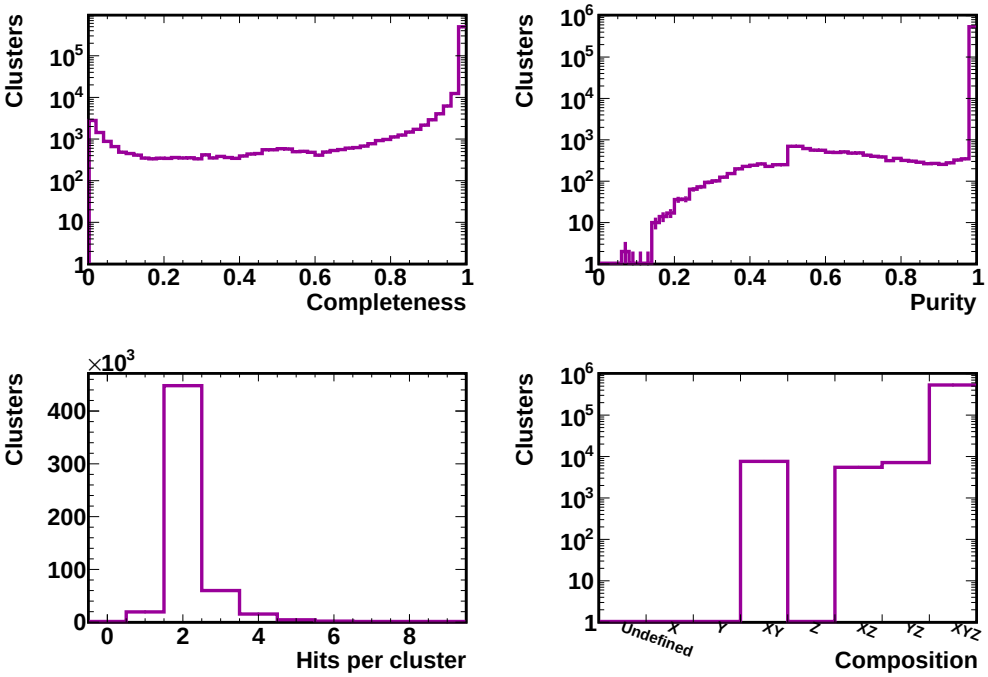


Figure 5.21: The performance of the CRT clustering algorithm. From top left, clockwise: the completeness of each cluster according to equation 5.7, the purity of each cluster according to equation 5.8, the composition of the clusters in terms of the coordinate directions they constrain and finally the number of hits per cluster.

however, result in a full three-dimensional constraint. The arrangement of two perpendicular layers of modules in each wall was designed to do just this, and it is clear that this is by far the dominant case.

A separate algorithm is used to 'characterise' the clusters and produce space point objects, representing a three-dimensional location. This allows for future workflows which may want to make use of the clustering output with alternative methods for characterising them. Different methods are applied for clusters containing just a single strip hit compared to those with multiple strip hits. The single hit method was written with the CRT♭ in mind, most of which only has single layer coverage. Single hit clusters in other walls are dropped at this point.

The single hit method takes the lateral position from the strip hit and converts it to a three-dimensional position by considering the geometry of the strip in question. Its position is therefore well constrained in two dimensions, the width of the strip and the depth of the strip, but is only constrained to the order of metres in the length dimension. The summed photoelectron count, as a representation of the space point's energy scale, is calculated to be

$$N_{\mathrm{PE}} = \sum_i g_i \mathrm{ADC}_i \tag{5.9}$$

where ADC_i is the ADC recorded on channel i and g_i is that channel's calibrated gain value. Finally, one of the clocks is selected to provide the time (which one is used is configurable) and a configurable offset can also be applied. The space point objects therefore have a single time that has a common reference (likely the trigger time) to objects from other systems allowing for combined use of the reconstruction outputs from the whole detector. This marks a transition from more 'raw' focused objects to more analysis focused objects.

If there are multiple hits then the first stage is to construct all possible combinations of two overlapping strips, each is made into a temporary space point. The position is defined as the centre of the area of overlap, with the size encoded in the error on

the position. This is shown on the green cluster in figure 5.22. The overlap area is defined via the lateral position (as calculated in equation 5.6) and error associated with the strip hit, not the full width of the strip. The construction of a three-dimensional position allows the reconstruction to account for the attenuation and delay of the light pulse along the length of the strip. The photoelectron count, N_{PE}, is constructed from the ADC as with single hit clusters and then corrected to

$$N'_{\mathrm{PE}} = N_{\mathrm{PE}} \cdot \frac{(l+K)^2}{K^2}, \tag{5.10}$$

where l is the reconstructed distance from the energy deposit to the SiPMs along the length of the strip and $K = 1085\,\mathrm{cm}$ is a model parameter derived from test stand results and used in the simulated application of the attenuation as well [255].

A similar process is applied to the time reconstruction. The time of each strip hit is corrected for two effects: propagation delay and time walk. The propagation delay is the time taken for photons to travel the length of the optical fibre to reach the SiPMs whilst the time walk effect is a result of larger pulses passing the threshold quicker and therefore appearing to have occurred earlier. The final corrected time is given by

$$\begin{aligned} t' = t & \\ & - (l \cdot D_{\mathrm{Prop}}) \\ & - C_{\mathrm{TW}} \cdot e^{-\frac{1}{2}\cdot\left(\frac{N'_{\mathrm{PE}} - P_{\mathrm{TW}}}{\sigma_{\mathrm{TW}}}\right)^2} + t_{\mathrm{TW}} \end{aligned} \tag{5.11}$$

where t is the uncorrected time and $D_{\mathrm{Prop}} = 0.061\,\mathrm{ns/cm}$, $C_{\mathrm{TW}} = 4125.74\,\mathrm{ns}$, $P_{\mathrm{TW}} = -300.31\,\mathrm{PE}$, $\sigma_{\mathrm{TW}} = 90.392\,\mathrm{PE}$ and $t_{\mathrm{TW}} = -1.525\,\mathrm{ns}$ are all parameters resulting from test stand results accounting for the propagation delay (Prop) and time walk effects (TW) respectively [255]. Both of these corrections are, at the most, of the order of 20 ns.

The temporary space points for each of the possible overlaps are then combined

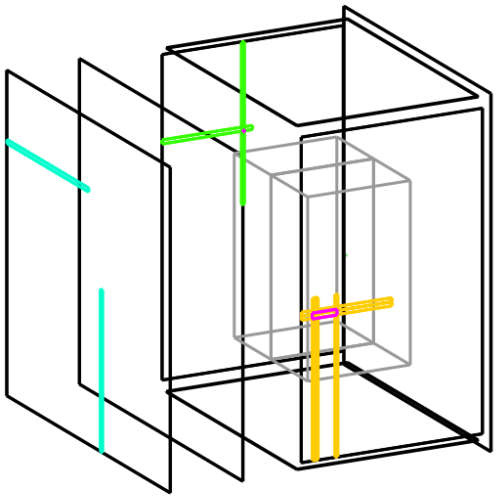


Figure 5.22: An event display demonstrating the construction of space points from CRT clusters. The grey boxes in the centre indicate the TPC, the black lines indicate the 7 CRT tagger walls, the clusters of strip hits are then represented in a range of colours, one colour per cluster, and space points are shown by the magenta areas at regions of overlap.

into a final space point for each cluster. The positions and times are averaged and the photoelectron counts are summed. The errors on the positions and times are assigned as the error on the mean. The space point associated with the yellow cluster in figure 5.22 is a typical example of a space point created from more than two strip hits.

Figure 5.23 shows the performance of space point reconstruction for multi-hit clusters. The positions are well constrained, with a resolution of 5 cm and the vast majority within the 11 cm×11 cm area created by a pair of overlapping strips. Note that, due to the construction of some space points from more than just two perpendicular strips, this resolution is coarser than the simple combination of the resolutions of two single strips. The timing resolution, of around 2.7 ns, is also close to the intended ns-level with further improvements expected, as mentioned in section 4.4.3. The importance of the corrections applied in equation 5.11 are clear given that scale of these effect would significantly increase that resolution if not corrected for. The PE is clearly representative of the deposited energy, the linear relation applies well throughout the high statistics region before saturation of the SiPM ADC takes over at the higher energies. Finally, the bottom right plot, demonstrates that for energy depositions of > 4 MeV the reconstruction efficiency is ~ 100 %, making for an integrated efficiency of 91.0 %. There are two populations in the energy distribution. Minimum ionising particles passing through the polystyrene scintillator will deposit around 1.94 MeV/cm [32], so a particle passing through both layers (2 cm) of one of the taggers will deposit around 4 MeV. This causes the activation of the second peak at this value, as well as driving the reconstruction efficiency, as a space point requires energy in both layers to be reconstructed. The lower peak represents particles, primarily radiogenic backgrounds, that only deposit energy in a single layer of scintillator thus not passing the CRT's internal trigger requirement.

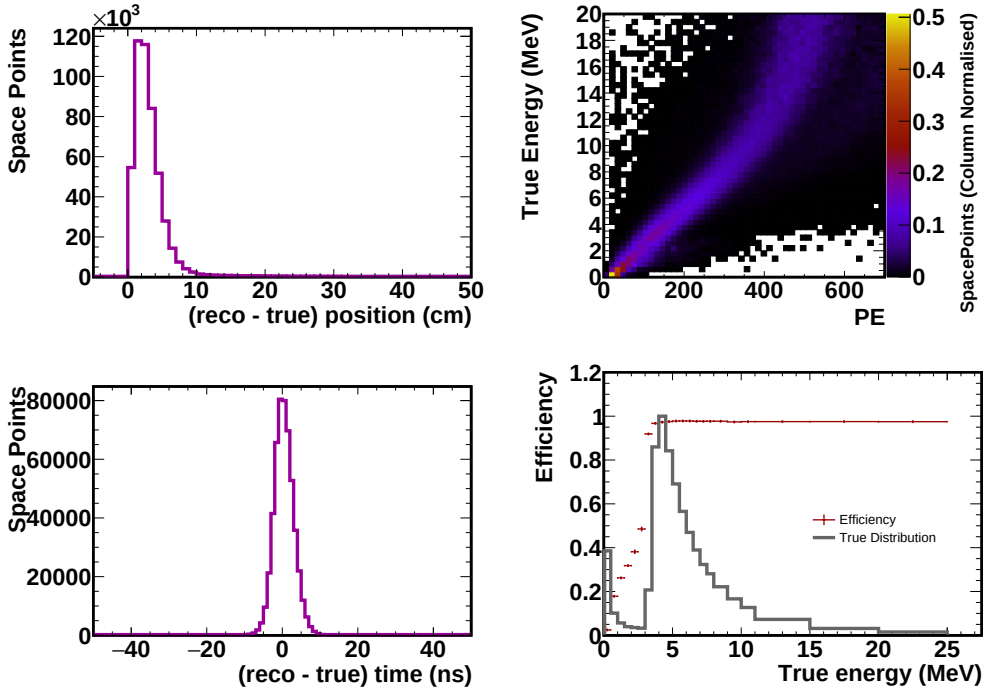


Figure 5.23: The performance of the CRT space point creation. From top left, clockwise: the accuracy of the reconstructed space point positions, the linear relation of reconstructed photoelectron count to the true energy deposited, the reconstruction efficiency of true energy depositions in the tagger and the accuracy of the reconstructed time.

5.2.3.3 Track Reconstruction

The final stage of the CRT reconstruction aims to combine space points from multiple tagger walls into tracks representing the trajectory of a particle, usually a muon, passing through the detector region. Firstly, the space points in each tagger wall are ordered in time and then candidate tracks are produced from any combination of space points that occur in different taggers within 100 ns of each other. These combinations can contain two *or* three space points. The latter occurs when a track uses space points produced from both of the two top walls and one of the other five walls. In this scenario the third hit is subject to the same time coincidence requirement but must also lie within 50 cm of the extrapolated track formed by the initial two space points. For the simpler tracks, the start and end points are merely the positions of the two space points. For the three space point scenario, a best fit line is found between the three points. The points at which that line intersects the three taggers are used as the track points.

The candidates are ordered by the goodness of agreement of their timing (accounting for time of flight) with three point tracks being prioritised over two point tracks. The candidates are then saved if none of their space points have been used in any preceding candidates with better timing agreement. Two demonstrations of the CRT tracking are shown in event displays in figure 5.24 where the reconstructed CRT track clearly follows the trajectory of the true muon. As with the clustering, the purity and completeness of the reconstructed tracks are useful metrics to assess the quality of the decision making process in the tracking algorithm. In the tracking case they are defined as

$$\text{Completeness} = \frac{\text{Total true energy from the leading particle accounted for in the track}}{\text{Total true energy deposited by the leading particle across all tagger walls}} \tag{5.12}$$

and

$$\text{Purity} = \frac{\text{Total true energy from the leading particle accounted for in the track}}{\text{Total true energy from any particles accounted for in the track}}. \tag{5.13}$$

Figure 5.25 shows the resulting distribution of these metrics. About 93 % of tracks

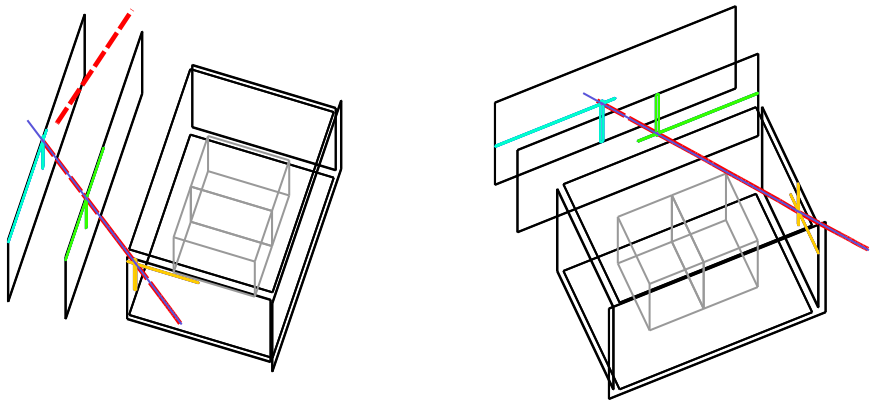


Figure 5.24: Two event displays showing a short period of simulated cosmic-ray particle flux. They demonstrate the tracking of cosmic-ray muons with the CRTs. The grey boxes in the centre indicate the TPC, the black lines indicate the 7 CRT tagger walls, the strip hit clusters are represented in a range of colours, one colour per cluster, with the tracks represented by the blue line and the true muon trajectory in red dashes. Each shows good tracking of a single cosmic-ray muon. The left-hand display also shows a second true muon which escapes the CRT system having only intersected one wall and thus is not tracked.

have both completeness and purity above 90 %. Also shown are the length resolution and reconstruction efficiency. Similarly to the space points the reconstruction efficiency switches on above a threshold (∼8 MeV) corresponding to the energy deposited across the 4 cm of polystyrene required to form a reconstructable track (two layers in each of two walls). The total integrated efficiency is 96.2 %. The length resolution is about 10 cm and is dependent on the particular pair of walls considered due to geometric effects.

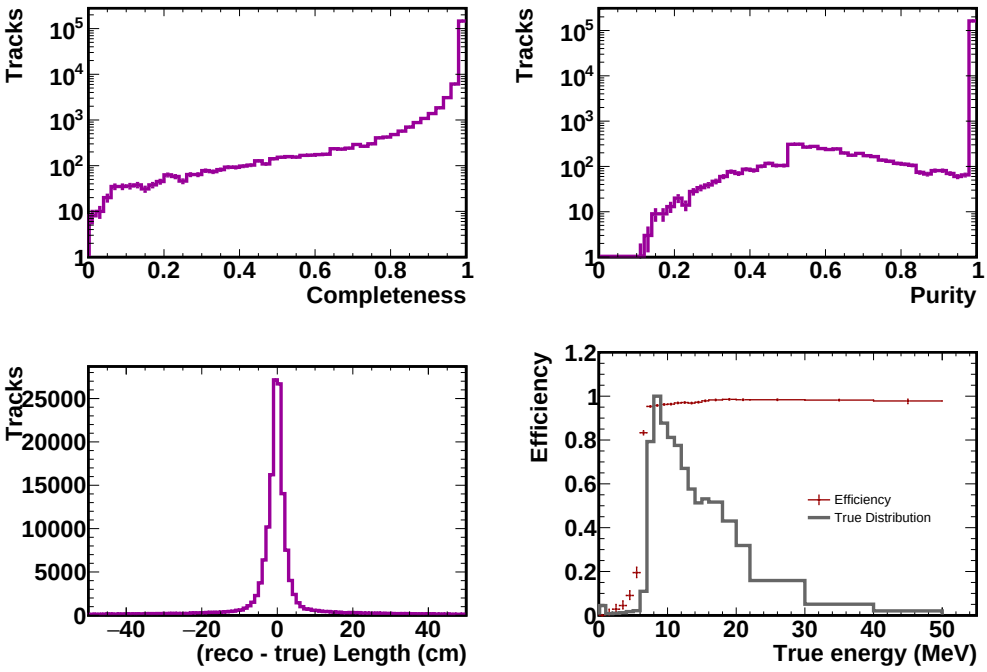


Figure 5.25: The performance of the CRT track creation. From top left, clockwise: the completeness of each track according to equation 5.12, the purity of each track according to equation 5.13, the reconstruction efficiency for particles that leave true energy depositions across multiple taggers and the accuracy of the reconstructed track length.

5.2.4 Multi-System Matching

The final stage of the reconstruction chain involves high-level tools that aim to link information between the three different sub-systems by correlating their respective positional reconstruction information. The primary goal of these tools is that of cosmic rejection. The CRT and PDS both provide nanosecond level timing resolution that can be used to tag an associated TPC slice as arriving before or after the beam, and therefore rule it out as a neutrino candidate.

SBND has two 'flash-matching' algorithms, both of which aim to associate a flash recorded by the PDS with a TPC slice, *OpT0Finder* and *SimpleFlash*. They both consider the charge recorded in a slice and use that to make a hypothesis of what they would expect the photodetectors to see from the corresponding scintillation emission.

The SimpleFlash algorithm constructs its own flash objects - like those described in section 5.2.2 they aim to represent the total light emitted by a single interaction. A large sample of simulated neutrino events is used to produce metrics describing the average yz span of the flashes, dependent on their charge distribution. It then constructs a score that represents the agreement of the measured flash with the values obtained from the pre-simulated library. The OpT0Finder algorithm instead utilises the output of the dedicated flash reconstruction, considers the slice hit-by-hit and uses lifetime and recombination to model the corresponding energy deposition of charged particles in the TPC. The related light production from that energy deposition is calculated using either an assumption of linearity or the more physically correct anti-correlation relationship depending on the scenario. The latter can be used for contributions from track-like objects due to the ease of calculating a $\mathrm{d}E/\mathrm{d}x$ along their length, however, for the contributions from shower-like objects this is not the case and therefore the former is utilised. This light production is passed through the semi-analytic photon library used in the detector's light

simulation, and the relevant photodetector efficiencies are then applied to create a hypothesis of the observed photoelectrons in each photodetector. Finally, a χ^2 score is used to assess the agreement between the hypothesised and measured photoelectron counts, allowing the slice that agrees best with the light signature to be identified. Due to the larger computational intensity of this algorithm it is performed on flashes during the neutrino beam window only.

In order to link the TPC and CRT information only the track-like PFOs are considered. This is due to both the likelihood of the particle entering / escaping the TPC and appearing in the CRTs, as well as the ease with which TPC tracks can be extrapolated to evaluate their likely intersection points with CRT tagger walls. There are two matching algorithms, one which attempts to match the TPC tracks to CRT space points and a second which attempts to match TPC tracks to CRT tracks. Both use the time of the CRT object as the basis for determining the x position of the TPC track. The space point matching compares the extrapolated intersection of the TPC track with the relevant CRT tagger and computes the distance of closest approach (DCA) to the space point in question. A loose cut is applied to the DCA but the value is retained for analysis level harsher cuts. The algorithm was updated by the author such that each track and space point can only be used once, retaining only the match with the shortest DCA. For the track matching algorithm, a compound score is created from the sum of the average DCA stepping along the TPC track and (four times) the angle between the directions of the two tracks. The same philosophies of applying a loose cut and ensuring no duplication are used to decide which matches are saved.

Figure 5.26 shows the efficiency and quality of the CRT-TPC matching, defined as

$$\text{Efficiency} = \frac{\text{Number of TPC tracks with a CRT match}}{\text{Number of TPC tracks whose truth matched particle created a CRT space point / track}} \tag{5.14}$$

and

$$\text{Quality} = \frac{\text{Number of CRT matches with the same truth matched particle as the TPC track}}{\text{Number of CRT matches}}. \tag{5.15}$$

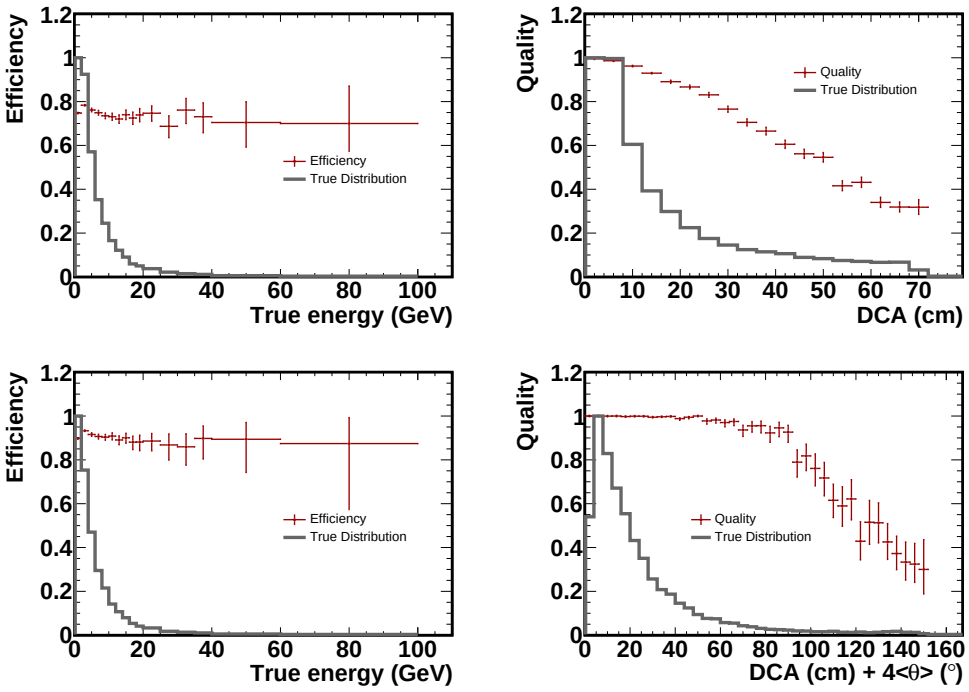


Figure 5.26: The performance of the CRT-TPC matching. From top left, clockwise: the efficiency of CRT space point to TPC track matches according to equation 5.14, the quality of CRT space point to TPC track matches according to equation 5.15, the quality of CRT track to TPC track matches according to equation 5.15, the efficiency of CRT track to TPC track matches according to equation 5.14.

The quality plots demonstrate how the quality of the matches decreases as the respective scores increase, introducing more false positives at higher score values. The loose cuts applied in the algorithm allow these distributions to fall off significantly but preserves the very strong efficiency seen in the other plots. This combination preserves the ability of each analysis to make informed decisions on the correct compromise for their physics goal. Even before such cuts the space point matching has an integrated quality of 87.5 % and the track matching 98.0 %. The efficiencies sit at 75.3 % and 90.7 %, respectively, with the lower space point matching efficiency mainly driven by smaller TPC tracks with poorer pointing information.

SBND has also developed a CRT-PMT matching algorithm that constructs a time-

of-flight metric to distinguish between incoming stopping cosmic-ray muons and exiting muons from neutrino interactions, both of which generate a CRT space point and PMT flash but can otherwise look very similar. This tool is not currently utilised in the downstream analysis tools or the selection presented in chapter 7 so further details are omitted.

Chapter 6

Rejecting Cosmic Backgrounds, the CRUMBS Tool

This chapter describes an analysis agnostic cosmic rejection tool developed by me and widely used amongst most current SBND Monte Carlo studies. The credit for the idea must go to Dominic Brailsford, and it relies heavily on reconstruction work performed by many SBND collaborators and others in the liquid argon community, but the development, validation and implementation were all carried out by myself.

6.1 CRUMBS Concept

As has been touched upon in previous chapters, SBND is a surface located detector, with no shielding in place to reduce the cosmic ray flux. This is particularly challenging given the 'slow' nature of a LArTPC readout. At the nominal drift field strength of $500\,\mathrm{V\,cm^{-1}}$ it takes 1.3 ms for an electron to cross the entire drift length. Extra readout length known as a 'porch' is added to either side of this full drift time such that a full readout will be about 1.7 ms. An envelope calculation

using the cosmic-ray muon flux at sea level of $\sim 1\,\mathrm{cm}^{-2}\,\mathrm{min}^{-1}$ [256] and the SBND 'footprint' of $20\,\mathrm{m}^2$ yields a rate through the detector of 5-6 muons per readout. A similar calculation suggests a muon will cross the detector once in every 190 BNB spills (of $1.6\,\mu$s each). This compares to the likelihood of a neutrino interaction once in every 20 spills. This implies that around 10 % of activity observed to be coincident with the beam actually originates from the cosmic-ray background.

These two calculations speak to two distinct categories of cosmic background that will be regularly referred to throughout the following chapter.

- **In-time Cosmics**: Events in which a cosmic-ray muon crosses the detector coincident with the neutrino beam spill passing through the building, thus inducing a 'beam + light' trigger. This background can be evaluated using runs taken in the absence of the neutrino beam, with a fake beam window used to create a 'fake beam + light' trigger.
- **Out-of-time Cosmics**: Regardless of what induced the trigger, each event will contain a number of out-of-time cosmics that entered the detector outside of the beam arrival but within the detector readout.

The Pandora unambiguous cosmic removal described in section 5.2.1.2 removes hits associated with very cosmic-like activity but is deliberately cautious, so many of the slices processed through the full neutrino pass are still of cosmic origin. In particular, the unambiguous removal essentially removes all through-going muons but retains many stopping or exiting events. The first step in any SBND neutrino analysis is to remove these backgrounds and identify strong neutrino candidate slices. The Pandora reconstruction output contains a BDT score referred to as the 'Nu Score' which indicates how neutrino-like each remaining slice is. Figure 6.1 shows the resulting distribution of this score for true neutrino and cosmic slices in a simulated sample. It can be seen that this parameter provides a reasonable degree of separation between neutrino interactions and their cosmic-ray backgrounds.

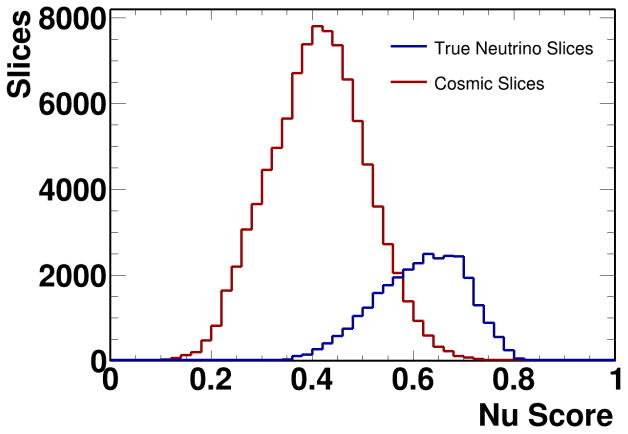


Figure 6.1: The Pandora 'Nu Score' for slices created by simulated neutrinos or cosmic-ray muons respectively.

By definition, the Pandora Nu Score only considers information provided by the TPC. Most 'traditional' analyses began with a cut on this score and proceeded with cuts on variables provided by one of the flash matching algorithms. The CRUMBS tool, standing for *Cosmic Rejection Using a Multi-system BDT Score*, aims to harness the capabilities of all three of SBND's sub-systems in one go, therefore avoiding the summed inefficiencies of each cut and utilising the complementarity of the different rejection variables. It can be thought of as an enhanced version of Pandora's Nu Score - it produces a single BDT score output and uses the Nu Score variables as the majority of its TPC inputs.

6.2 Input Variables

The CRUMBS tool takes inputs from a series of variables from each of the sub-systems. The TPC variables provide topological and calorimetric discrimination between signatures of neutrino interactions and cosmic-ray backgrounds. The PDS and CRT variables are constructed from the tools that match their outputs to that of the TPC. They characterize the timing and topology of the matched slice

and that timing information can be linked to the known time of the beam spill to discriminate in-time neutrino interactions from out-of-time cosmic-ray backgrounds. Whilst selected plots will be shown through this section, all of the input variables are shown in appendix B.

6.2.1 TPC

The full set of inputs that create the Nu Score are also used in CRUMBS. Pandora reconstructs every slice under two hypotheses, 6 of the variables are calculated from the neutrino hypothesis reconstruction and 4 from the cosmic hypothesis. In general, the neutrino hypothesis variables attempt to identify key characteristics of neutrino interactions such as particle multiplicity at the vertex and general momentum along the beam axis whilst the cosmic hypothesis variables attempt to identify more cosmic-like characteristics such as the dominance of a single large track and general downward momentum.

Neutrino Reconstruction Variables

- The total number of space points in the reconstructed slice.
- The number of PFOs in the slice.
- The vertical (y) position of the reconstructed interaction vertex.
- The z-component (beam direction) of the space point weighted direction of the immediate children of the reconstructed neutrino.
- The number of space points within 10 cm of the vertex.
- The ratio of the secondary to the primary eigenvalues from a PCA of the space points within 10 cm of the vertex. These values represent the strength of the correlation of the space points along the primary and secondary axes, thus the

ratio represents a metric of the local directionality around the vertex. Higher values representing more spread space points, lower values representing most space points being correlated along a single direction from the vertex.

Cosmic Reconstruction Variables

- The fraction of the slice's space points that belong to the longest track.
- The difference in the initial and final directions of the longest track (technically $1 - \cos(\theta_{if})$)
- The y-component (vertical) of the longest track's direction.
- The number of space points in the largest PFO.

One further TPC variable is used, making use of the calorimetric variables produced after the pattern recognition. A particularly problematic cosmic background occurs when a cosmic-ray muon enters the TPC but comes to a stop without passing through. In this scenario, especially when a decay electron is not identified, the signature can look geometrically identical to a muon produced in a CC ν_μ interaction which proceeds to escape the detector. However, these two cases can be distinguished by the presence of the distinctive 'Bragg peak', in which the energy deposited by an ionising particle increases as its velocity decreases, culminating in a sharp peak as the particle comes to a stop [257]. Figure 6.2 shows how this can be identified in reconstruction by evaluating how a track's $\mathrm{d}E/\mathrm{d}x$ varies with respect to the distance from the end of the track, a quantity referred to as residual range. For CRUMBS, a variable is constructed by fitting both a 0^{th}-order polynomial and an exponential function to the $\mathrm{d}E/\mathrm{d}x$ - residual range distribution for the longest track in the slice and then taking a ratio between the χ^2 values for these two fits.

TPC Calorimetry Variables

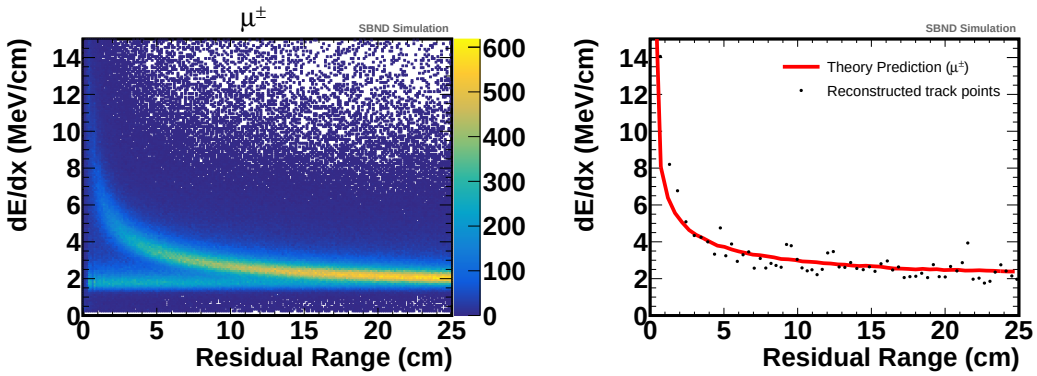


Figure 6.2: The left hand plot shows the reconstructed $\mathrm{d}E/\mathrm{d}x$ with respect to the residual range for true muons in SBND Monte Carlo. The two populations show the rising deposition for stopping muons at low residual range whilst exiting muon show constant energy deposition right down to zero residual range. The right hand plot shows this distribution for a single muon candidate with the profile from a theoretical prediction for muons overlayed.

- The ratio between the χ^2 agreement for 0^{th}-order polynomial and exponential fits to the longest track's $\mathrm{d}E/\mathrm{d}x$ -residual range distribution.

Figure 6.3 shows two of the TPC input variables, one each from the neutrino and cosmic-ray hypotheses. Both show significant discriminating power to contribute to the CRUMBS BDT. They show that generally cosmic-ray slices are more likely to be reconstructed with a vertex at the top or bottom of the TPC than neutrino slices, and that neutrino slices tend to have a moderate number of reconstructed space points in their largest reconstructed object with cosmic-ray slices having a much wider range of sizes. The latter results from the fact that cosmic-ray slices have a range of populations such as those created by muons that 'clip' the edge of the TPC or smaller particles thrown off from cosmic-ray air showers as well as the more traditional large crossing muon tracks.

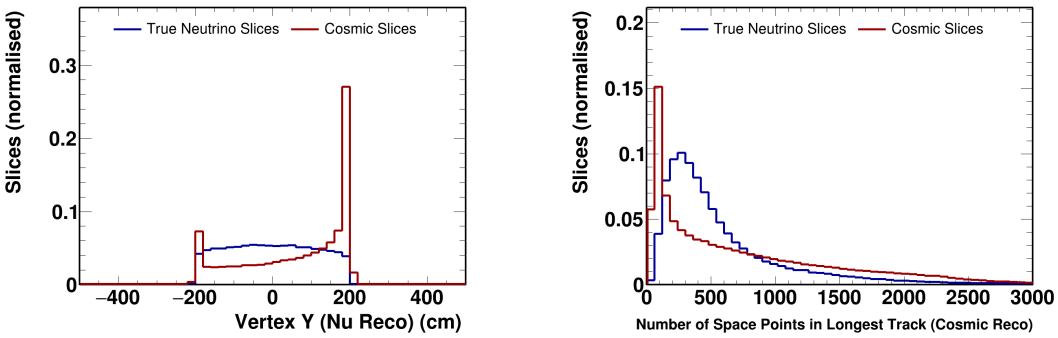


Figure 6.3: Two TPC variables used as CRUMBS inputs. The position of the neutrino vertex in the vertical y coordinate under the neutrino hypothesis reconstruction (left) and the number of space points in the largest track under the cosmic-ray hypothesis (right).

6.2.2 PDS

The fundamental unit of the CRUMBS tool is a reconstructed TPC slice, that is, all the inputs must reflect the properties of that slice, and the output represents how neutrino-like that slice is. Incorporating information from the auxiliary sub-systems (CRT and PDS) requires the use of the matching tools described in section 5.2.4. In the case of the PDS, SBND's flash matching tools allow us to associate the slice to the precise timing associated with the reconstructed flash, in order to help categorise activity as in or out of coincidence with the beam spill.

The original version of CRUMBS utilised the *SimpleFlash* tool for the PDS variables, whilst subsequent versions had the capability to use either of the flash matching tools. The selection presented in chapter 7 uses *OpT0Finder* and this is currently the default configuration for other SBND analyses. No version of CRUMBS uses both flash matchers simultaneously due to their identical intended outcomes.

SimpleFlash Variables

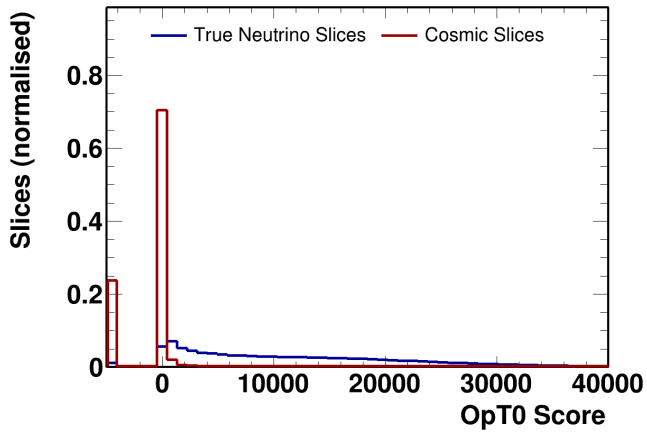


Figure 6.4: The OpT0Finder reciprocal χ^2 score used as a CRUMBS input.

- The score describing the agreement between the measured flash and the flash properties pulled from the pre-simulated library for the given slice's charge span.
- The number of PE recorded in the measured flash.
- The time of the measured flash with respect to the start of the beam spill.

OpT0Finder Variables

- The reciprocal of the χ^2 agreement between the measured flash and the flash predicted from the slice's reconstructed energy depositions.
- The number of PE recorded in the measured flash.

Figure 6.4 shows the reciprocal χ^2 variable from the OpT0Finder tool. It is clear that the neutrino slices tend to agree much better with the beam spill flash than the cosmic-ray slices, resulting in a smaller χ^2 and thus a larger score value. This offers a significant separation power between in- and out-of-time activity, helping remove many cosmic-ray slices.

6.2.3 CRT

Both of the CRT-TPC matching tools are used to generate inputs to CRUMBS. All of the track-like reconstructed PFOs in the slice are checked for matches to CRT space points or tracks. If there are multiple matches of one type within the slice then the match with the best score is retained. As with the PDS variables, these allow separation of cosmic-ray slices out-of-time with the neutrino beam spill.

CRT-TPC Matching Variables

- The score constructed from the average DCA and angle between matched TPC and CRT tracks.
- The time of the matched CRT track with respect to the start of the beam spill.
- The DCA between the extrapolation of the TPC track and the matched CRT space point.
- The time of the matched CRT space point with respect to the start of the beam spill.

Figure 6.5 shows the time of CRT space points matched to a reconstructed track from the TPC slice of interest. It is clear from the default value spike (at -3000 μs) that the majority of TPC slices do not receive a match. This is logical as many slices will be fully contained in the TPC. The timing distribution for those that do receive a match tells us a number of things. Firstly, if a neutrino slice is matched it is most likely, by an order of magnitude, to match to CRT activity in the beam spill. The cosmic activity also has a small peak here, created by in-time cosmic triggered events. Finally, it is clear that there is a small bias to cosmic activity after the beam spill compared to that before. This occurs as the TPC is read out for one full drift time after the beam spill plus a small porch on either side, so cosmics after

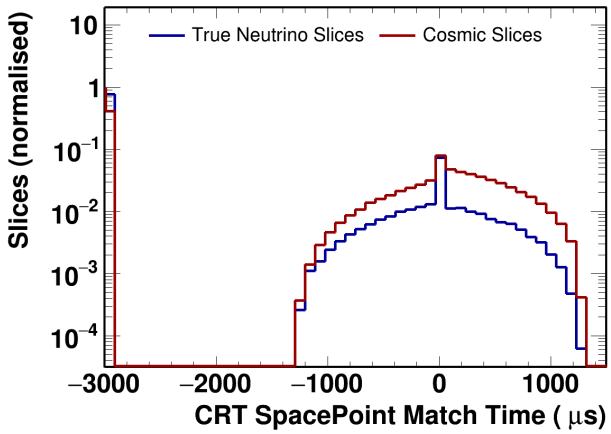


Figure 6.5: The time of the CRT space point matched to a TPC track in a slice. This is used as a CRUMBS input. The spike at -3000 μs results from the default value assigned for slices with no CRT match. The spike at 0 μs is the spike due to the additional activity that occurs at the simulated trigger time.

the beam spill are more likely to be fully reconstructed in the TPC. If a cosmic ray passes through the CRT and then near the TPC anode well before the beam spill we will see the activity in the cosmic ray taggers but the TPC signals will have been induced on the anode before the designated readout window. The same occurs with cosmics near the cathode well after the beam spill. However, the TPC readout is centered after the beam and therefore creates this uneven shape.

6.3 Training the CRUMBS BDT

The CRUMBS BDT is trained using the TMVA toolkit [258] on a sample of both neutrino and cosmic 'triggered' SBND Monte Carlo events. All events contain the full out-of-time cosmic flux regardless of their trigger origin and the neutrino samples are evenly balanced between muon and electron type neutrino events. In total 100,000 muon neutrino induced events, 100,000 electron neutrino induced events

and 50,000 in-time cosmic events are used with half forming the training sample and half a test sample. Each slice is matched to the true particles responsible for the charge depositions it contains, and the completeness and purity of the neutrino slices are assessed as

$$\text{Completeness} = \frac{\text{Number of hits produced by the neutrino interaction products contained in the slice}}{\text{Total number of hits produced by the neutrino interaction products}} \tag{6.1}$$

and

$$\text{Purity} = \frac{\text{Number of hits produced by the neutrino interaction products contained in the slice}}{\text{Total number of hits contained in the slice}}. \tag{6.2}$$

For the purposes of training the BDT, the signal category is defined as slices produced by a neutrino interaction with at least 80 % completeness and purity. This is in order to ensure that poorly reconstructed slices that do not well reflect the label of 'neutrino' or 'cosmic' are excluded from the training to avoid the BDT learning characteristics of mis-reconstruction. Such slices *are* considered in the validation of the tool's performance. The background category contains all slices that were matched to cosmic activity.

It is usually important to avoid much correlation between the input variables for a BDT. By definition, highly correlated variables tend to be providing very similar information to the BDT, and although the BDT will discover this correlation it will not do so in an efficient way. This increases the complexity of the trees for no performance gain. Figure 6.6 shows the correlation matrices for the input variables for the version of the CRUMBS BDT used in the selection in chapter 7. The clearest area of correlation is around the CRT variables. This is predominantly driven by the fact that most slices will not have any CRT matching information so these variables will be regularly set to default values. Many of the cosmic backgrounds that would create a CRT match are removed before CRUMBS during Pandora's unambiguous removal, whilst most neutrino slices will not create a CRT match due to being fully contained in the TPC. Other areas of higher correlation are focussed around some of the Pandora geometric variables, such as the number of space points

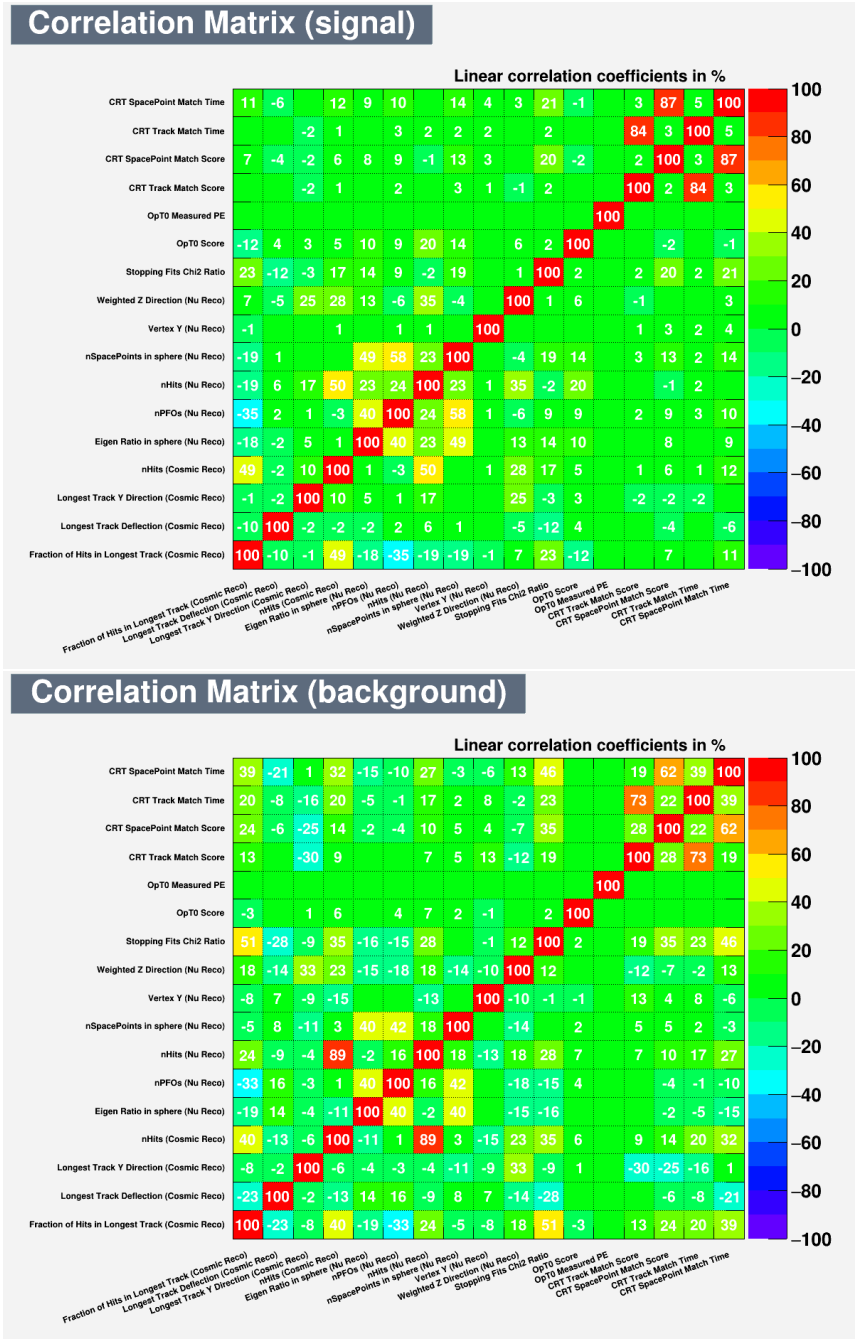


Figure 6.6: The correlation matrices (signal and background) for the variables used to train the version of the CRUMBS BDT used in the NC π^0 selection.

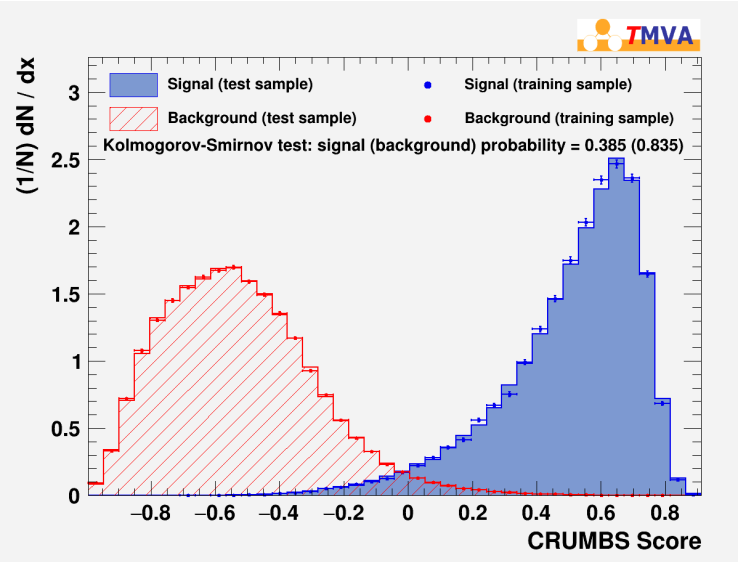


Figure 6.7: The output CRUMBS score for signal and background distributions in both the training and testing samples.

in the entire slice under the neutrino hypothesis or in the largest track under the cosmic hypothesis. This correlation is stronger in background slices, where a single cosmic-ray muon track will contain the vast majority of the slice's space points. In general, the correlations are minimal and those that remain have particular benefit in identifying specific topologies.

The CRUMBS BDT is trained with an architecture of 100 individual decision trees, each with a maximum depth of 3 nodes and the resulting output distributions are shown in figure 6.7. The red distributions show the background cosmic distributions for both the training sample and the testing sample and the blue distributions show the same for the neutrino signal. It is immediately clear that there is very good separation between the signal and background slices and that the testing and training distributions are in good agreement. Two-sample Kolmogorov-Smirnov tests [259] are performed to compare the testing and training distributions for both the signal and background cases (again shown in figure 6.7). This helps to simultaneously check for poor agreement, indicated by very low values, or evidence of the distributions agreeing so well that they may not be statistically independent,

indicated by values close to 1. The former is often a sign of overtraining, a phenomenon usually caused by a lack of statistics or too complex a set of trees. It results in a BDT becoming overly tailored to the specific training sample, to the degree that natural statistical fluctuations in any other sample will result in significantly worse performance. In this scenario, the tests yield reasonable values away from the extremes of 0 and 1 for both the signal and background cases.

It is also instructive to examine the importance of each variable. The TMVA software calculates this importance by summing the nodes in which each variable is used, weighted by its separation power and the number of units assessed at that node. Table 6.1 shows the ranked importance of the variables in the same iteration that produced figure 6.7. Previous iterations of the tool yielded some variables with zero importance, an indicator that was useful in thinning the number of inputs for the final configuration. It is clear that the primary objective of the CRUMBS tool is being well achieved, with all three sub-systems represented in the first three variables. When training alternative versions utilising SimpleFlash in place of OpT0Finder, the flash match score variable remains the most significant variable throughout.

6.4 CRUMBS Performance

The score distribution in figure 6.7 demonstrated very significant separation between signal and background but to truly understand the performance of the CRUMBS tool it must be evaluated in a more realistic context than the training samples. This was done with significant detail for the first production version of CRUMBS and will be presented here. It should be noted that this version of CRUMBS was subtly different in configuration from the training cycle presented in the previous section but that the performance level has been consistent throughout.

Rank	Variable	Variable Importance
1	OpT0 Score	0.14320
2	CRT Track Match Time	0.08438
3	Number of Space Points in Longest Track (Cosmic Reco)	0.08341
4	Vertex Y (Nu Reco)	0.06963
5	Longest Track Y Direction (Cosmic Reco)	0.06862
6	OpT0 Measured PE	0.06415
7	Number of Space Points (Nu Reco)	0.06386
8	Weighted Z Direction (Nu Reco)	0.05976
9	CRT SpacePoint Match Time	0.05939
10	Fraction of Hits in Longest Track (Cosmic Reco)	0.05551
11	CRT SpacePoint Match Score	0.05101
12	Number of Space Points in Sphere (Nu Reco)	0.04509
13	Eigen Ratio in Sphere (Nu Reco)	0.03993
14	Stopping Fits Chi2 Ratio	0.03326
15	CRT Track Match Score	0.03244
16	Longest Track Deflection (Cosmic Reco)	0.02517
17	Number of Final State PFOs (Nu Reco)	0.02117

Table 6.1: The relative importance of each variable used in the CRUMBS BDT.

6.4.1 Cosmic Rejection and Neutrino Efficiency

The following analysis compares the effect of a set of traditional cuts against the single CRUMBS score cut. The traditional cuts were taken from the most recent SBND Monte Carlo CC ν_μ selection [260], utilising variables from the TPC reconstruction and independent matches to the PDS and CRT reconstruction, and represent a realistic baseline to judge the improvements against. The cuts used for both workflows are shown in figure 6.8. The cut position for the CRUMBS score was tuned in order to closely match the level of in-time cosmic rejection provided by the traditional cuts, in order to assess the resulting increase in neutrino selection efficiency against a consistent baseline. This is demonstrated in figure 6.9 which shows cosmic rejection defined in two different ways. The left hand plot shows the rejection of in-time cosmic events in terms of the longest track in the event, requiring *all* slices in that event to fail the cosmic rejection cuts. This demonstrates the total removal of events with in-time cosmic induced triggers, at the level of $\sim 96\,\%$ for

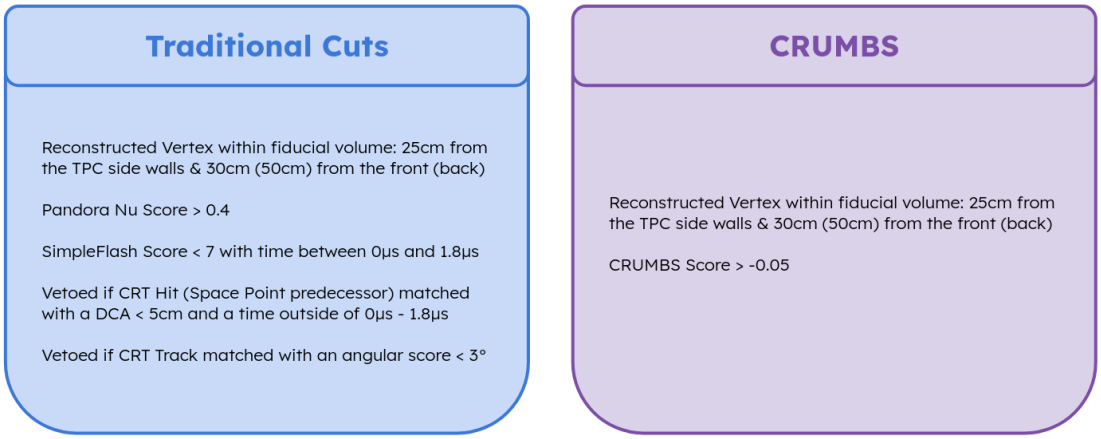


Figure 6.8: The two cut workflows used to compare the CRUMBS cosmic rejection performance to traditional SBND cut methods.

both methods, with the CRUMBS cut tuned to achieve this similarity. However, it is important to also consider out-of-time cosmic rejection, especially those present in events with a neutrino induced trigger. The right hand plot shows the rejection of out-of-time cosmic slices in terms of the longest track in the slice. This removal is extremely high at $> 99.8\,\%$ for both methods. Given initial samples will consist of an order of magnitude more cosmic slices than neutrino slices this high performance is critical to successful SBND physics.

Having evaluated the cosmic rejection power, the same sets of cuts can now be used to investigate the related effects on neutrino selection efficiency. Using exactly the same selections, figure 6.10 demonstrates the efficiency of retention for slices from three different neutrino interaction categories: the inclusive charged-current channels in ν_μ and ν_e as well as the NC $1\pi^0$ channel on which this thesis focuses. All three channels receive a very significant increase in efficiency when moving to the CRUMBS approach. This is most pronounced for CC ν_μ which is the most susceptible channel to cosmic cuts due to the fact that both CC ν_μ interactions and cosmic signals are dominated by the presence of a muon track. The increase in efficiency of $14\,\%$, from $80\,\%$ to $94\,\%$, will contribute to a considerable increase in

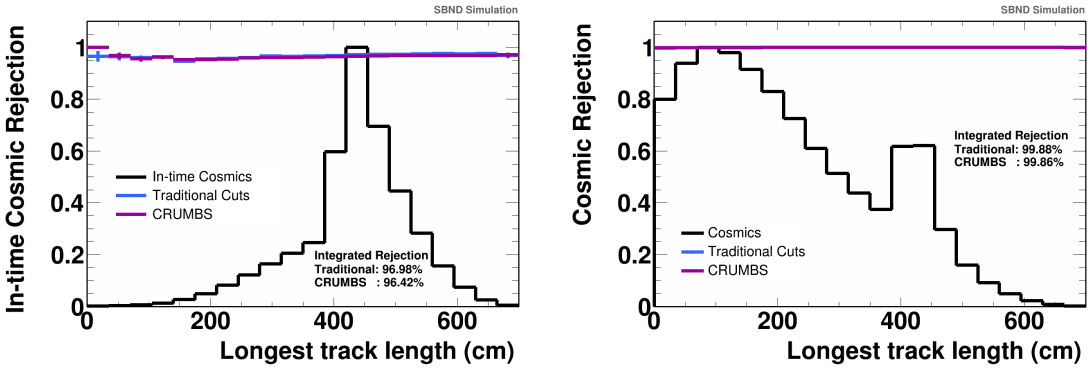


Figure 6.9: The left hand plot shows the proportion of in-time cosmic events for which all slices were rejected using the traditional set of cosmic rejection cuts compared to using the CRUMBS cut. The right hand plot shows the proportion of all cosmic slices rejected using the same sets of cuts. Overlaid on each plot (in the black curve) is the shape of the distribution of the variables before any cuts are applied.

the precision of a number of SBND analyses; particularly those that consider more exclusive CC ν_μ channels or use large double / triple differential spaces to make forensic assessments of cross section models. The efficiency boost for the shower-dominated channels, CC ν_e and NC $1\pi^0$, is somewhat less significant, at 6 % and 4 % respectively, but still represents a large material improvement in performance, all the more important in these lower statistics channels.

6.4.2 Phase Space Dependency Assessment

As with any machine learning tool, CRUMBS has potential for dependence on the modelling used to create the Monte Carlo samples on which it is trained. One way in which this risk can be assessed is by looking at the performance in different regions of the neutrino interaction phase space. If neutrino interaction slices from a certain phase space region are preferentially selected by the CRUMBS BDT then this indicates that an efficiency prediction created with Monte Carlo from one generator

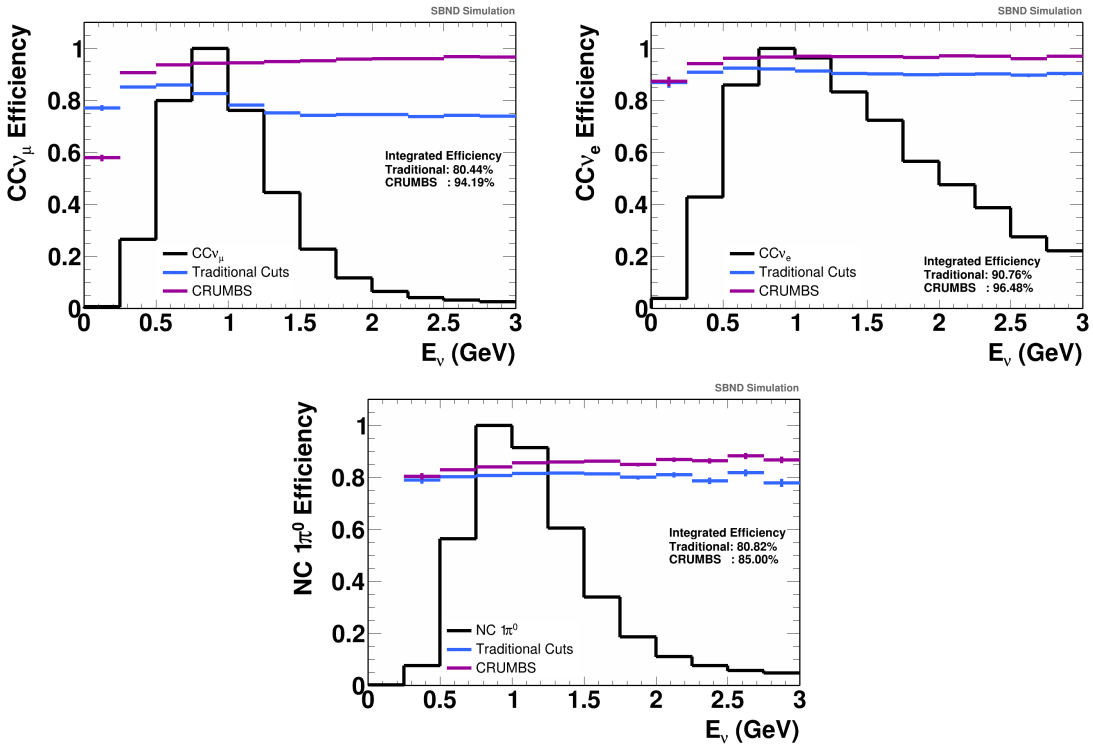


Figure 6.10: The proportion of neutrino slices retained after using the traditional set of cosmic rejection cuts and using the CRUMBS cut (top left: CC ν_μ interactions, top right: CC ν_e interactions and bottom: NC $1\pi^0$ interactions). Overlaid on each plot (in the black curve) is the shape of the distribution of the true neutrino energy before any cuts are applied.

or tune may be biased in a way that could obscure data discrepancy with this model.

Figure 6.11 shows the same CC ν_μ interaction slices subject to the same parallel sets of cuts used in the previous section but distributed across six different variables related to the kinematics of neutrino interactions: the energy of the initial neutrino (E_ν), the squared four-momentum transfer (Q^2), the Bjorken variable (X), the inelasticity (Y), the angle between the initial neutrino and the final lepton ($\theta_{\nu l}$) and the momentum of the final lepton (p_l). It is abundantly clear from these distributions that the selection efficiency of CRUMBS for these CC ν_μ events is very consistent across each of the variables, demonstrating an absence of bias towards events with particular kinematics. In fact, relative to the traditional selection, CRUMBS represents an improvement in this regard, with flatter selection efficiencies in certain regions such as moderate inelasticity.

Another related but not totally parallel approach is to look at the underlying modes of neutrino interactions, in this case coherent, quasi-elastic, meson exchange current, resonant and deep inelastic scattering. These categories relate to the variables assessed previously, particularly the four-momentum transfer, but critically are normally handled by distinct models within the generators. This, therefore, offers another angle to test for problematic model dependence in the CRUMBS tool response. The plots in figure 6.12 show the same CC ν_μ interactions in terms of the squared four-momentum transfer (Q^2) separated by their underlying interaction mode. The performance is very similar for all 5 modes providing another reassurance that the performance of CRUMBS is not inherently tied to the particulars of the generator models.

For the NC $1\pi^0$ channel, we also examine the efficiency improvements in the phase space of three other variables: the invariant mass of the hadronic system (W), the momentum of the neutral pion (p_{π^0}) and the cosine of the angle the pion's direction makes with respect to the beam axis ($cos(\theta_{\pi^0})$). These can be seen in figure 6.13. The first shows the characteristic peak of the dominant $\Delta(1232)$ resonance whilst

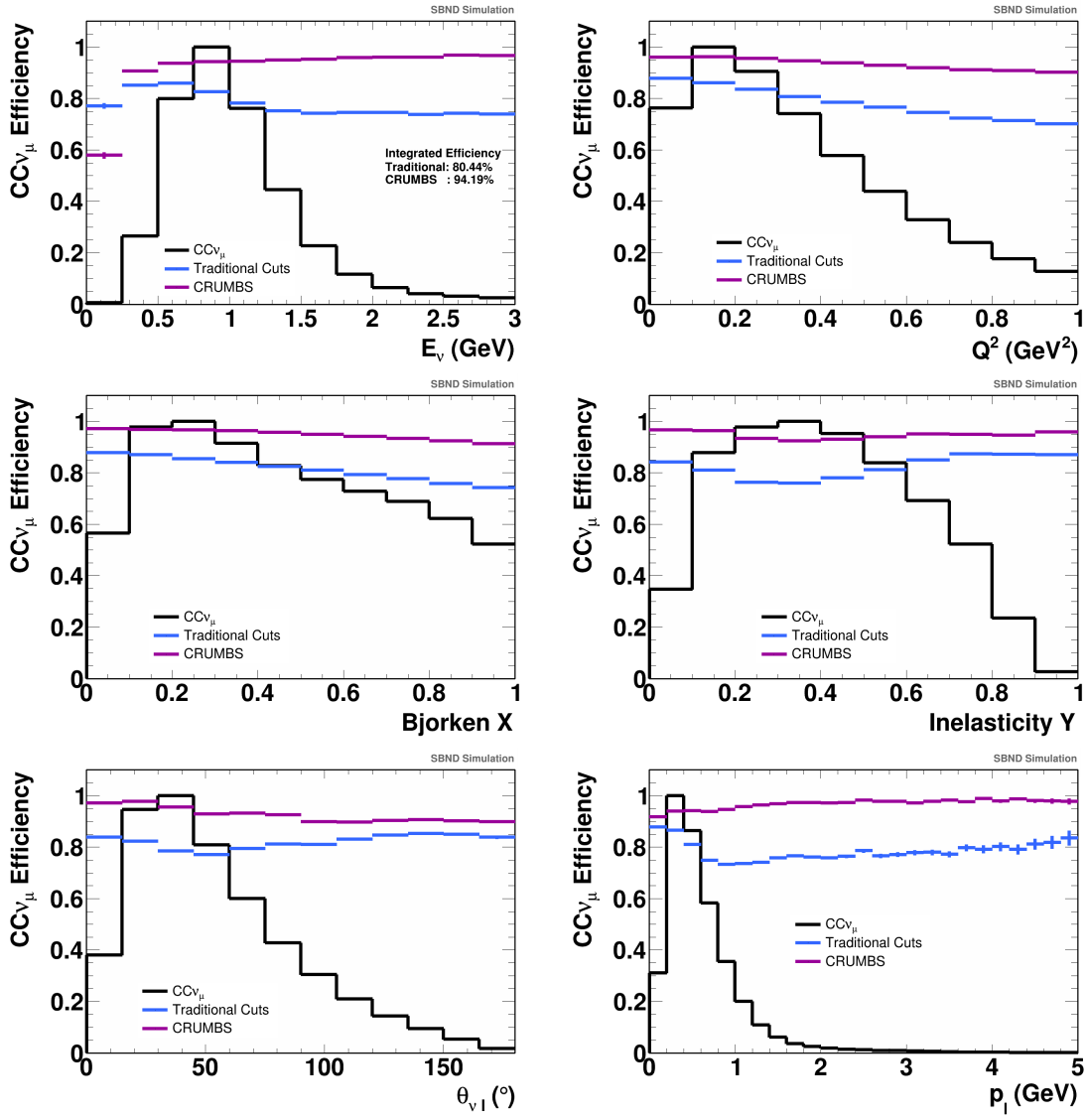


Figure 6.11: The proportion of CC ν_μ interaction slices retained after using the traditional set of cosmic rejection cuts and using the CRUMBS cut. From left to right, top to bottom the slices are distributed in terms of the energy of the initial neutrino (E_ν), the squared four-momentum transfer (Q^2), the Bjorken variable (X), the inelasticity (Y), the angle between the initial neutrino and the final lepton ($\theta_{\nu l}$) and the momentum of the final lepton (p_l). Overlaid on each plot (in the black curve) is the shape of the distribution of the variables before any cuts are applied.

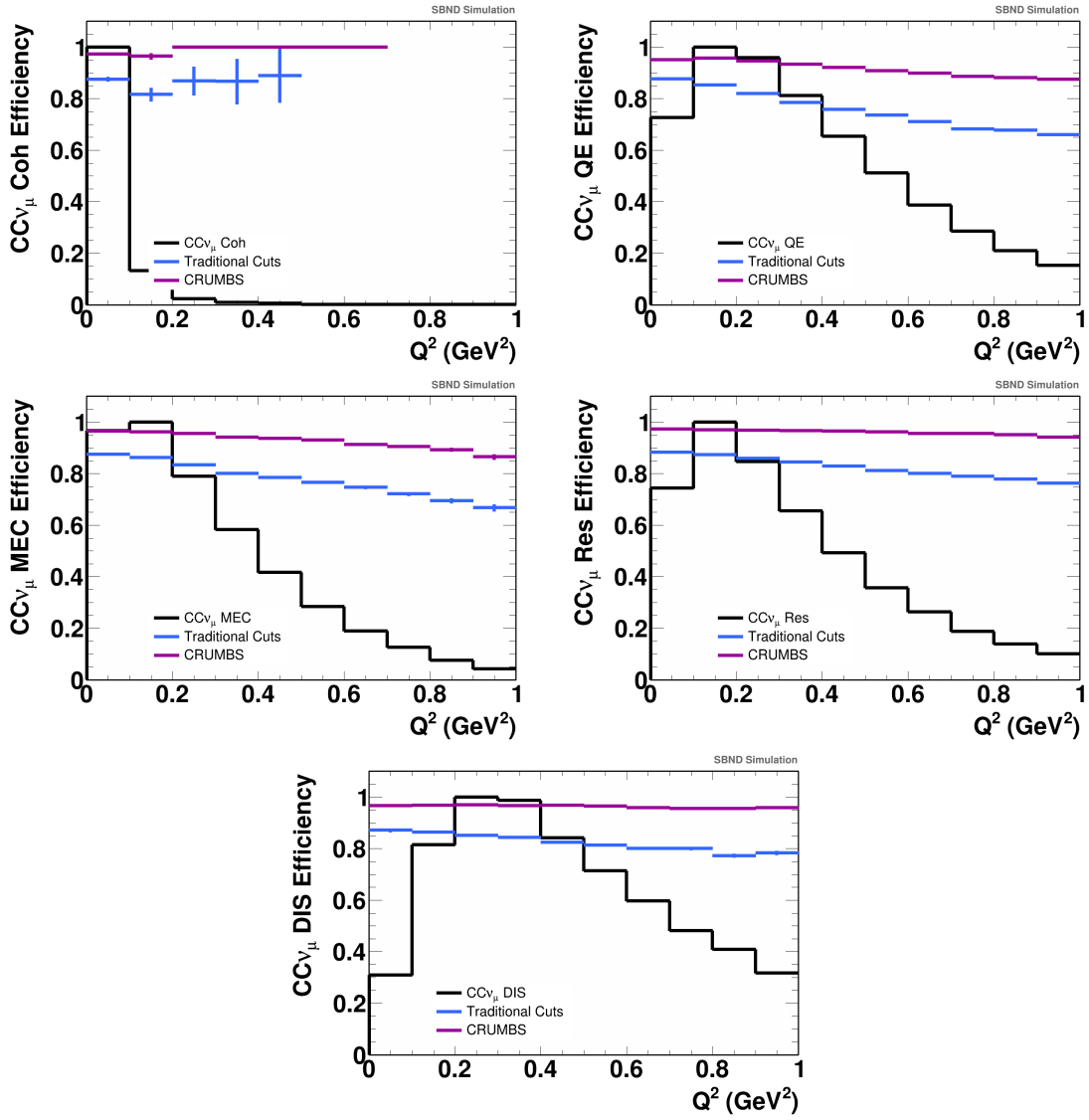


Figure 6.12: The proportion of CC ν_μ interaction slices of different underlying interaction modes retained after using the traditional set of cosmic rejection cuts and using the CRUMBS cut, distributed in terms of the squared four-momentum transfer (Q^2). From left to right, top to bottom the underlying interaction was coherent, quasi-elastic, meson exchange current, resonant and deep inelastic scattering. Overlaid on each plot (in the black curve) is the shape of the distribution of the four-momentum transfer before any cuts are applied.

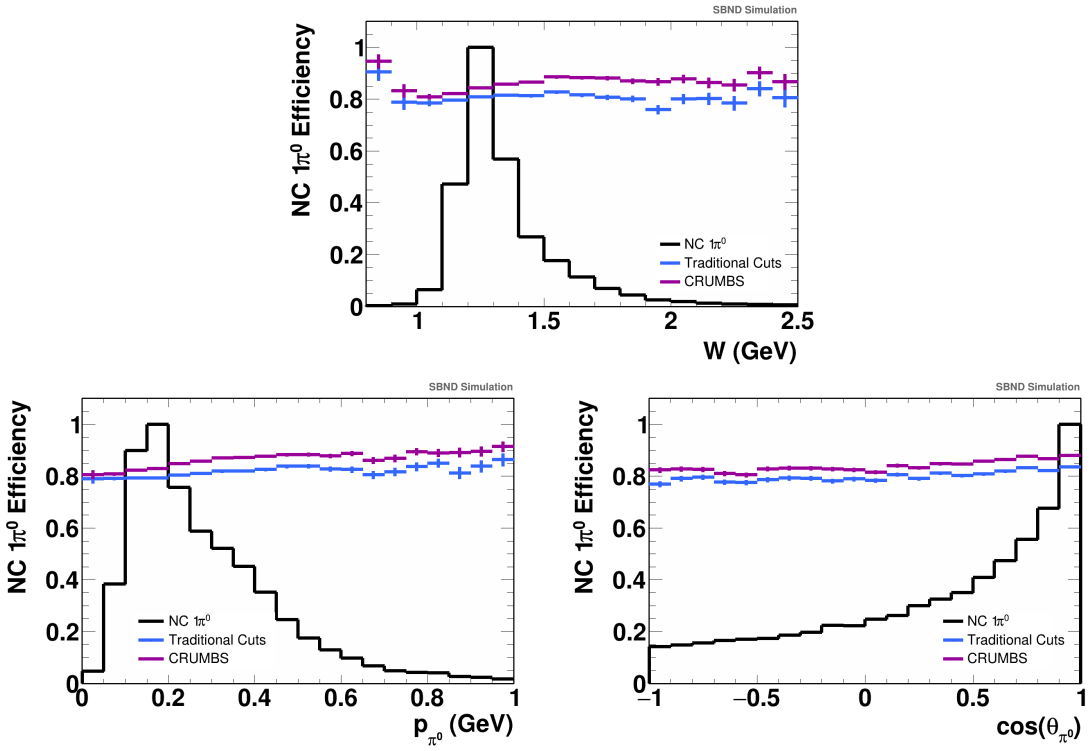


Figure 6.13: The proportion of NC $1\pi^0$ interaction slices retained after using the traditional set of cosmic rejection cuts and using the CRUMBS cut. The efficiency is expressed in terms of the hadronic invariant mass (top), the neutral pion's momentum (bottom left) and the cosine of the angle the pion's direction makes with respect to the beam axis (bottom right). Overlaid on each plot (in the black curve) is the shape of the distribution of the variables before any cuts are applied.

the other two are the observables used by the cross sections studies presented in chapter 8. Again, all three show consistent efficiency improvements across the phase space.

6.4.3 Full CC ν_μ Selection Context

To fully test the impacts of CRUMBS on a physics analysis, the context of a CC ν_μ inclusive selection was chosen. Firstly, this is the largest statistics channel

available to SBND and will form the basis of a huge number of analyses from exclusive cross section measurements to near detector constraints in ν_μ disappearance studies. Secondly, as discussed previously, CC ν_μ is the most vulnerable channel to cosmic rejection cuts due to the presence of a final state muon.

The inclusive CC ν_μ signal was defined to be all interactions in the fiducial volume of the form

$$\overset{(-)}{\nu}_\mu + \mathrm{Ar} \rightarrow \mu^\pm + X, \tag{6.3}$$

where X can consist of any number of final state nucleons and mesons.

A very simple CC ν_μ selection was designed using the cosmic rejection cuts from the previous section (see figure 6.8). Both selections then used a pair of cuts to affirm the presence of the final state muon: firstly that the slice contained at least one track-like PFO and then secondly that at least one of those track-like PFOs was identified as a muon candidate by the Dazzle tool. Dazzle is a particle identification (PID) tool that uses a multi-class BDT to identify tracks as either muon, charged pion or proton candidates. It will be discussed further in chapter 7. The respective results are shown in figure 6.14 indicating a vast improvement in the final selection efficiency of 13 % is possible, with the purity remaining consistently around 87 % for the two selections. For the planned full SBND exposure of 1×10^{21} POT this increase in efficiency corresponds to 500,000 more CC ν_μ events available to analyses. The efficiency curves also show that the gain in efficiency is most pronounced for longer tracks, indicating that by using CRUMBS a wider phase space is being retained in the selected sample.

6.4.4 Sub-system Contributions

The idea for the CRUMBS tool was to exploit the complementarity of the different SBND sub-systems and reduce the inefficiencies associated with the cosmic rejection cuts from each applied individually. This has been conclusively demonstrated.

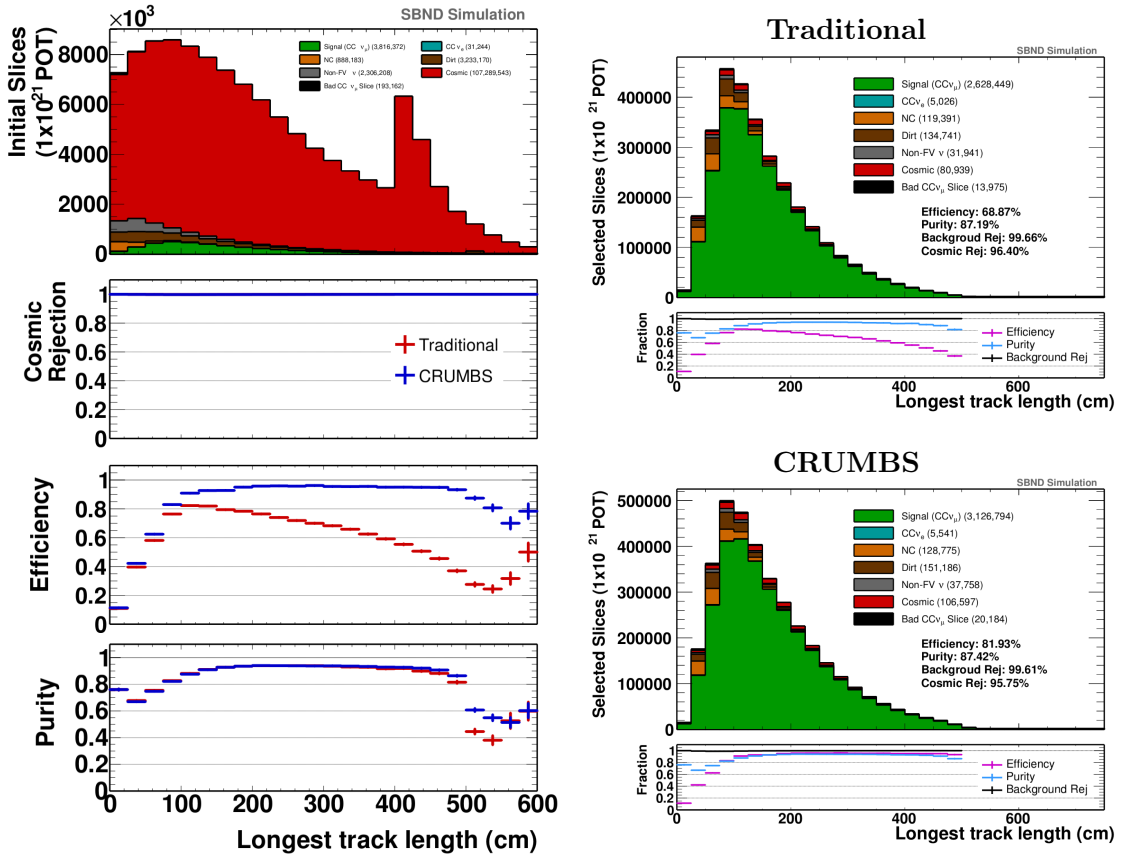


Figure 6.14: Alternative selections for CC ν_μ events using traditional cosmic rejection and CRUMBS, respectively. The left hand panel shows from top to bottom the initial slices before any selection cuts are applied, followed by comparisons of the cosmic rejection, signal efficiency and selection purity for the two approaches. The right hand panel shows the final selected samples for both approaches.

However, it is of interest to know the relative contributions of the different sub-systems. In order to achieve this, a series of retrainings of CRUMBS were performed, using the same training sample. Each permutation used the variables from a different combination of sub-systems - TPC only, PDS only, CRT only, TPC + PDS, TPC + CRT, PDS + CRT and TPC + PDS + CRT.

To evaluate the performance of each permutation, a receiver-operator-characteristics curve (ROC) was created, showing the relationship between the signal efficiency and background rejection as the cut value is moved. It nicely represents the compromise between the two concepts that is always at play when selecting a cut value, but also, via the area under the curve, demonstrates the difference in power between different variables, in this case different versions of the CRUMBS score. The efficiency and rejection are both defined with respect to the number of slices that pass some common pre-selection cuts (namely the existence of a Pandora neutrino slice in the fiducial volume), this means their absolute values are different to those used in the previous section, but provides a cleaner assessment of the effect of the CRUMBS cut. The samples used are identical in type but statistically independent of the training samples.

Figure 6.15 shows these ROCs for all 7 permutations. The solid orange line represents the nominal version of CRUMBS, and, as expected, is the most performative version as well. Looking at the dashed lines representing the individual sub-systems in isolation it is clear that the PDS provides the most separation, followed by the TPC, and that the CRT is the least consequential. This, and the fact that the CRT line follows a non-smooth trajectory, both result from the nature of the CRT variables which are set to default values for the many instances in which no CRT match was made. The right hand version of the plot includes only those events for which all of the variables were well defined. Whilst it is clear that the CRT variables are much less significant than those of the TPC or PDS, it is also clear that they do make some improvement to the overall picture when used in

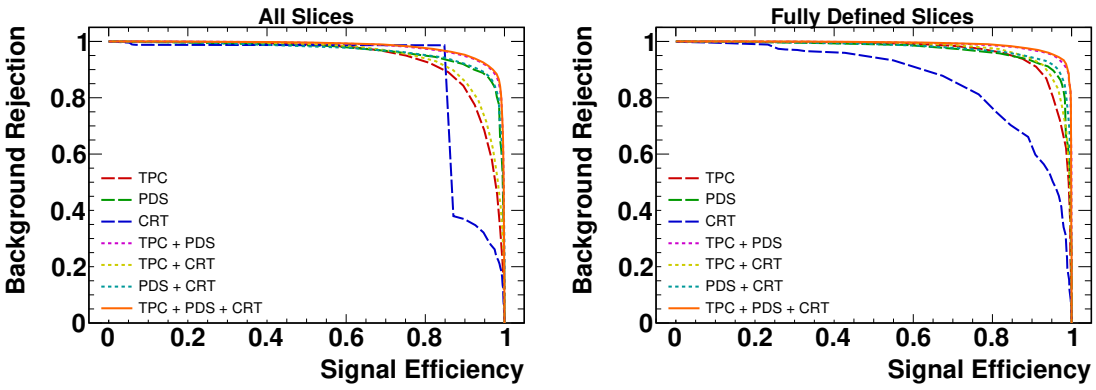


Figure 6.15: Receiver-operator-characteristics curves showing the behaviour of 7 permutations of the CRUMBS score using different combinations of sub-systems. The left hand version is for all slices that pass Pandora's unambiguous cosmic removal and a fiducial volume cut. The right hand version also requires that all the variables are well defined, meaning that matches were found using both the CRT and the PDS.

conjunction with the TPC and PDS. The fact that the PDS outperforms the TPC is initially surprising, but logical in the context of CRUMBS' position in the workflow. The Pandora unambiguous cosmic removal, occurring before CRUMBS, successfully tackles the cosmics for which the TPC information is most useful and so, at the point that CRUMBS is used, the PDS' high resolution timing is the most powerful discriminator against the remaining cosmics.

6.4.5 Signal Variations

Different classes of neutrino interactions have quite distinct spatial signatures. This is particularly true for the CC ν_μ and CC ν_e interactions due to the different ways that their charged leptons behave in the detector. The CRUMBS tool provides 4 flavours of the score for different analyses to employ: the original general score and then a score trained on each of just CC ν_μ slices, just CC ν_e slices and just NC slices. The background for each of the scores is trained purely on cosmic slices, the

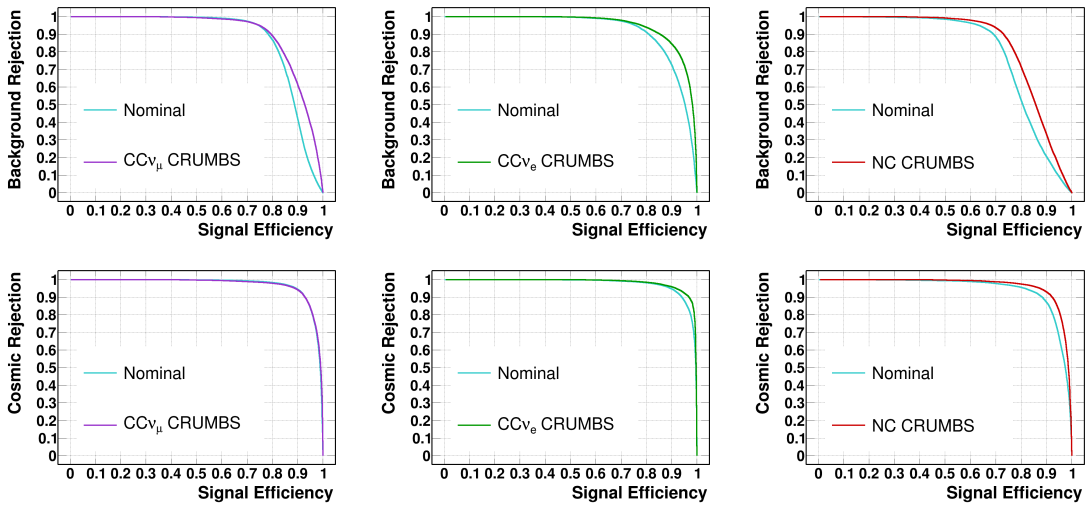


Figure 6.16: Receiver-operator-characteristics curves showing the behaviour of the three 'signal-specific' versions of CRUMBS, relative to the nominal (left CC ν_μ, central CC ν_e, right NC). The top row uses total backgrounds (including those from neutrinos) as the rejection metric, whilst the bottom row uses cosmic rejection.

exclusion of *other* neutrino slices is to avoid too much correlation between signal and background, particularly for the timing variables provided by the PDS and CRT. Figure 6.16 contains ROC curves for each of the 'signal-specific' scores compared to the nominal. The top and bottom rows separate out the effects on neutrino and cosmic backgrounds respectively. It is clear the the improvements using the CC ν_μ score are purely in the ability to throw away more CC ν_e and NC slices whilst the CC ν_e and NC scores both improve the cosmic rejection as well as rejecting CC ν_μ backgrounds. This is consistent with the ability of these scores to be 'harsher' on rejecting slices with any muon-like track without concern about removing signal slices. It is the NC-specific score that will go on to be used in section 7.3 for the cosmic rejection in the NC $1\pi^0$ selection due to its enhanced performance relative to the nominal CRUMBS score.

6.5 CRUMBS Conclusions and Development

The CRUMBS tool is demonstrably a significant improvement on the previous generation of cosmic rejection techniques for SBND. The sharp increases in efficiency for comparable cosmic rejection power are maintained across all categories of neutrino interactions. A wide validation campaign has indicated the consistency and breadth of its performance on SBND Monte Carlo, with strong indications of an absence of model dependency in its response to different areas of the accessible neutrino interaction phase space. Since the studies presented here were completed, GiBUU has been integrated as a second available generator for SBND. A valuable study that should soon be carried out would be to assess the response of the GENIE-trained CRUMBS tool to GiBUU events. This is something the author feels would help to further understanding of any areas in which CRUMBS could be susceptible to model dependence.

As SBND moves towards physics quality data taking, it will become possible to begin performing the necessary steps to use the CRUMBS approach in data analyses. The first step in this process will be the validation of the data-MC agreement for each of the input variables, as well as for other cosmic sensitive variables that could form part of an adapted configuration. It is clear from the results presented here that the ability to combine the complementary responses of SBND's three powerful sub-systems to cosmic backgrounds offers a vast reduction in the inefficiencies associated with each system alone.

Chapter 7

Selecting NC $1\pi^0$ Events

This chapter outlines a procedure for selecting NC $1\pi^0$ events from an SBND Monte Carlo sample. It was developed entirely by myself using inputs from the various reconstruction tools described in chapters 5 and 6 as well as particle identification methods discussed during this chapter. The latter builds on significant work performed by Edward Tyley in the context of his thesis analysis of a CC ν_e selection.

7.1 Signal Definition

The ultimate goal of this thesis is to evaluate the potential performance of a future SBND measurement of the NC $1\pi^0$ cross section. The first stage of any particle physics analysis requires the events of interest, the signal, to be isolated from the overall dataset. The signal events in this scenario are defined as any that satisfy

$$\begin{aligned} \nu \text{ Ar} \longrightarrow \nu\ \pi^0\ \text{X}, \\ \hookrightarrow \gamma\ \gamma \end{aligned} \tag{7.1}$$

where X represents rest of the hadronic final state (excluding further π^0s). The decision to exclude further neutral pions was motivated by the difficulty in

Signal	π^0	p ($p_p > 400\,\mathrm{MeV/c}$)	$\pi^\pm$ ($p_{\pi^\pm} > 150\,\mathrm{MeV/c}$)
NC $1\pi^0$ "Inclusive"	1	0+	0+
Exclusive NC $1\pi^0 0p 0\pi^\pm$	1	0	0
Exclusive NC $1\pi^0 \mathrm{N}p 0\pi^\pm$	1	1+	0

Table 7.1: The requirements on the hadronic final states for the three NC $1\pi^0$ channels covered in this thesis.

reconstructing states with a large number of electromagnetic showers. Note that the neutral pion is required to decay via the 2γ channel as this is the signal searched for in the selection. This is the pion's leading decay mode with a branching fraction of 98.823 % [32].

The selection actually considers three channels in parallel; firstly an 'inclusive' channel in which X can consist of any combination of hadrons (other than further neutral pions), and a pair of 'exclusive' channels referred to as NC $1\pi^0 0p 0\pi^\pm$ and NC $1\pi^0 \mathrm{N}p 0\pi^\pm$. The exclusive channels both require that X contains no charged pions above a momentum threshold of 150 MeV/c as well as requiring that there are no (NC $1\pi^0 0p 0\pi^\pm$) or at least one (NC $1\pi^0 \mathrm{N}p 0\pi^\pm$) protons above 400 MeV/c. These requirements are summarised for simplicity in table 7.1. The momentum requirements reflect the threshold above which the reconstruction efficiency of these particles is relatively consistent. This is demonstrated in figure 7.1. These exclusive channels provide two assets to the study of the cross section. Firstly, their more tightly defined signatures increase the chances of higher purity selections which in turn reduce the systematic uncertainties associated with modelling the background contributions. Secondly, the simultaneous measurements of channels with small differences in hadronic final states is one manner in which we can study the impacts of final state interactions.

A fiducial volume (FV) requirement is also placed on the true interaction vertex to ensure that there is a reasonable chance that the event is reconstructable. This volume (shown in figure 7.2) is defined by removing 20 cm from the x and y

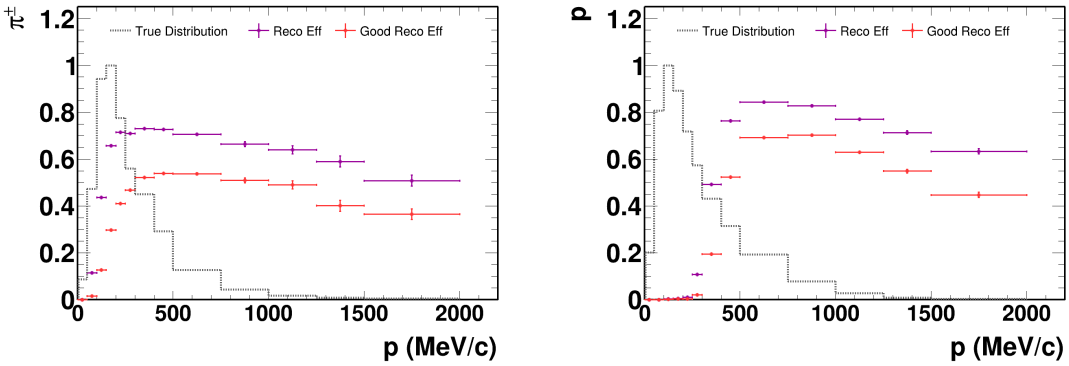


Figure 7.1: The reconstruction efficiency of charged pions (left) and protons (right) in SBND Monte Carlo as a function of their true momentum. The particle is deemed 'reconstructed' if there exists a reconstructed three-dimensional particle (PFO) with completeness and purity above 50% and 'well reconstructed' with completeness and purity above 80%.

detector edges, 10 cm from the upstream edge, 50 cm from the downstream edge and 5 cm from the central cathode. The asymmetry in the z cuts reflects the forward momentum of beam induced neutrino interactions.

7.2 Selection Procedure

As discussed extensively in previous chapters, a single SBND readout will contain a number of distinct physics 'events' resulting from both the neutrino beam and cosmic rays. The reconstructed slices produced by Pandora and used in the CRUMBS tool in chapter 6 represent isolated regions of reconstructed space and time, and therefore the selection considers every slice in the sample independently regardless of the properties of other slices in that detector readout. Note that due to SBND's proximity to the BNB target, the rate of neutrino pile-up, in which multiple neutrinos interact within the FV in the same beam spill, is about 1 in every 25 neutrino triggered events. This makes this effect non-negligible and further

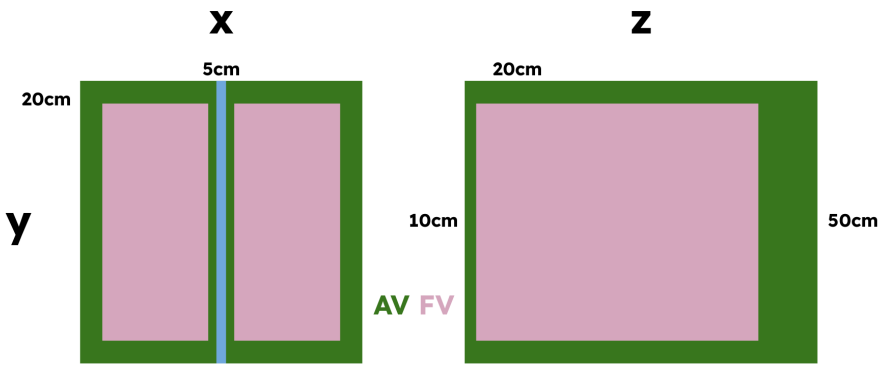


Figure 7.2: The fiducial volume (FV) considered in the selection visualised as cuts with respect to the full active volume (AV) in the yx and yz planes.

motivates the independent analysis of each slice in order to allow for multiple potential neutrinos to be considered in each event. Each slice can be placed into a true category by matching it to the simulated energy which resulted in the relevant signals. In order to be categorised as a signal slice, a requirement of 50% completeness is also applied. This requirement is designed to ensure double counting is avoided. Otherwise, in theory, multiple signal slices resulting from the same true interaction could be selected, invalidating the efficiency correction applied to the cross section in chapter 8. The signal and background truth categories assigned to slices are as follows:

- Signal
- Other NC $1\pi^0$ (only relevant for the two exclusive channels)
- Other NC
- CC ν_μ (in the inclusive channel this is split into those with a π^0 and those without)
- CC ν_e
- Dirt ν (slices resulting from neutrino interactions outside of the detector active volume)
- Non-FV ν (slices resulting from neutrino interactions within the active volume

but not the fiducial volume)

- Cosmic
- Mis-Reco Signal (slices matching to neutrino interactions satisfying the signal requirements but with less than 50 % completeness)

These categories can be appropriately covered via the use of two simulated samples. The first sample (denoted the neutrino sample) simulates neutrino interactions resulting from the BNB flux in a wide volume surrounding the detector, including an extended upstream volume. The continuous cosmic flux is also simulated throughout each of the readouts. A filter is applied to require that there is at least 100 MeV of true energy deposited in the active argon volume. This sample provides the signal NC $1\pi^0$ interactions as well as all varieties of neutrino background (including dirt interactions) and some cosmic backgrounds, including the potential for cosmic muons to overlap signal neutrino interactions in the TPC readout and cause reconstruction difficulties. This is of particular relevance to NC $1\pi^0$ interactions where a small isolated photon is more at risk of being attached to a nearby cosmic muon in reconstruction as a suspected delta-ray or Michel electron.

The second sample (denoted the in-time cosmic sample) simulates only the cosmic flux but with filters that require that a particle intersects the detector volume during the beam window, and at least 10 simulated photons are observed in the photodetection system. This sample provides the in-time cosmic background in which events are recorded due to a cosmic muon firing the trigger. Future iterations of SBND studies will include a full emulation of the trigger used to collect data in SBND operations, verified against a minimum bias trigger.

Combining these samples for analysis is non-trivial due to the need to calculate their relative normalisation. The neutrino sample contains a POT value representing the total integrated neutrino flux used by GENIE in order to produce this number of events. It also records the number of beam spills to which this corresponds, including

those for which no event was saved due to no neutrino interactions passing the filter requirements. The in-time cosmic sample also contains the total number of spill windows required to produce this number of in-time cosmic 'triggered' events. The neutrino sample can be scaled to the intended POT, and the number of spills scaled by the same amount. Using the approximation that each beam spill consists of 5×10^{12} POT the total number of spills required to reach the targeted POT can be calculated. This approximation is based on the existing data for the BNB which has run in a stable configuration for around 20 years. Beamline monitoring during SBND operations will remove the need for this approximation in future analyses. Finally, the targeted number of spills, removing those already accounted for by the neutrino sample, can be used to scale the in-time cosmic sample to account for the remaining spills.

A third independent sample of just NC $1\pi^0$ events is also used at points to evaluate signal-only efficiencies but never in conjunction with the other samples and, as such, does not require a similar scaling procedure.

Assuming a total SBND exposure of 1×10^{21} POT the nominal prediction results in ∼280,000 events satisfying the inclusive NC $1\pi^0$ definition and ∼145,000 and ∼105,000 satisfying the NC $1\pi^0 0p 0\pi^\pm$ and NC $1\pi^0 \mathrm{N}p 0\pi^\pm$ definitions respectively. These values reduce by ∼5-7 % when the slice matching procedure and 50 % completeness requirements are applied, resulting in 260,000, 135,000 and 100,000 signal slices at the start of each selection respectively. At this point the signal events are dwarfed by 4.6 million CC ν_μ slices, 3.5 million dirt neutrino slices, 2.1 million non-FV neutrino slices, 800,000 other neutrino slices and 185 million cosmic slices.

The NC $1\pi^0$ selection can be coarsely grouped into three stages: rejection of the plethora of cosmic-ray induced backgrounds, particle identification to isolate events with neutral pions from other neutrino interactions and remove any charged-current background, and finally the use of charge-light matching information to cleanly remove any remaining cosmics and poorly reconstructed slices.

7.3 Cosmic Rejection

The first stage of any neutrino analysis is to reject as many of the overwhelming number of cosmic slices as possible, in order to identify candidate neutrino interactions. Three cuts are used to achieve this:

- **Pandora Unambiguous Cosmic Removal** - Any slices that the Pandora TPC reconstruction labelled as unambiguously cosmic are removed.
- **Fiducial Volume** - Any slices with a reconstructed vertex position outside of the fiducial volume defined in figure 7.2 are removed - this targets non-fiducial neutrino backgrounds such as dirt muons as well as cosmic backgrounds.
- **CRUMBS** - The CRUMBS NC score described in chapter 6 is used with a cut value tuned to maximise efficiency $\times$ purity for each of the three selections separately.

Figure 7.3 shows the optimisation of the CRUMBS cut for each of the three channels to ensure the maximisation of efficiency $\times$ purity (equivalent to minimising the statistical error). The combination of these three cuts are able to retain a signal efficiency above 85 % for all three selections, whilst reducing the cosmic backgrounds by more than 99.85 %. For the NC $1\pi^0$N$p0\pi^{\pm}$ channel this rises to 99.95 % as the cut can be more strict whilst retaining the high signal efficiency. This is due to the presence of guaranteed activity at the vertex, making the signal easier to reconstruct.

Figure 7.4 shows the distribution of slices according to their fiducial status and the CRUMBS NC score to demonstrate the impact of the cosmic rejection cuts. The unambiguous cosmic removal has *already* reduced the cosmic-ray backgrounds by a factor of almost 10. It is clear from these figures that the two different cuts successfully target different topologies. The dirt and non-FV neutrino slices have signatures that will predominantly look more like neutrino interactions than cosmic

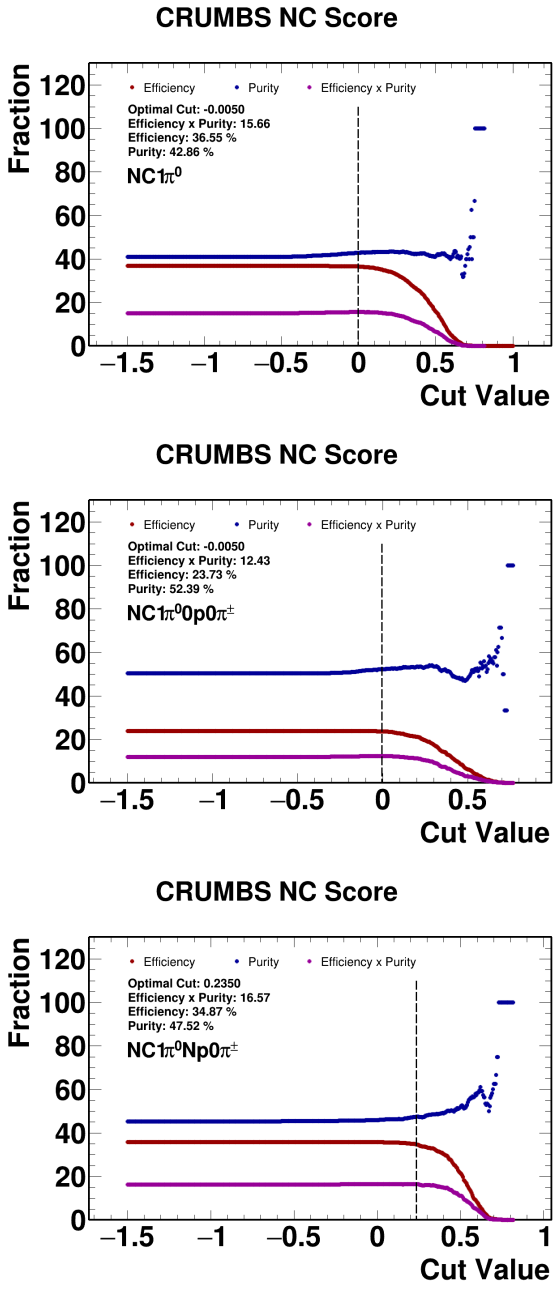


Figure 7.3: Efficiency, purity and efficiency × purity of the selections (top: NC $1\pi^0$, centre: NC $1\pi^0 0p 0\pi^\pm$, bottom: NC $1\pi^0 \mathrm{N}p 0\pi^\pm$) as a function of CRUMBS NC score cut value.

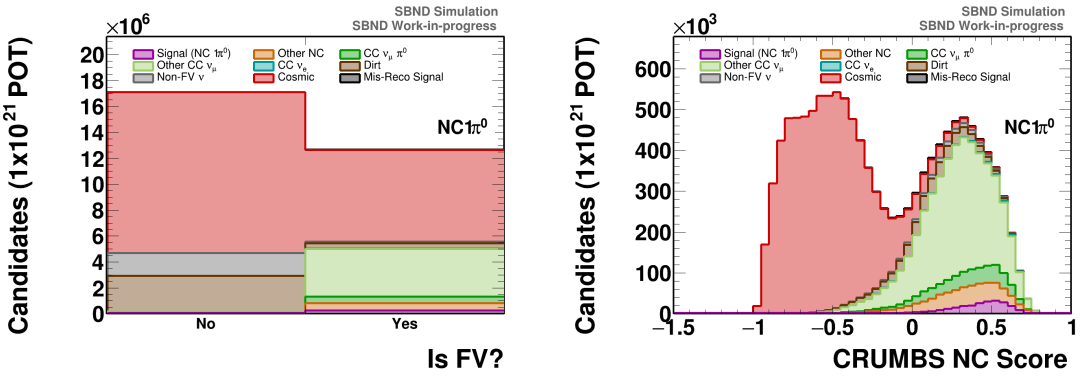


Figure 7.4: Left: the classification of slices as either fiducialised or not, following the unambiguous cosmic removal. Right: the CRUMBS NC Score distribution for the NC $1\pi^0$ selection after the unambiguous cosmic removal and the fiducial volume cut.

activity. The fiducial volume cut deals very well with these backgrounds, reducing them by 88 % and 94 % respectively. The cosmic slices are instead targeted with the CRUMBS NC score where they are heavily piled to the left hand side of the distribution and are almost entirely removed via the CRUMBS cut. In total over 99.8 % of all cosmic slices are removed by the three cuts described here, ensuring they are no longer the most significant background at this point.

7.4 Particle Identification

At this stage of the selection, the remaining slices are predominantly of neutrino origin, with CC ν_μ interactions forming the largest part of this by quite a margin. Successfully isolating the signal events from this pool of neutrino interactions requires identifying the differing signatures of the particles produced. The ability to successfully identify different particle types and therefore select different complex final state topologies is something that sets LArTPCs apart from many of the previous generation of experimental setups.

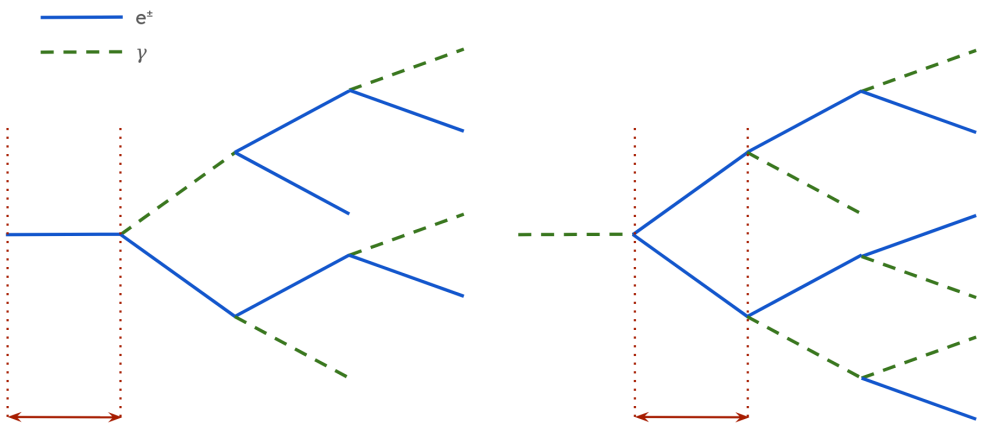


Figure 7.5: Simplified schematics of electromagnetic showers induced by electrons/positrons (left) and photons (right). The red zones show the first visible sections of the showers.

7.4.1 Identifying Features

It is important to understand the physical reasons that might identify a certain type of particle's signature in a LArTPC. In section 4.3.1 we discussed that, at the energies accessible to SBND, only electrons and photons are relativistic enough to undergo radiative processes and create showers of particles. The left panel of figure 7.6 shows that, for electrons, radiative effects become the dominant mode of energy loss between 10 MeV and 100 MeV. Broadly classifying a reconstructed particle as a track or shower is usually the first step to identifying the particle type.

There are two features that are then usually used to distinguish between electron (or positron) and photon induced showers. Figure 7.5 shows simplified schematics of electromagnetic showers beginning to develop via bremsstrahlung, pair-production and Compton scattering from an initial electron or photon. The first key difference between the two is the point at which the shower becomes visible. The electron shower is immediately visible due to the fact that the primary particle itself is charged, whereas the photon shower is only visible once the photon has undergone

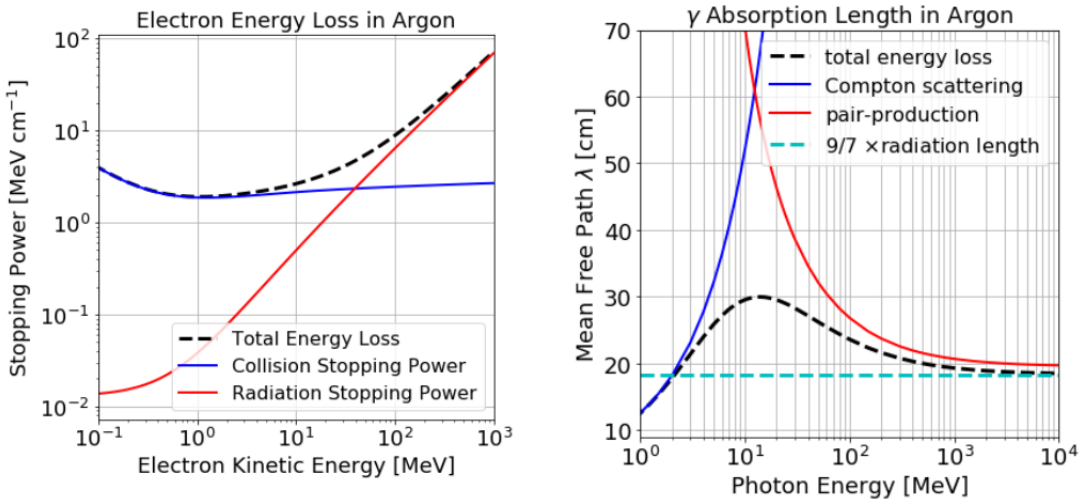


Figure 7.6: EM particle behaviour in liquid argon. Left: the relative stopping power of direct and radiative effects for electrons as a function of their kinetic energy. Right: the mean free path of photons with separate contributions from Compton scattering and pair-production as a function of their kinetic energy. Figure from [261].

pair-production or Compton scattering. If the photon's production point is known, usually due to visible activity at the neutrino interaction point, then a 'conversion gap' before the start of a shower is a good indicator it was produced by a photon. The scale of this gap is characterised by the photon's radiation length in liquid argon, 14.1 cm.

Figure 7.6's right hand panel shows that pair-production becomes the leading effect around 10 MeV and by 100 MeV is dominant. This means that most photon showers at SBND energies initiate via this process and so the first visible sections of the showers, the red regions of figure 7.5, will consist of different levels of ionisation. The beginning of the electron shower should deposit energy at a rate consistent with a single minimum ionising particle (MIP), whilst the beginning of a photon shower, due to the overlapping electron and positron, should deposit energy at roughly double this rate. A measurement of $\mathrm{d}E/\mathrm{d}x$ in the first few cm's of a shower, referred to as the track stub, can be used to try and identify this difference.

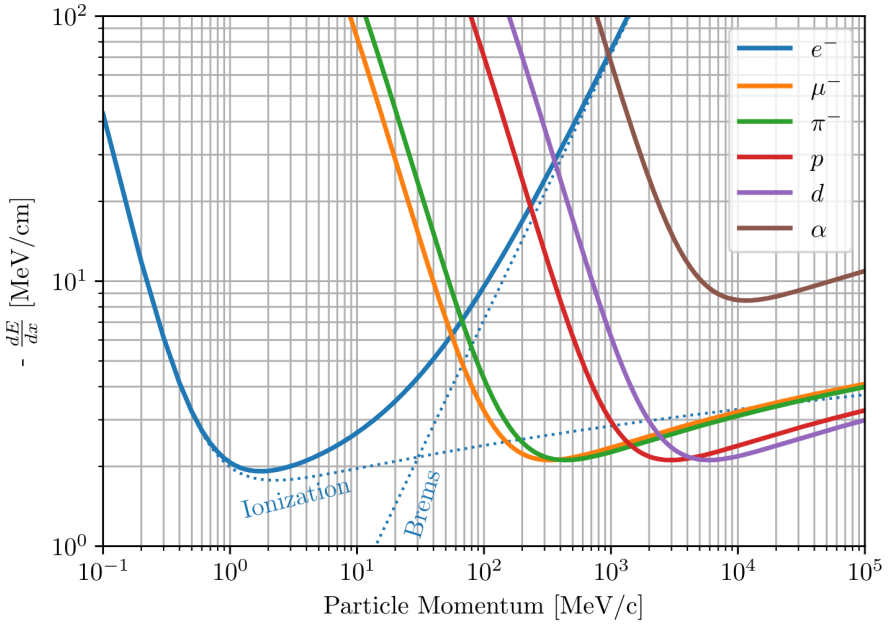


Figure 7.7: The energy deposition rates of different particles in argon as a function of their momentum. Figure from [262].

The heavier particles, muons, charged pions and protons, are less relativistic at these energies and tend to deposit energy directly through collisions, creating track-like topologies. As with showers there are distinctive features of each type used to try and identify them. The first is their relative ionisation strengths; due to their similar masses, both muons and charged pions deposit energy at MIP-like rates, whilst protons are much heavier, highly ionising particles. This is illustrated via figure 7.7 which shows that for a given momentum in the region typical of few-GeV neutrino interactions ($p \sim$100 MeV-1 GeV), the muon and pion depositions are very similar whilst the proton deposition is much larger.

The scattering along the length of the track also provides two indicators. Firstly, all charged particles will undergo multiple coulomb scattering (MCS) as they propagate through the argon. This scattering is momentum dependent and, as such, the average scattering 'wiggle' is expected to be largest for muons and smallest for protons. Secondly, whilst protons and pions are expected to have less small MCS,

they are susceptible to hard hadronic scatters resulting in much larger angular deflections. There is still separation between the two for, given the same momentum, the pion's average scattering angle will be less than that of the proton due to the dependence on particle velocity [263]. Finally, decay products can distinguish muons and pions. If the reconstruction identifies a decay into a new particle then a smaller, shower-like particle indicates a Michel electron from a muon decay whilst a larger track-like particle indicates a muon produced from the pion decay.

7.4.2 Existing Tools

Previous iterations of SBND MC selection studies have harnessed and developed a series of useful particle identification tools, all exploiting the particles' various characteristics outlined in the previous section. The most recent development was the combination of some of these tools into a pair of boosted decision trees: Razzle and Dazzle. Razzle tackles PFOs that Pandora determined to be shower-like and Dazzle those it determined to be track-like [245]. The selection presented in this chapter will use a combined version of these tools, but first we consider them independently.

7.4.2.1 Razzle

The Razzle BDT assesses any Pandora PFO labelled as shower-like using five representative variables. The BDT uses a multi-class approach, defining a series of categories rather than just the traditional signal and background labelling. A score is produced for each category representing how much the input variables conform to that label. The scores sum to unity, and as such can be treated comparably as 'probabilities' of belonging to each category. In the case of Razzle, the categories are Electron, Photon and Other, the latter mainly consisting of mis-classified 'track-like' particles. The algorithm's output for each shower contains a PDG code indicating

the most likely category for it to belong to, as well as the three BDT scores to allow more flexible cuts for different selections.

Razzle uses the following input variables:

- The $\mathrm{d}E/\mathrm{d}x$ for the initial track stub, calculated for the readout plane with the most hits in this region.
- The gap between the reconstructed neutrino interaction point and the start of the shower in question. In the case of a photon-only signal (such as that in the NC $1\pi^0 0p 0\pi^\pm$ channel) this handle does not offer any e/γ separation.
- The opening angle of the shower, calculated from its length and width.
- The density of hits per wire, corrected for the effective pitch (angle of the shower with respect to the wire plane).
- The density of energy along the shower, calculated by dividing the shower energy by its length squared.

The first two variables access the key distinguishing features between electron and photon induced electromagnetic showers detailed in section 7.4.1. The following three intend to distinguish objects that may be mis-classified track type particles. Tracks should be narrower and less dense in hits and energy. The final variable also offers some sensitivity to electron photon separation due to the increased energy density of photon showers. Figure 7.8 shows each of the output PID distributions for recently updated Razzle training and testing samples. It is clear that there is good separation in each of the distributions. The Other category is clearly more distinct than the Electron and Photon categories which is logical, given they do not result from genuine electromagnetic showers. It is worth noting that the Razzle tool is trained only on showers of 100 MeV or more in order to increase the reliability of the input variables. Given the typical energies of π^0 decay photons regularly sit

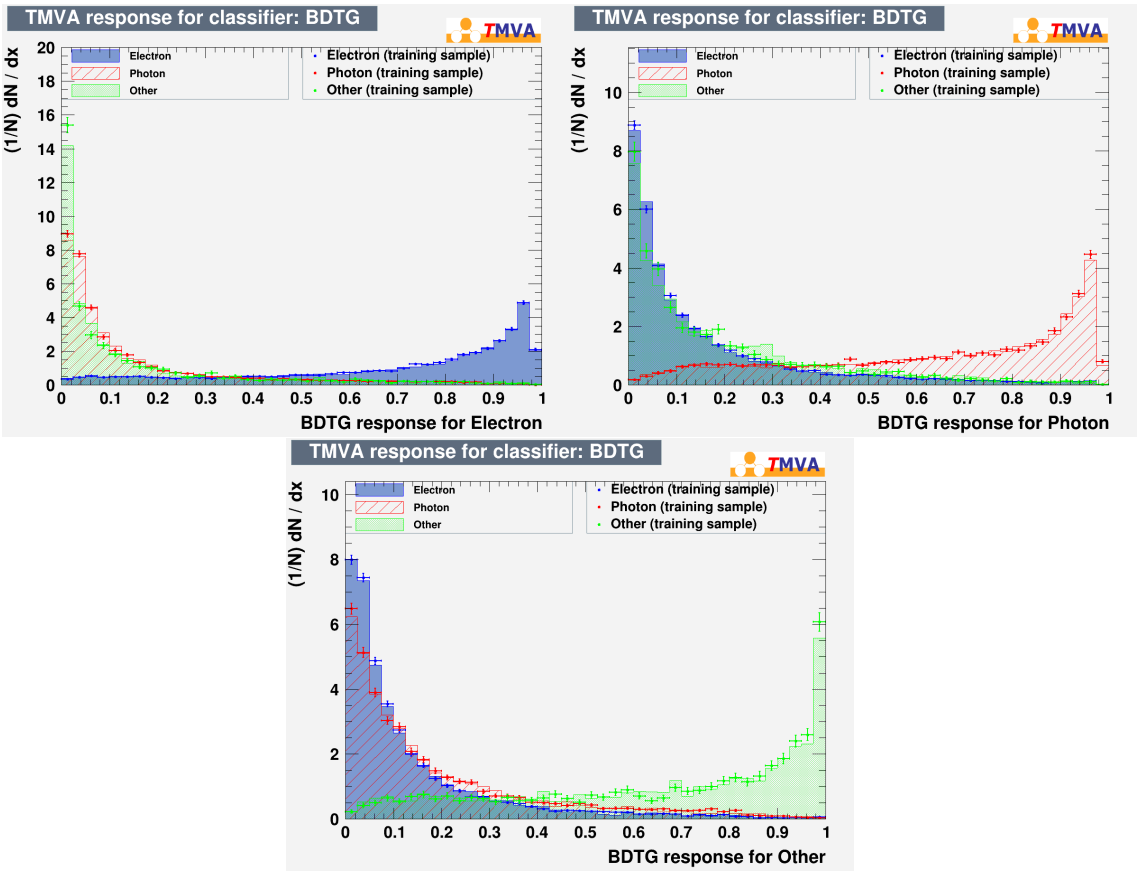


Figure 7.8: The distributions for the Razzle output scores (top left: Electron, top right: Photon, bottom: Other) for both training and testing samples.

below this value, as shown in figure 7.9, this provides a challenge to π^0 identification with Razzle.

7.4.2.2 Dazzle

As with Razzle, Dazzle uses a multi-class approach to assess track-like particles, resulting in an overall classification alongside scores for the categories of Muon, Charged Pion, Proton and Other. The Dazzle BDT assesses any Pandora PFO labelled as track-like using the following twelve variables:

- The fitted track length.

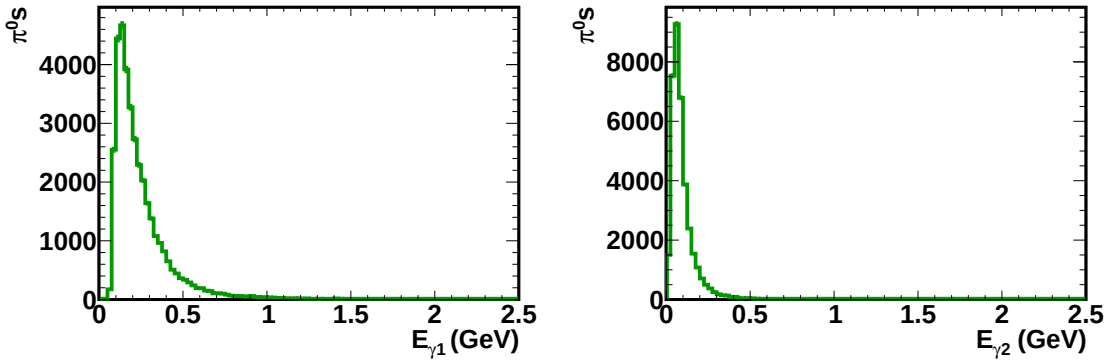


Figure 7.9: The energy of leading (left) and sub-leading (right) photons resulting from neutrino induced π^0 decays in the SBND fiducial volume.

- The χ^2 agreements for the track's $\mathrm{d}E/\mathrm{d}x$ as a function of residual range compared to theoretical distributions for muon, charged pion and proton hypotheses. The muon and proton values are inputted directly whilst the charged pion score is inputted as a difference with respect to the muon score. This is due to the close similarity in the energy deposition associated with both MIP-like particles.

- The mean scattering angle calculated from a tool that evaluates the many small scatters associated with MCS.

- The ratio of the maximum scattering angle to the mean angle from the MCS tool.

- The mean distance of closest approach between the fitted track points and the straight line between the track's start and end. This is a further assessment of the larger MCS effects expected in muons.

- A tool to identify Bragg peaks performs fits of a 0th-order polynomial and exponential curve to the track's $\mathrm{d}E/\mathrm{d}x$ residual range distribution (the same tool described and used in section 6.2). The input variables are the ratio between the χ^2 values for each fit (indicator of the existence of a Bragg peak)

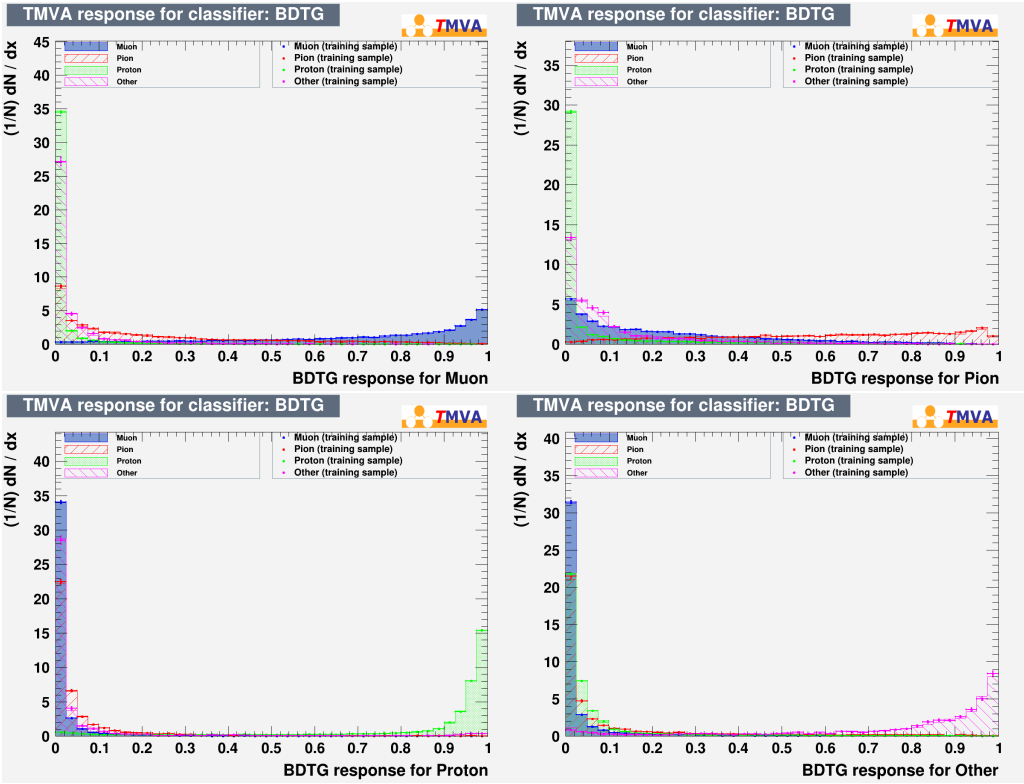


Figure 7.10: The distributions for the Dazzle output scores (left top: Muon, right top: Charged Pion, left bottom: Proton, right bottom: Other) for both training and testing samples.

and the constant coefficient from the polynomial fit (good representation of the average $\mathrm{d}E/\mathrm{d}x$ along the track).

- The difference between the muon hypothesis momentum calculated via two techniques: range and MCS. These should agree well for contained muon tracks and poorly for all other categories.

- The number of children this PFO has and the number of hits in the largest child PFO.

The Dazzle score distributions are shown in figure 7.10 and clearly indicate that the Proton category is the most separated from all the alternatives, followed by Other and then Muon and Charged Pion. This is consistent with the fact that, due to

the protons being significantly more ionising than charged pions and muons, they exhibit much more distinct differences than between the two lower ionising particles. As the proton momentum increases the chances of inelastic scattering processes also increase, thus reducing the chance of a clear Bragg peak and making them harder to differentiate from charged pions.

7.4.3 Razzled

As mentioned in section 5.2.1.4, SBND has recently implemented a dual-characterisation approach in which all PFOs, regardless of what label Pandora assigned them, undergo characterisation as both showers and tracks. It is left to each analysis to determine the best characterisation to use. The Razzle and Dazzle tools were developed at a time when only single characterisation was applied. In order to make best use of the available information the author constructed a new BDT tool, imaginatively named Razzled, combining the variables from each of its predecessors.

The Razzled BDT takes a single PFO as its unit input, creating the Razzle and Dazzle input variables from the shower and track characterisations respectively. The final Razzled input variable is the Pandora TrackScore which is itself obtained from a BDT and describes how much the PFO resembles a track or shower (see section 5.2.1.2). The intention of this combined tool is to regain identification inefficiencies near the track-shower boundary. This is particularly prevalent for lower energy electromagnetic particles such as the sub-leading photon in π^0 decay as they often sit below the highly relativistic threshold at which point radiative effects become dominant, and thus look like small high scatter tracks (see figure 7.11 to see how increasing energy changes the geometric and therefore reconstructed appearance).

Razzled also lowers the shower energy and track length cuts used in Razzle and Dazzle respectively, in order to attempt to better cover the phase space in which

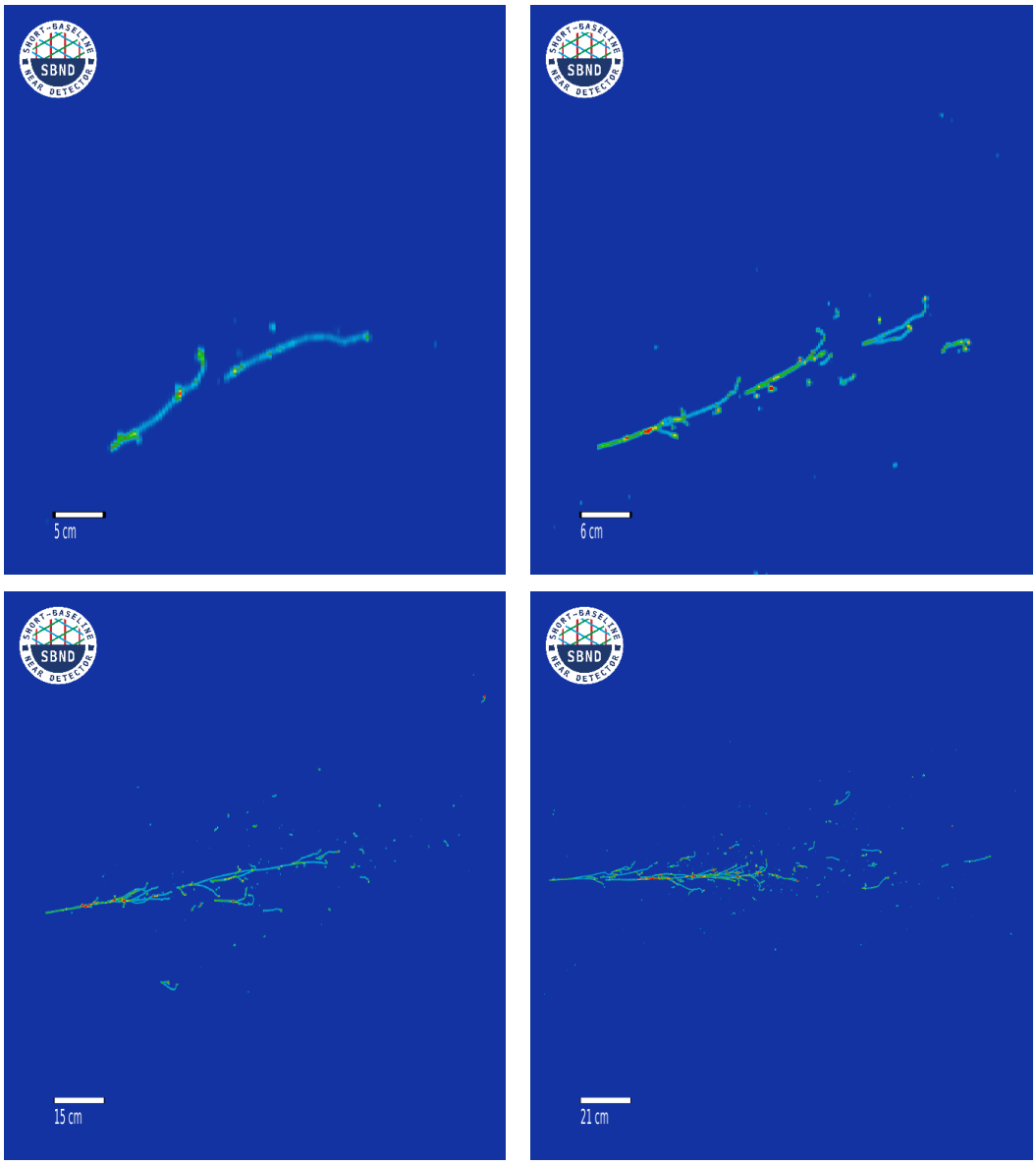


Figure 7.11: Sections of SBND Monte Carlo event displays showing simulated photons of increasing energy left to right, top to bottom.

the NC π^0 events exist. The shower energy cut is lowered to 10 MeV and the track length to 3 cm. The former should ensure many more lower energy photons from π^0 decays are encapsulated in the training set, whilst the latter should ensure that charged pions and protons of momentum large enough to be counted in the exclusive signal definitions are also covered. Unlike in the two separate tools there is no Other category as this mainly covered particles placed in the alternative characterisation, a concept that does not exist once both are run for all PFOs. As such, there are five output scores: Electron, Muon, Photon, Pion and Proton, all shown in figure 7.12.

The performance of the Razzled particle identification, as will be used in the final selection, is well summarised via the confusion matrices shown in figure 7.13. The efficiency matrix shows that, in general, most true particles of each class were identified as the correct type by Razzled The Proton class was the best performing at 88 % and Charged Pion the least at 70 %. The largest areas of confusion are in the Muon-Pion quadrant and the Electron-Photon quadrant. This is consistent with the all the previous discussions concerning the similar properties of these pairs of particles.

The purity normalised matrix tells a somewhat more complicated story regarding the make-up of each Razzled class. The Proton, Electron and Muon classes are highly pure, a strong indication of the power of this tool for the majority of key signals for both oscillation and cross section analyses. The less pure classes of Photon and Charged Pion are both more commonly included, or excluded, in the signal definitions of exclusive cross section measurements such as the one pursued in this work. The largest confusion is in the Charged Pion class in which 44 % of the identified pions are in fact, muons. This primarily results from the very similar calorimetric behaviour of the two particles and represents a significant challenge for the liquid argon community requiring improved methods of separation. Much ongoing work in this area harnesses the power of image based machine learning algorithms such as convolutional neural networks to identify complex distinguishing

Figure 7.12: The distributions for the Razzled output scores (left top: Electron, right top: Muon, left middle: Photon, right middle: Pion, bottom: Proton) for both training and testing samples.

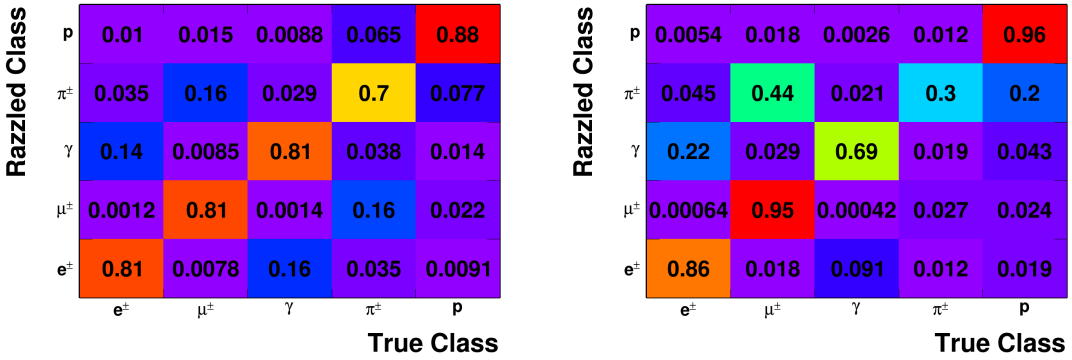


Figure 7.13: Confusion matrices showing the Razzled classification of PFOs matched to each class of true particle. Both contain the same sample but are normalised differently. The left version is column normalised so represents the efficiency of each Razzled class, whilst the right version is row normalised so represents the purity of each Razzled class.

features that have been difficult to encode in classical analytic variables. It must also be noted that the combination of the two matrices indicates a much larger population of muons in this validation sample, compared to the charged pions which occur much less regularly, especially once the presence of cosmic muons is accounted for. The training sample used to produce the Razzled tool accounted for this offset by reweighting each particle class to the same scale (with a large enough sample size to do so without prohibitively inflating statistical fluctuations). The confusion matrices also represent the integral of the performance across the momentum phase space. Both track-like and shower-like particles possess some variation in their behaviour across different energies. Whilst Razzled does not explicitly account for such variation, some of its variables (such as track length) offer some insight into different energy particles and how to account for them.

It is important to note the improvements achieved by the Razzled PID relative to the previous Razzle and Dazzle PID tools. Figure 7.14 shows receiver-operator-characteristics curves for the four particles of interest to this selection - muons,

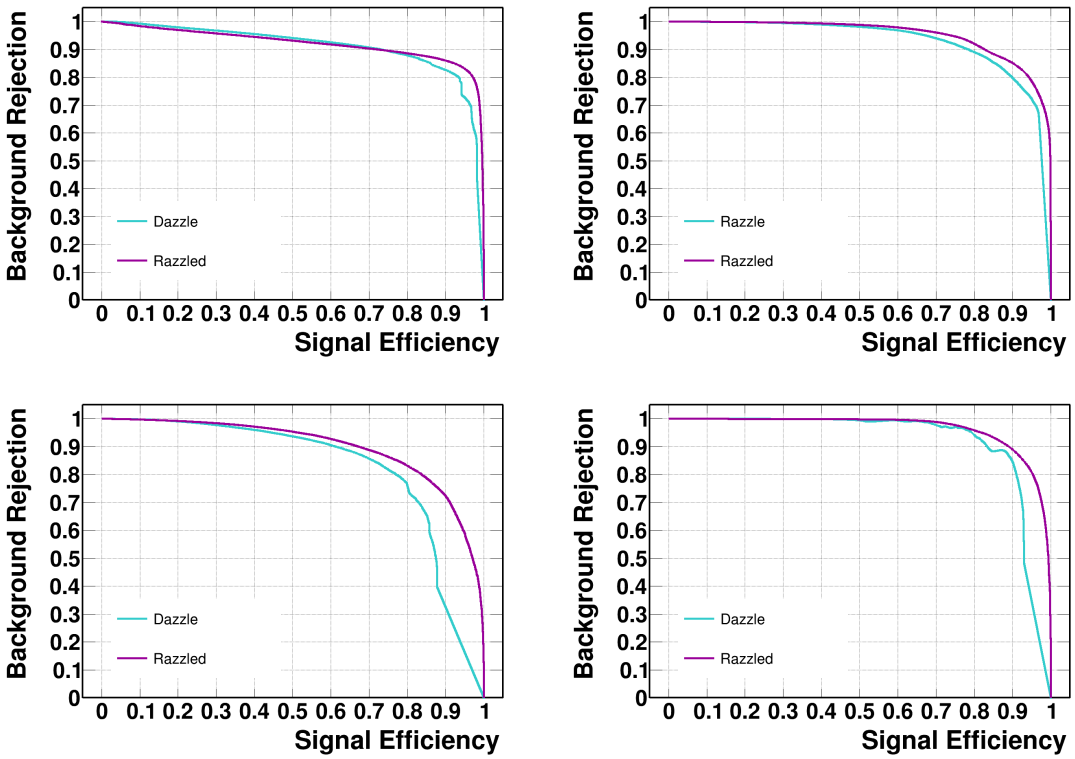


Figure 7.14: Receiver-operator-characteristics curves showing the behaviour of the previous Razzle/Dazzle PIDs relative to the new Razzled PID, for muons (top left), photons (top right), charged pions (bottom left) and protons (bottom right).

photons, charged pions and protons. It is clear that for all four categories the Razzled PID outperforms the previous iteration. The area under the curve, a metric that expresses the power of the score, increases from 0.929 to 0.953 for the key category, signal photons. The performance is similar for muons and even better for charged pions and protons. The curves for previous PIDs show more jumps, due to their dependence on a more limited set of variables and harsher cuts on their training samples. It is also clear, if you follow the curves from right to left, that the previous PIDs always begin with a single significant decrease in signal efficiency. This is the difference that corresponds to the inclusion of mis-characterised PFOs which was the primary motivation for combining the previous tools. Whilst Razzled has been designed to be a PID tool that is agnostic to the channel of study, it is clear that using Razzled as the primary tool in the NC $1\pi^0$ selections will improve their performance in rejecting muons, selecting pion-decay photons, and selecting hadronic final states.

7.5 Identifying NC $1\pi^0$

Having removed the vast majority of cosmic and non-fiducial backgrounds at the cosmic rejection stage, the particle identification techniques discussed in section 7.4 can now be used to identify NC $1\pi^0$ events and reject other neutrino backgrounds. Firstly, any events containing a primary PFO that falls into Razzled's Muon class are removed. This reduces the number of CC ν_μ slices by over 68 % but still leaves them as the most dominant background. The primary reason for the remaining CC ν_μ slices is the muon falling into Razzled's Charged Pion class, particularly those of lower momentum and therefore track length. For the two exclusive channels, charged pions also lie outside of the signal definition and thus removing any slices containing a primary Razzled Charged Pion dramatically reduces the remaining CC ν_μ backgrounds by a further 80 %.

The next selection stage focuses on identifying the key signature of the signal events,

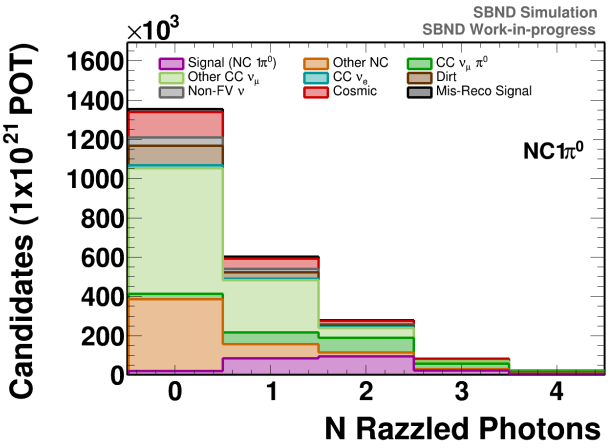


Figure 7.15: The number of primary PFOs identified as photons in each slice that has passed the cosmic and muon rejection cuts for the inclusive NC $1\pi^0$ selection.

the two photon showers aligned to a common vertex. The first cut requires that at least two PFOs in the event have been identified as photons by Razzled. This is the most severe cut of the selection, it places a high burden on the Razzled tool and is not a value based cut that can be tuned to optimise any performance metric. This reduces the number of remaining signal slices by roughly 50 % for all of the three selections but makes much more significant reductions in all background categories, increasing the purity by a factor of 4. This is well visualised in figure 7.15 which illustrates that, whilst the signal distribution strays significantly from the expected peak at 2 Razzled Photons, the backgrounds are very clearly peaked at zero. At this point 21 % of the remaining signal has at least 3 photon candidates.

It is worth considering for a moment the width of the signal distribution. Leaving aside any inefficiencies resulting from the Razzled tool's imperfect identification of photons, there are driving factors in the reconstruction of showers that complicate the picture, particularly in the case of neutral pion decays. The algorithms that determine shower clustering have to balance the risks of over and under clustering. If the parameters are set in such a way that all of the sparse elements of a large electromagnetic shower are clustered together then inevitably there will be

cases in which other small particles or parts of particles are also gathered up, resulting in missing information and poor shower purity. On the flip side, if the clustering is too conservative then shower completeness suffers, and with it energy reconstruction, and 'fake' shower multiplicity arise from splitting activity from one true electromagnetic cascade into multiple reconstructed showers. These effects can be exacerbated in the scenario of a neutral pion decay where two photons originate from the vertex, heightening the risk of merging or splitting the showers.

Two scenarios dominate the reconstruction inefficiencies, the first involves higher momentum neutral pions whose decays result in very forward boosted photons, as illustrated by the high energy tail of figure 7.16. The second case involves highly asymmetric decays where the energy of the sub-leading photon is so low that it proves difficult to identify as a distinct particle, often being grouped into the leading photon or missed entirely. When requiring the identification of two Razzled Photons, the NC $1\pi^0 0p 0\pi^\pm$ channel suffers a noticeably larger hit to the efficiency than the NC $1\pi^0 \mathrm{N}p 0\pi^\pm$ channel (see tables 7.3 and 7.4 later). This stems from the fact that one of the key discriminators between photons and electrons, the conversion gap, requires identification of the neutrino interaction vertex. This is possible in the Np case but not in the $0p$ case due to the absence of any charged particles emanating from the interaction point. In the latter case, the reconstructed vertex is often identified as being located at the first visible activity from one of the photon showers. For this shower, the only remaining handle with which to identify it as a photon, not an electron, is the $\mathrm{d}E/\mathrm{d}x$ for the initial track stub.

Having identified the existence of at least two photon candidates, these must be combined to form a neutral pion candidate. It is possible for the PFO characterisation algorithms described in section 5.2.1.4 to fail to produce elements of the characterisation, often occurring when the number of hits and/or three-dimensional space points are very low. If any of the photon candidates suffer from this issue in relation to their energy or direction they are removed from consideration.

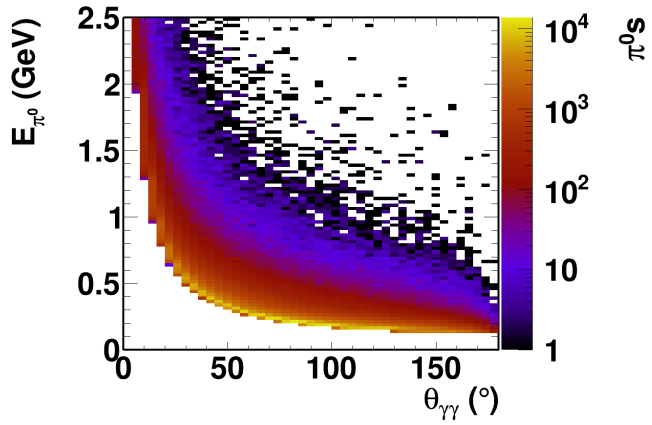


Figure 7.16: The relationship between the energy of a neutral pion and the angle between the two photons produced when it decays.

If at this stage there are less than two photon candidates remaining, the slice is discarded. If not, the reconstructed invariant mass of all pair permutations of remaining photon candidates is calculated as

$$m_{\gamma\gamma} = \sqrt{2 \cdot E_1 \cdot E_2 \cdot (1 - \cos\theta)}, \tag{7.2}$$

where E_1 and E_2 are the energies of the two photon candidates and θ the angle between their directions. The pair with the closest invariant mass to $134.9769\,\mathrm{MeV/c^2}$ is retained as the neutral pion candidate.

A further cut is applied to the two exclusive selections, NC $1\pi^0 0p 0\pi^\pm$ and NC $1\pi^0 \mathrm{N}p 0\pi^\pm$, in order to specify the number of protons as being zero or non-zero respectively. This determination is also made using the Razzled tool. In both this cut, and the zero charged pions cut mentioned previously, the kinetic energy of the reconstructed PFOs is required to be above 65.3 MeV (for charged pions) and 81.7 MeV (for protons) respectively. These requirements match the momentum thresholds applied to the signal definitions. At this stage the selection purity has reached 36.5 % for the inclusive selection and above 45 % for the exclusive selections.

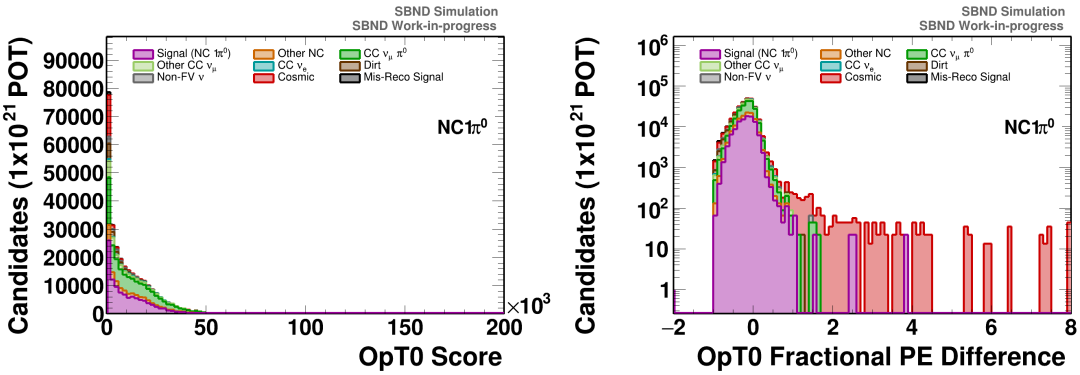


Figure 7.17: The score and fractional agreement variables from the OpT0Finder tool for all slices remaining after the neutral pion candidate identification stage in the inclusive NC $1\pi^0$ selection.

7.6 Cleaning Up

The final stage of the selection is predominantly made up of quality cuts utilising variables supplied by the OpT0Finder flash matching tool. Two metrics are used that assess the quality of agreement between the slice in question and the flash recorded during the beam spill. The first is the agreement score which is the inverse of the χ^2 agreement between the measured and hypothesised photoelectron count at each photodetector. The other metric is the fractional difference between the summed photoelectron counts for the measured and hypothesised flashes ($\frac{\text{Hypothesis}-\text{Measured}}{\text{Measured}}$). Distributions of each of these variables can be found, for the inclusive selection in figure 7.17. Both variables indicate the quality of agreement between the slice in question and the flash recorded during the beam spill, however, with subtly different purposes. The score variable considers each photodetector individually and thus generates a fine grain assessment of the agreement in both scale *and* distribution. The fractional photoelectron count difference is, by nature, coarser but gives direct information on the direction of the overall agreement, a positive value indicating over-estimation and a negative value indicating under-estimation.

For all three selections, cut values are chosen by optimising the product of efficiency and purity, as was the case for the CRUMBS score cut. Two cut values, a minimum and maximum, are chosen for the fractional metric to exclude both over and under estimations. Appendix C contains plots that demonstrate this optimisation in the equivalent manner to figure 7.3. These cuts are relatively small in effect compared to the earlier parts of the selection but do ensure the removal of a number of stray remaining cosmic slices, as well as the majority of the slices representing very poor reconstruction.

Due to the need to allow for the inclusion of charged pions in the inclusive NC $1\pi^0$ signal definition, the CC ν_μ background is still the most significant in the remaining sample. These are mainly interactions that did produce a neutral pion but for which Razzled labelled the primary muon as a charged pion. A final cut is applied to this sample to target this stubborn background. Other than the two PFOs identified as the two photons, all other primary PFOs that Pandora labelled as track-like must be contained in the detector. This cut is applied by ensuring that both the start and end of the fitted track lie no closer than 10 cm from the detector walls. This cut reduces the remaining CC ν_μ backgrounds by around 30 %, less effective than the charged pion veto applied in the exclusive selections but significant at this stage nonetheless.

7.7 Selection Results

Following the cuts described over the course of this chapter, and the optimisation of all value based cuts, the final samples of each of the three selections are presented in figure 7.18 in terms of their reconstructed diphoton invariant mass. The final efficiency of each selection lies at 34 % for the NC $1\pi^0$, 22 % for NC $1\pi^0 0p 0\pi^\pm$ and 32 % for NC $1\pi^0 \mathrm{N}p 0\pi^\pm$ with purities of 44 %, 53 % and 48 % respectively. For a total exposure of 1×10^{21} POT this corresponds to 96,000 selected

Cut Name	ϵ (%)	ρ (%)	$\epsilon\rho$	Selection ϵ (%)	Selection $\epsilon\rho$	BR (%)
No Cut	93.96	0.13	0.13	100.00	0.13	0.00
Not Clear Cosmic	93.64	0.88	0.82	99.66	0.87	84.95
FV	88.66	1.95	1.73	94.36	1.84	93.66
CRUMBS Cut (> -0.005)	84.52	4.58	3.87	89.96	4.12	97.50
No Razzled Muons	81.15	9.64	7.82	86.36	8.32	98.92
Has Two Razzled Photons	41.34	35.22	14.56	44.00	15.49	99.89
Good PiZero Kinematics	38.98	36.52	14.24	41.49	15.15	99.90
Good OpT0 Frac High (< 0.512)	38.77	36.95	14.32	41.26	15.24	99.91
Good OpT0 Frac Low (> -0.676)	37.91	38.13	14.46	40.34	15.38	99.91
OpT0 Score (> 45)	37.46	38.32	14.36	39.87	15.28	99.91
All Other Tracks Contained	34.38	43.50	14.96	36.59	15.92	99.94

Table 7.2: Efficiency and purity metrics for the cuts making up the NC $1\pi^0$ selection.

signal NC $1\pi^0$ slices, 33,000 selected signal NC $1\pi^0 0p 0\pi^\pm$ slices and 34,000 selected signal NC $1\pi^0 \mathrm{N}p 0\pi^\pm$ slices. Table 7.2 shows the evolution of the efficiency and purity for the NC $1\pi^0$ selection. Full breakdowns and background rates can be found in appendix D for all three selections. The higher purity in the exclusive selections results almost entirely from the increased ability to reject CC ν_μ backgrounds, whilst the poorer efficiency in the NC $1\pi^0 0p 0\pi^\pm$ selection reflects the increased difficulty in reconstructing neutral pions with no visible vertex activity to guide the pattern recognition. In all of the selections the background from NC events producing multiple neutral pions forms just 3-5 % of the total selected sample.

Figure 7.19 shows how the efficiency of the inclusive selection varies as a function of the opening angle and energy asymmetry of the two neutral pion decay photons. It highlights a number of themes discussed in this chapter. Firstly, it is crystal clear that by far the most significant stage is the requirement of at least two Razzled

Cut Name	ϵ (%)	ρ (%)	$\epsilon\rho$	Selection ϵ (%)	Selection $\epsilon\rho$	BR (%)
No Cut	92.77	0.07	0.06	100.00	0.07	0.00
Not Clear Cosmic	92.45	0.46	0.42	99.66	0.46	84.89
FV	85.39	0.99	0.85	92.04	0.92	93.61
CRUMBS Cut (> -0.005)	79.49	2.28	1.81	85.69	1.95	97.44
No Razzled Muons	77.89	4.90	3.82	83.96	4.11	98.86
Has Two Razzled Photons	36.77	16.59	6.10	39.63	6.57	99.86
Good PiZero Kinematics	33.83	16.78	5.68	36.47	6.12	99.87
No Razzled Pions	31.23	27.82	8.69	33.67	9.37	99.94
No Razzled Protons	23.39	45.10	10.55	25.21	11.37	99.98
Good OpT0 Frac High (< 0.400)	23.28	46.16	10.75	25.10	11.58	99.98
Good OpT0 Frac Low (> -0.676)	22.52	51.96	11.70	24.27	12.61	99.98
OpT0 Score (> 110)	22.16	53.47	11.85	23.88	12.77	99.99

Table 7.3: Efficiency and purity metrics for the cuts making up the NC $1\pi^0 0p 0\pi^\pm$ selection.

Cut Name	ϵ (%)	ρ (%)	$\epsilon\rho$	Selection ϵ (%)	Selection $\epsilon\rho$	BR (%)
No Cut	94.83	0.05	0.05	100.00	0.05	0.00
Not Clear Cosmic	94.60	0.33	0.31	99.75	0.33	84.88
FV	92.08	0.76	0.70	97.10	0.74	93.59
CRUMBS Cut (> 0.235)	82.30	2.55	2.10	86.78	2.22	98.33
No Razzled Muons	79.54	4.98	3.96	83.88	4.18	99.19
Has Two Razzled Photons	44.09	16.83	7.42	46.49	7.83	99.88
Good PiZero Kinematics	42.65	17.71	7.55	44.98	7.96	99.89
No Razzled Pions	37.65	29.13	10.97	39.71	11.57	99.95
Has One Razzled Proton	34.07	45.20	15.40	35.93	16.24	99.98
Good OpT0 Frac High (< 0.512)	33.84	45.30	15.33	35.69	16.16	99.98
Good OpT0 Frac Low (> -0.440)	32.13	47.93	15.40	33.88	16.24	99.98
OpT0 Score (> 80)	32.06	47.98	15.38	33.81	16.22	99.98

Table 7.4: Efficiency and purity metrics for the cuts making up the NC $1\pi^0$N$p0\pi^{\pm}$ selection.

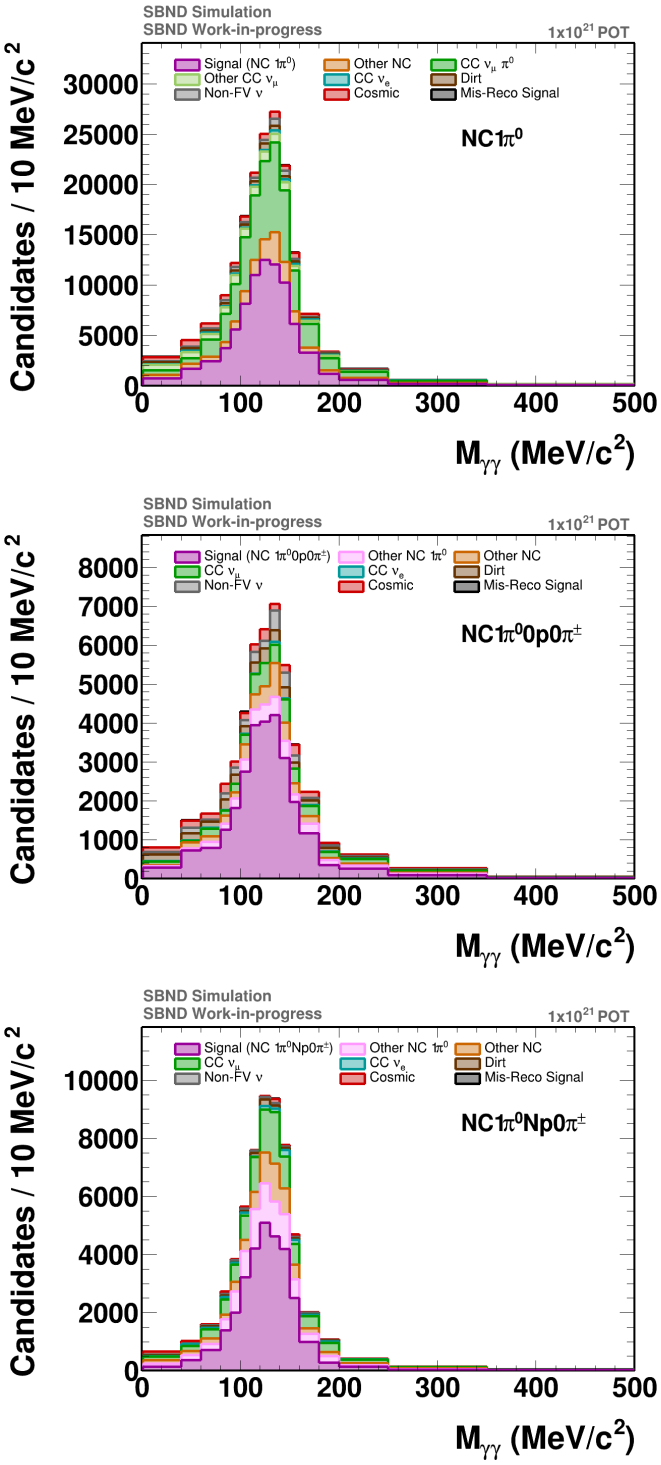


Figure 7.18: The reconstructed diphoton invariant mass distribution following the full selection process for the NC $1\pi^0$ (top), NC $1\pi^0 0p 0\pi^\pm$ (centre) and NC $1\pi^0 \mathrm{N}p 0\pi^\pm$ (bottom) selections.

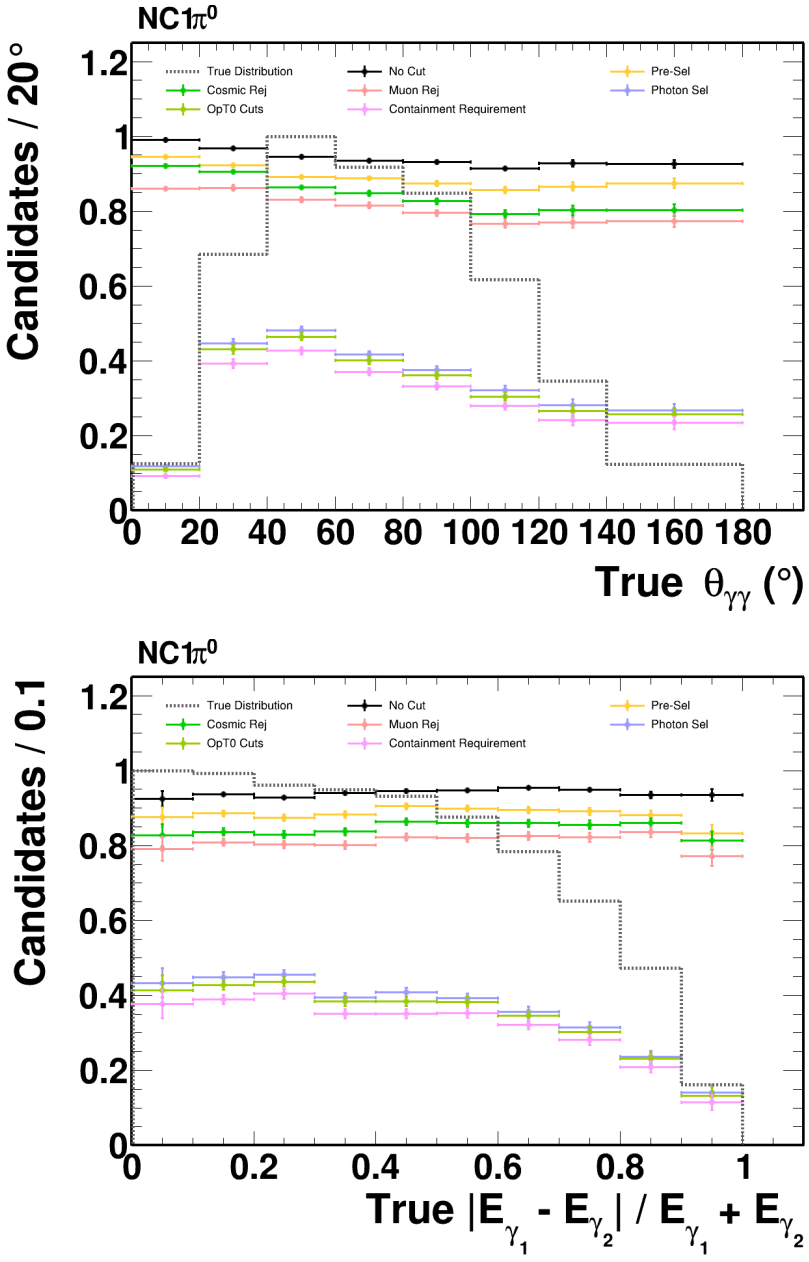


Figure 7.19: The efficiency of the NC $1\pi^0$ selection shown in terms of the broad stages of the selection, for both the opening angle (top) and the energy asymmetry (bottom) of the decay photons.

photon candidates. Whilst it results in a dramatic drop in efficiency it is also the critical step in rejecting most neutrino backgrounds. Many NC $1\pi^0$ events could be selected by requiring just a single good photon candidate, however, this approach restricts the achievable purity of the selection, reduces the average reconstruction quality of the retained signal events and prevents proper reconstruction of the neutral pion's kinematics. The choice of variables in the efficiency plots is not incidental, both show areas with much reduced efficiency from the photon selection stage onwards. For the opening angle this is in the first bin, where the angle is below 20° and the photons will overlap and be difficult to resolve individually. For the decay asymmetry this occurs in the later bins where the asymmetry is large and therefore one of the photons is likely to be very low in energy and difficult to well reconstruct or select.

In summary, despite the challenging nature of the signal, this selection rejects over 99.99 % of cosmic induced backgrounds and 99 % of neutrino induced backgrounds, preserving a strong balance between the efficiency and purity of the final sample. As the first data arrives for the SBND experiment, the building blocks that make up the pre-existing Monte Carlo selections (such as this one) can be examined to understand and reduce discrepancies. This includes the 'simple' box cut variables as well as the inputs to the different Boosted Decision Trees used in the selection. Any remaining discrepancies will need to be understood as systematic uncertainties.

Chapter 8

Demonstration of NC $1\pi^0$ Cross Section Extraction

This chapter demonstrates how the samples selected in chapter 7 can be used to extract the NC $1\pi^0$ cross section. The effect of systematic uncertainties relating to the interaction, flux and reinteraction models are discussed, and an evaluation of SBND's ability to discriminate between different generator predictions is made within a number of limitations. The systematic tools utilised are the product of the work of a wide group of people in the liquid argon and neutrino physics community over the last couple of decades. Whilst the work performed in this chapter was all my own it relies hugely on fruitful conversations with Andrew Blake, Vishvas Pandey, Andrew Furmanski, Jaroslaw Nowak, Dominic Brailsford, Afroditi Papadopoulou and many others. It was also made significantly easier by collaborating with Vu Chi Lan Nguyen and Rodrigo Alvarez-Garotte to produce the Monte Carlo files and a common analysis framework for our analyses.

The ultimate goal of this thesis is to assess how precisely SBND will be able to discriminate between different models of NC $1\pi^0$ production. Chapter 3 described the complex modelling landscape of neutrino-nucleus interactions in the $\mathcal{O}(1\,\text{GeV})$

'transition' regime and the key results from experiments measuring the cross sections of such interactions. A cross section describes the probability of a certain interaction occurring. This probability depends on the energy available to the interaction, in this case the neutrino energy. However, accurately reconstructing neutrino energy on an event-by-event level is highly challenging due to the effects of final state interactions, nuclear effects, unobserved neutral particles and detector resolution. Attempts to do so are likely to suffer significantly from model dependence as they account for this. As such, it is typical to calculate neutrino cross sections via the energy-integrated event count

$$S = n_t \cdot \int (\phi(E_\nu) \cdot \sigma(E_\nu)) \, \mathrm{d}E_\nu, \tag{8.1}$$

where n_t is the number of target nucleons available for interactions to take place, ϕ is the neutrino flux (encoding both the energy shape and total normalisation of neutrinos the experiment was exposed to), σ is the cross section of interest and E_ν is the neutrino energy.

Experiments therefore report 'flux-integrated' cross sections in which the shape of the neutrino flux to which they were exposed is convolved with the underlying cross section:

$$\sigma_{\mathrm{FI}} = \frac{S}{\Phi \cdot n_t}, \tag{8.2}$$

where

$$\Phi = \int \phi(E_\nu) \, \mathrm{d}E_\nu \tag{8.3}$$

is the integrated neutrino flux (cm^{-2}). Unfortunately experiments have limitations. Chapter 7 demonstrated that it is impossible to perfectly select all of the signal interactions with no other interactions. The S used to define our flux-integrated cross section is not what the detector observes as its count rate, N. In order to relate the two, Monte Carlo studies such as the one presented in chapter 7 are used to calculate the efficiency and purity (or background rate) of the selection, such that

the cross section can be expressed as

$$\sigma_{\mathrm{FI}} = \frac{N - B}{\epsilon \cdot \Phi \cdot n_t}, \tag{8.4}$$

where B is the number of selected background events in the Monte Carlo sample, and ϵ is the efficiency or fraction of true signal events selected. A variation of this method is to express the cross section as

$$\sigma_{\mathrm{FI}} = \frac{N \cdot \rho}{\epsilon \cdot \Phi \cdot n_t}, \tag{8.5}$$

where ρ is the purity of the selection. Whilst these methods appear identical, they have different model dependent effects on the result due to the manner in which the uncertainties are propagated in each method. When performing the subtraction method, the absolute error on the background is propagated through to the signal estimate, which, in bins of large relative background, can yield a very high uncertainty on the corrected rate. In the purity method, the fractional error is propagated which then results in much smaller uncertainties in such bins. However, this uncertainty is then larger than the background subtracted one for bins of high purity. Given the purity is not extremely high in any of the bins for this analysis, the purity technique is chosen to control the uncertainties.

When sufficient statistics are available, it is instructive for experiments to report 'differential' cross section measurements, in which the dependence of the cross section on some variable is assessed. This is done by extracting the cross section in each bin:

$$\left(\frac{\mathrm{d}\sigma_{\mathrm{FI}}}{\mathrm{d}X}\right)_i = \frac{N_i \cdot \rho_i}{\epsilon_i \cdot \Phi \cdot n_t \cdot (\Delta X)_i}, \tag{8.6}$$

where the number of selected events, N_i, the purity, ρ_i, and the efficiency, ϵ_i, are now evaluated in each bin separately. The term ΔX_i represents the bin width, in the case of double or higher order differential measurements this is the product of the bin's width in each variable, $(\Delta \boldsymbol{X})_i$. Extracting differential cross sections provides

a further complication in comparing the result to generator predictions, which will be made in true space, $\boldsymbol{Y}$, whilst the extracted result exists in reconstructed space, $\boldsymbol{X}$. Whilst equation 8.6 accounts for the inefficiencies of the detector and selection, it does not account for any bin migration due to smearing of the reconstructed observables in which the result is binned. This is accounted for via the use of a folding procedure to map one result onto the other; either forward folding, in which the generator predictions are mapped into reconstructed space, or unfolding in which the extracted result is mapped into true space.

The simplest approach to forward folding is summarised here via the relationship between a true observable y and its reconstructed counterpart x:

$$\begin{aligned} x_i &= \sum_j \epsilon_j R_{ij} y_j \\ &= \sum_j \Delta_{ij} y_j, \end{aligned} \tag{8.7}$$

where R_{ij} are the elements of the forward folding matrix describing the chance of reconstructing an event in bin i given a true location in bin j and ϵ_j is the efficiency of selecting events in true bin j. They are often represented in a combined form of the detector response matrix (Δ). If a generator predicts a true event count in bin j, $\widetilde{N}_j$, then the predicted cross section in reconstructed bin i will be given as:

$$\left(\widetilde{\frac{\mathrm{d}\sigma}{\mathrm{d}X}}\right)_i = \frac{\sum_j \Delta_{ij} \widetilde{N}_j}{\epsilon_i \cdot \Phi \cdot n_t \cdot (\Delta X)_i}. \tag{8.8}$$

The forward folding matrix is constructed using a simulated sample of selected signal events and column normalised to ensure each true event contributes a total of a single count to the forward folded prediction. This technique can be developed further to include extra bins for sideband background constraints [264]. Whilst forward folding ensures that the comparisons are made directly to the data, with all other effects applied to the predictions, the detector response matrix can suffer from model dependency. It is difficult to account for all of the variables on which

the efficiency and smearing depend, especially for high multiplicity final states or regions of rapid efficiency fluctuations [265].

The unfolding technique, which has been pursued in recent MicroBooNE, MINERνA and T2K cross section results, inverts the detector response matrix to perform the reverse process. Unfolding can be very susceptible to statistical fluctuations with small changes in the measured rate inducing much larger changes in the the unfolded true bins [265]. Regularisation techniques are often used with a flat or theory-driven prior [266] intended to prevent instabilities in the unfolded rate, at the cost of a level of model dependency bias. Popular approaches are d'Agostini [267] and Wiener SVD [268] unfolding, both of which attempt to balance the trade-off between the instabilities and the introduced bias. The procedure outlined in this chapter will utilise forward folding, to make simple generator comparisons.

8.1 Preparing the Inputs

Chapter 7 presented selected samples of NC $1\pi^0$, NC $1\pi^0 0p 0\pi^\pm$ and NC $1\pi^0 \mathrm{N}p 0\pi^\pm$. It is clear from equations 8.5 and 8.6 that, in order to extract Monte Carlo cross sections from these samples we need values for the integrated flux, the number of target nucleons and, in order to extract a differential cross section, reconstructed observables of final state kinematic quantities.

8.1.1 Integrated Flux

Section 5.1.1 outlined the simulation of the BNB flux used by SBND. The files produced by this simulation contain the flavours, locations, energies and directions of incident beam neutrinos (as well as information about the parent particles from which they were produced). As well as providing the input to neutrino interaction

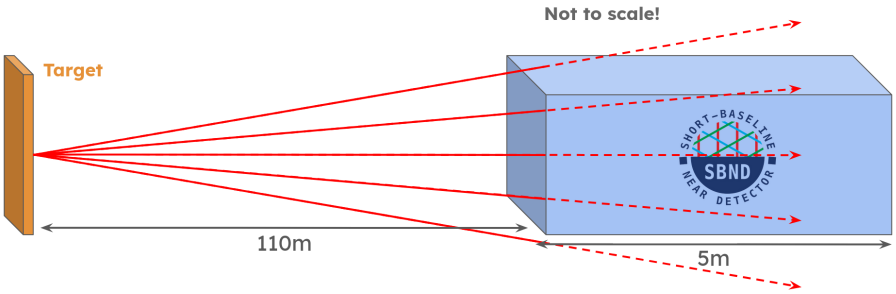


Figure 8.1: A schematic representation of the neutrino flux through the SBND detector.

generators, these files can be used to make an estimate of the integrated flux required for cross section extraction.

If the detector sees its flux source as point-like, i.e. from a large distance, then the integrated flux is just a sum of the total number of relevant neutrinos through the front face of the detector, scaled to the POT involved in the measurement. Relevant neutrinos just refers to the neutrinos which could produce the interaction of interest, for example in a CC ν_μ measurement only ν_μ should be considered. As figure 8.1 indicates, however, SBND is not subject to a point-like flux. At a distance of just 110 m and with an active volume 5 m long in the beam direction, the spread of the neutrino flux along SBND will be appreciable. Whilst the schematic suggests all neutrinos originate from the target point, in reality there is a distribution of neutrino production points depending on where the relevant decay occurs. In the simulation used to estimate this flux this is correctly handled. The direction and position used to trace each neutrino path are based upon these decay kinematics.

The integrated flux at a given baseline is evaluated using the face of the fiducial volume (LH pane of figure 7.2) to be

$$\Phi(L) = k \cdot \frac{n_{\nu_\mu+\bar{\nu}_\mu+\nu_e+\bar{\nu}_e}(L)}{A}, \tag{8.9}$$

where k is a factor to scale the sample exposure to the exposure of the measurement,

n is the sample count of all neutrino flavours through the area of the face ($A = 126,000\,\mathrm{cm}^2$) and L is the baseline distance from the target. All flavours are considered, as all can participate in this neutral-current process.

Figure 8.2 shows this integrated flux evaluated at a range of baseline values from the front to back faces of the SBND active volume (11000 cm to 11500 cm). As the flux spreads radially, the expected $1/r^2$ relationship is present and quantified via a fit of the form $\Phi = a + bL^{-2}$. The integrated flux reduces by $\sim 10\,\%$ from the front to the back face of the detector, demonstrating the significance of the effect. The fitted relationship can be used to calculate an expected value over the range of baselines encapsulated in the fiducial volume:

$$\begin{aligned}\langle\Phi\rangle &= \frac{1}{d_L}\cdot\int_{11010}^{11450}(a+bL^{-2})\,\mathrm{d}L \\ &= 1.6595\times10^{13}\,\mathrm{cm}^{-2},\end{aligned} \tag{8.10}$$

where $d_L = 440\,\mathrm{cm}$ is the length of the fiducial volume. Another method to account for the varying flux is to weight each neutrino according to its path length through the fiducial volume. Equation 8.9 becomes

$$\Phi = k\cdot\frac{\sum_\mu w_\mu}{A}, \tag{8.11}$$

where

$$w_\mu = \frac{d_\mu}{d_L} \tag{8.12}$$

is the weighting applied to a single neutrino, μ. This weighting is designed to normalise its path length through the fiducial volume (d_μ) relative to the path length of a neutrino that passes through the entire detector parallel to the beam axis ($d_L = 440\,\mathrm{cm}$). This method therefore accounts for the fact that each neutrino passes through a different amount of the volume that we consider for the analysis. In validation of the fitted method, this produces a value of $\Phi = 1.6597\times10^{13}\,\mathrm{cm}^{-2}$ within 0.01 % of the expectation value from the fit. It is this value of the integrated

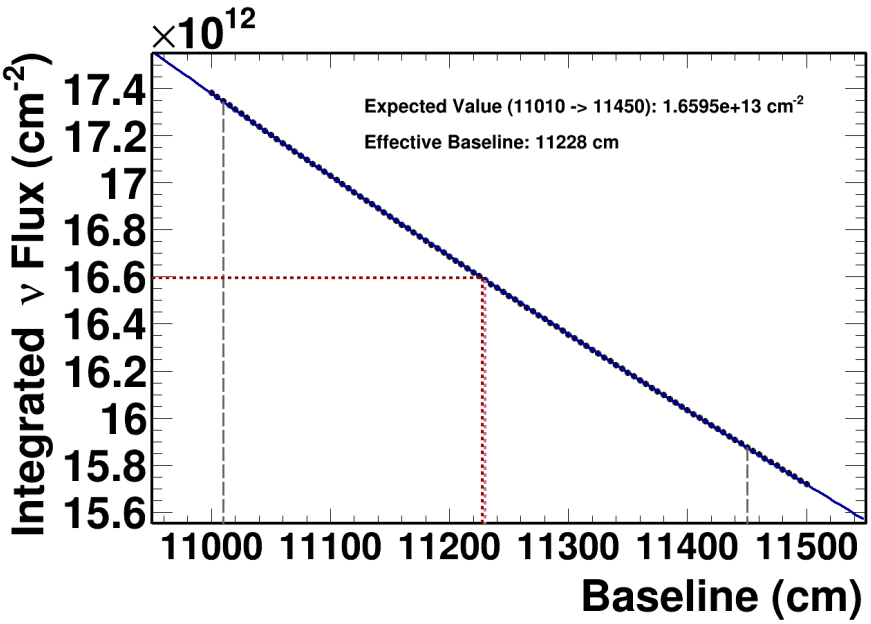


Figure 8.2: The integrated flux through the face of the SBND fiducial volume evaluated at a range of locations along the SBND detector for an exposure of $1{\times}10^{21}$ POT. A fit of the form $\Phi = a + bL^{-2}$ is shown in blue, with Φ as the integrated flux, L the baseline distance, $a = (-2.131 \pm 0.006) \times 10^{12}\,\mathrm{cm}^{-2}$ and $b = (2.3610 \pm 0.0008) \times 10^{21}$. The range of the fiducial volume in the baseline coordinate (11010 cm to 11450 cm) is indicated with the grey dotted lines. The expected value of the fitted function in this range is evaluated to be $1.6595 \times 10^{13}\,\mathrm{cm}^2$ and corresponds to an effective baseline of 11228 cm as shown via the red dotted line. This differs from the centre of the fiducial volume, 11230 cm, indicated with the pink dotted line.

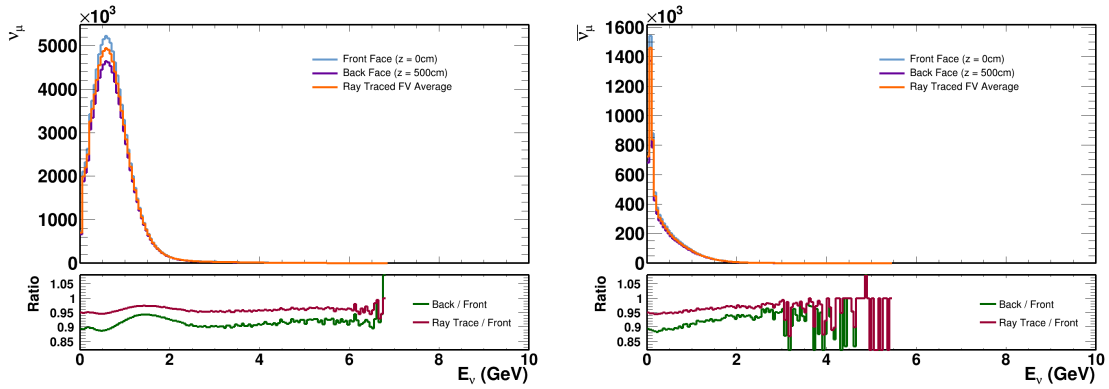


Figure 8.3: The energy dependence of the ν_μ (left) and $\bar{\nu}_\mu$ (right) fluxes when evaluated at the front and back faces of SBND, and when evaluated via weighting each flux neutrino according to its path length through the fiducial volume. The ratios of the two latter components are shown with respect to the front face evaluation.

flux that is used to calculate the final cross sections. The energy shape of the neutrino flux is also of importance. Figure 8.3 shows the neutrino flux for three scenarios (front face, back face and the weighted method) as a function of energy for the two largest contributions (ν_μ and $\bar{\nu}_\mu$). Whilst the length-dependent reduction is present across all energies, there is a larger reduction at lower energies, consistent with the fact that there is a shift toward lower energy neutrinos at wider angles, and it is these neutrinos that are more likely to 'escape' the fiducial volume over its length.

8.1.2 Number of Targets

Part of the normalisation of the cross section requires an estimate of the number of targets on which the interaction could occur. In this case that is the number of

nucleons in Argon nuclei in the fiducial volume considered. This is calculated as

$$\begin{aligned} n_t &= \frac{N_{Ar} \cdot V_{FV} \cdot \rho}{M_{Ar}} \\ &= 4.6712 \times 10^{31} \,\text{nucleons}, \end{aligned} \tag{8.13}$$

where $N_{Ar} = 40$ is the number of nucleons in an argon nucleus, $V_{FV} = 360\,\text{cm} \times 175\,\text{cm} \times 440\,\text{cm} \times 2 = 5.544 \times 10^7\,\text{cm}^3$ is the fiducial volume size, $\rho = 1.3973 \times 10^{-3}\,\text{kg}\,\text{cm}^{-3}$ is the density of liquid argon at 87 K [269] and $M_{Ar} = 39.948\,\text{AMU} = 6.63353 \times 10^{-26}\,\text{kg}$ is the standard atomic weight of argon [270].

8.1.3 Observables

8.1.3.1 What do they provide?

When considering a final state with a single observable particle, the π^0, there are two key variables covering the phase space of the interaction: its momentum (p_{π^0}) and direction, expressed as the cosine of the angle it makes with respect to the beam axis ($\cos\theta_{\pi^0}$). Figure 8.4 shows the true distributions of these variables for the three signal definitions. These plots highlight that the dominant contribution to all three channels is via resonance production, with important deep inelastic scattering and coherent contributions. Critically, they demonstrate why extracting differential cross sections enables more sophisticated analysis of generator predictions and their use of different models for each component. In particular, the coherent contribution to the NC $1\pi^0$ rate can be probed in the forward angle regions due to its small energy transfer. It can be further isolated by probing the alternative exclusive channels NC $1\pi^0 0p 0\pi^\pm$ and NC $1\pi^0 \text{N}p 0\pi^\pm$, being more significant in the former and, by definition, absent in the latter.

As discussed in chapter 4, SBND will collect a vast neutrino-argon interaction dataset, allowing for fine grain cross section measurements to be performed. This

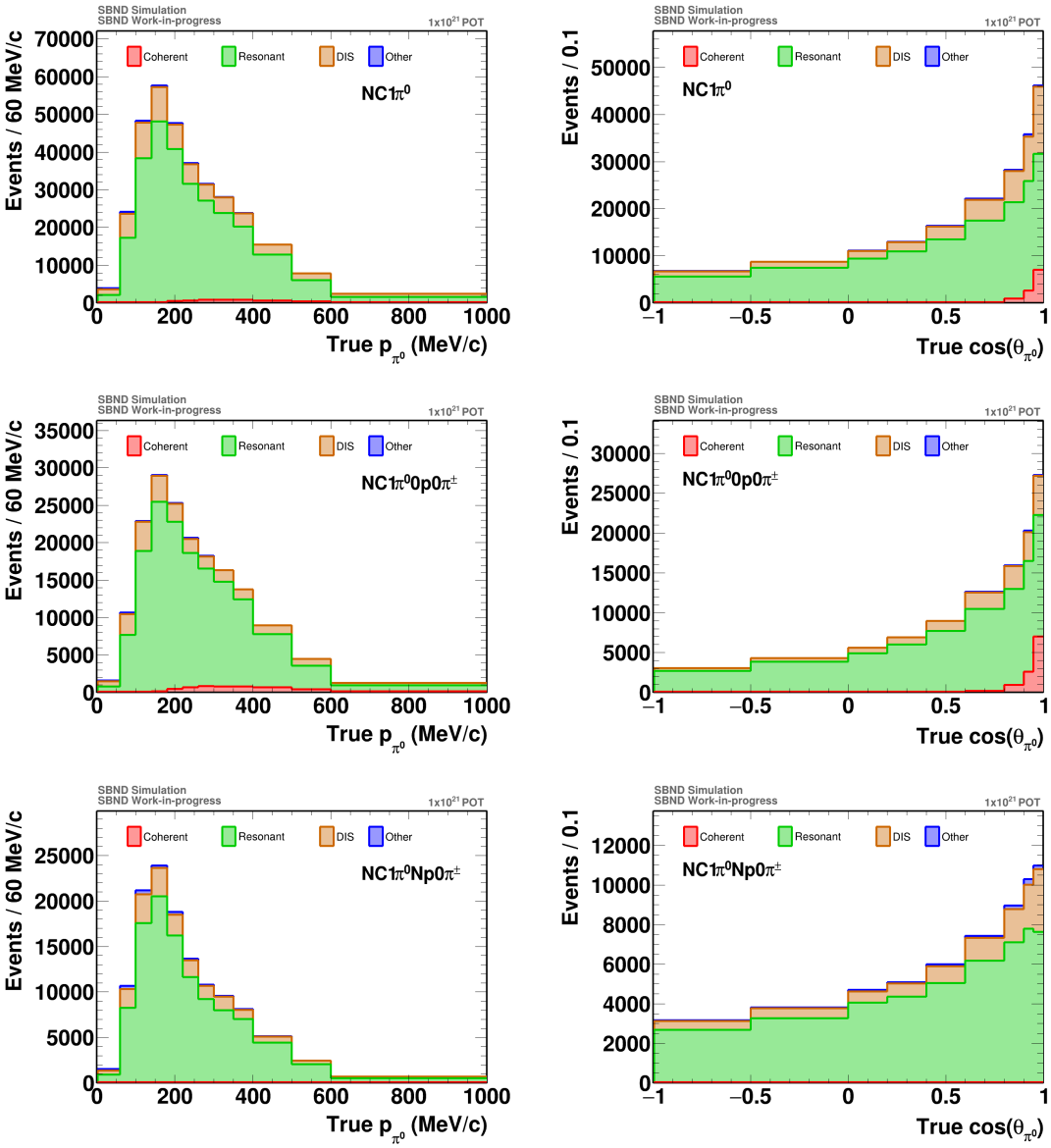


Figure 8.4: Distributions in terms of the neutral pion momentum (left) and direction (right) for true signal events in the NC $1\pi^0$ (top), NC $1\pi^0 0p 0\pi^\pm$ (middle), and NC $1\pi^0 \mathrm{N}p 0\pi^\pm$ (bottom) signal definitions.

is indicated in figure 8.5 which shows the two aforementioned variables in a double differential form for the expected exposure in the NC $1\pi^0$ channel. Similar plots for the exclusive channels can be found in appendix E and it is clear that even with detector and selection inefficiencies applied the statistics will still be sufficient for precise measurements of exclusive channels in the double differential space.

8.1.3.2 How do we reconstruct them?

In section 7.7 it was clear that the reconstruction and identification of both electromagnetic showers from a $\pi^0 \longrightarrow \gamma\gamma$ decay is challenging due to the large number of events in which the showers overlap, or the energy imbalance ensures one shower is much smaller and easily missed or merged. This makes the challenge of reconstructing the neutral pion's kinematics significant.

The neutral pion momentum can be reconstructed as:

$$\vec{p}_{\pi^0} = E_1 \cdot \hat{r_1} + E_2 \cdot \hat{r_2} \tag{8.14}$$

where E is the photon candidate's reconstructed energy and $\hat{r}$ is its direction unit vector. This is then used to reconstruct the two desired observables:

$$\begin{aligned} p_{\pi^0} &= |\vec{p}_{\pi^0}| \\ \cos\theta_{\pi^0} &= \hat{p}_{\pi^0} \cdot \hat{z}, \end{aligned} \tag{8.15}$$

where $\hat{z}$ is the unit vector along the beam axis.

Section 5.2.1.4 touched on the existing tools used to reconstruct quantities such as energy, $\mathrm{d}E/\mathrm{d}x$ and direction from a PFO under track and shower hypotheses. The energy of a shower is reconstructed in each plane by taking the integral of each of the contributing hits and correcting for lifetime recombination effects. A calibrated factor for each plane is then used to convert this summed ADC into a total number of produced ionisation electrons and finally into the equivalent energy,

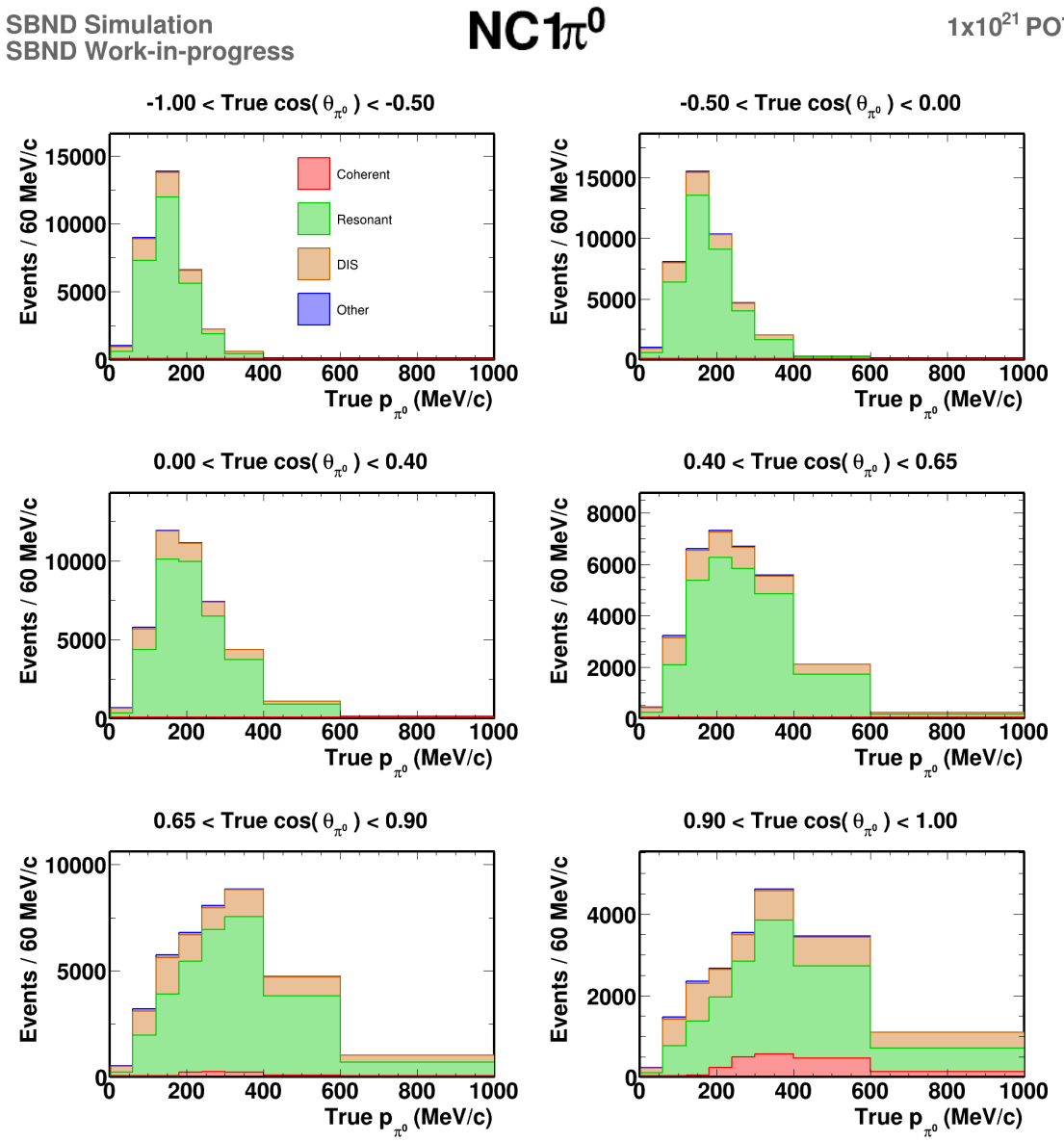


Figure 8.5: Distribution of the neutral pion momentum in slices of direction for true signal events in the NC $1\pi^0$ signal definition.

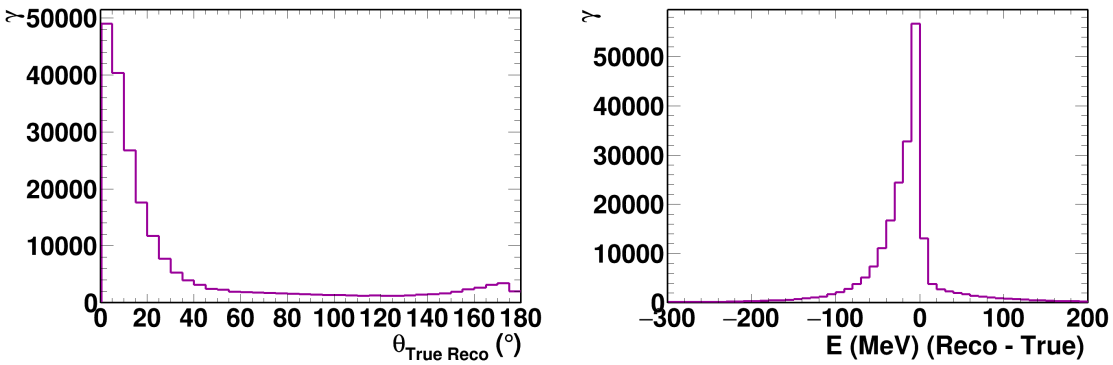


Figure 8.6: The resolution of the standard direction (left) and energy (right) reconstruction for a sample of photon showers.

assuming each electron ion pair requires 23.6 eV to produce. The shower's direction is calculated both using a principal component analysis to find the primary axis or via a sliding linear fit of the initial track stub. If they disagree strongly the former is kept, otherwise the latter is kept.

Figure 8.6 shows the resolution of the two quantities (photon shower energy and direction) that will be used to reconstruct the π^0 kinematics. Whilst both are peaked around zero, demonstrating accurate reconstruction, they also have clear failure modes. The direction has a second, much smaller, spike around 180° indicating an inverted reconstruction direction and the energy has a significant left hand tail representing under-estimation of energy. Given the shower energy reconstruction consists of summing the calorimetric energies associated with each reconstructed TPC hit, this is driven primarily by under-clustering occurring in the pattern recognition.

When analysing such variables for 'well reconstructed' photons, i.e. those with high completeness and purity, there is a noticeable uplift in performance and a corresponding decrease in the identified failure modes. This indicates that when the pattern recognition successfully clusters the majority of hits belonging to the photon, without any major impurities, then the tools that characterise this cluster

as a particle, attributing a direction and energy, do so well. A key area for future development will be the creation of tools to refine the reconstruction of showers more broadly but with a particular emphasis on splitting and searching tools targeted at the neutral pion topology.

However, in the shorter term, the reconstruction of the neutral pion observables can be improved via a trio of measures. The first two points correspond to simple corrections to the individual photon kinematics. Firstly, if the reconstructed energy is below 150 MeV, the direction of the photon is taken from the track characterisation of the reconstructed particle, rather than the shower characterisation. The improvement is shown in figure 8.7 to be relatively mild, but stems from the fact that, for lower energy photons, which by nature look more track-like, the track fitting is more representative. At moderate energies both characterisations offer similar performance, whilst at the highest energies they invert. The point of inversion is represented in the plot by the dashed red line, above this value the shower characterisation is used, and below it, the track characterisation.

The other simple correction applies to the shower energy. As mentioned previously, the energy reconstruction suffers from a general under-estimatation due to the fact that some shower hits in the sparse outer edges of the shower are often unclustered in order to balance the opposite risk of over-clustering. This can be crudely corrected for via the use of an energy dependent scaling factor. Figure 8.8 shows the fractional energy resolution profiled across reconstructed energy, the values in each bin provide a corrective factor, C_E, that can be used to construct a 'corrected' energy:

$$E_{\text{Corr}} = \frac{E_{\text{Reco}}}{C_E + 1} \tag{8.16}$$

where E_{Reco} is the original reconstructed energy. Whilst the statistically dense region (below 500 MeV) shows the expected negative factor, the factor becomes positive for the high energy tail, driven by the merging pathology, in which a highly correlated

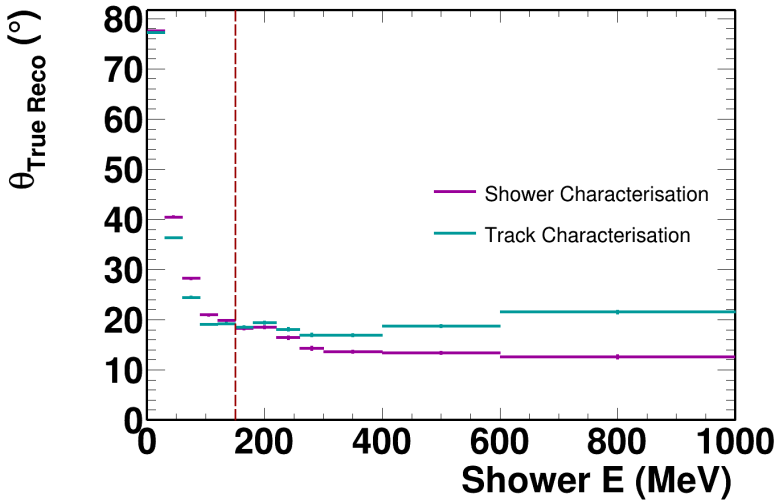


Figure 8.7: The angle between the true photon direction and the reconstructed photon direction profiled over the reconstructed photon energy. The point at which the resolution begins to favour the shower characterisation over the track characterisation is shown with the red dashed line at 150 MeV.

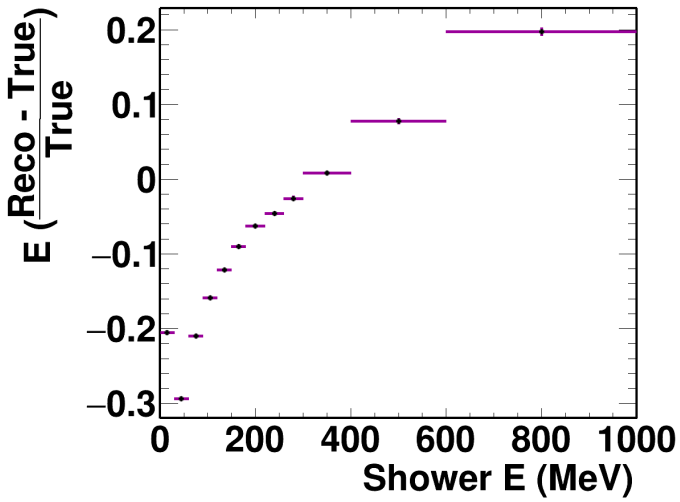


Figure 8.8: The fractional energy resolution of reconstructed photon showers profiled across reconstructed energy.

pair of photons from a higher energy π^0 are reconstructed as a single object.

The combined effect of both corrections is shown in figure 8.9 for the two observables constructed using equation 8.15. It is clear that the targeted measures, particularly the energy correction factor, reduce the bias tail for the momentum without introducing significant smearing. Both corrections have a negligible effect on the direction reconstruction, which did not have an initial bias like the momentum.

A final, more sophisticated correction is applied to the pion kinematics exploiting the fact that the reconstructed diphoton invariant mass (see equation 7.2) should agree well with the π^0 mass of $134.9769\,\mathrm{MeV/c^2}$. The constraint is applied using a Lagrange multiplier method, following the approach of ProtoDUNE [271]. The minimisation equation is constructed as:

$$\chi^2 = \lambda \cdot \mathcal{H} + (\boldsymbol{\alpha} - \boldsymbol{\alpha}_0)^T V_{\mathrm{cov}}^{-1} (\boldsymbol{\alpha} - \boldsymbol{\alpha}_0) \tag{8.17}$$

where $\mathcal{H} = |2 \cdot E_1 \cdot E_2 \cdot (1 - \cos\theta_{\gamma\gamma}) - m_{\pi^0}^2|$ is the invariant mass constraint, $\boldsymbol{\alpha} = (E_1, E_2, \theta_{\gamma\gamma})$ is a vector of the kinematic variables being varied, $\boldsymbol{\alpha}_0$ is the initial values of the variables and V_{cov} is the diagonal-only covariance matrix encoding the inherent measurement uncertainty in each of the kinematic variables. When considering a full covariance matrix, the convergence rate of the correction dropped and as such the resulting improvements were reduced.

For a given starting value of λ the optimal $\boldsymbol{\alpha}$ is found to minimise the χ^2. The value of λ is then iterated according to the range and step size provided and $\boldsymbol{\alpha}$ re-optimised. The optimal λ is selected as the value which minimised χ^2 whilst satisfying the invariant mass constraint. This double minimisation process is implemented using the MIGRAD algorithm in ROOT's implementation [272] of the Minuit minimisation system [273].

The minimisation converges in $\sim 85\,\%$ of selected NC $1\pi^0$ candidates, with similar

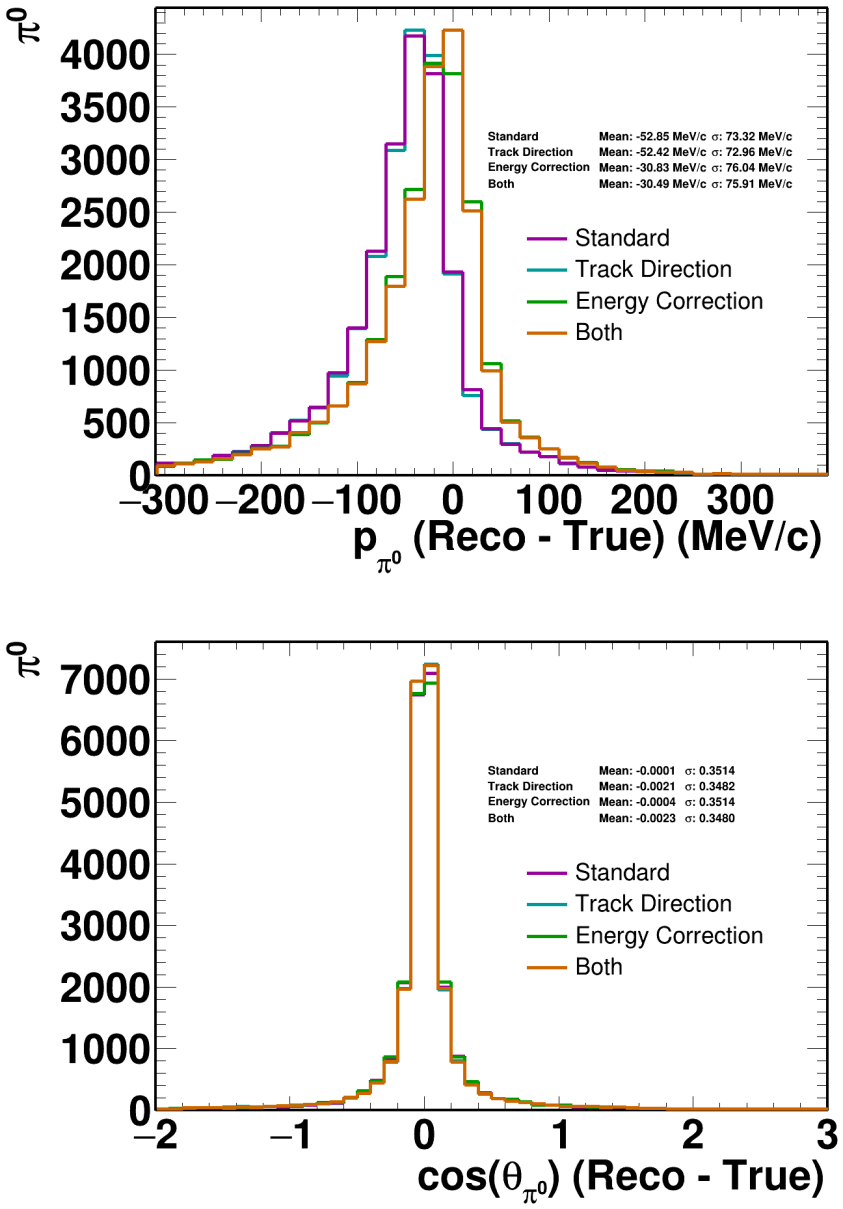


Figure 8.9: The resolution of the reconstructed π^0 momentum (top) and direction (bottom) under the standard reconstruction, and with the application of two corrections described in the text.

rates for both the signal and background cases. There is a marginal difference between backgrounds without a true π^0 (convergence $\sim 84\,\%$) and those without (convergence $\sim 81\,\%$) but it is not enough of a difference to give a handle to further background rejection. When the minimisation does converge, then the minimised value of $\boldsymbol{\alpha} = (E'_1, E'_2, \theta'_{\gamma\gamma})$ can be used to update the π^0 momentum to

$$p_{\pi^0} = \sqrt{(E'_1 + E'_2)^2 - m^2_{\pi^0}}. \tag{8.18}$$

The results of applying this improvement can be seen in figure 8.10 to make a significant difference to the p_{π^0} resolution. Overall the improvements result in a reduction of the overall bias from $53\,\mathrm{MeV/c}$ to $15\,\mathrm{MeV/c}$ driven by reductions in the under-estimation tail. Improving the quality of the reconstructed observables reduces the amount of bin migration that will occur in the extracted differential cross section therefore increasing the sensitivity to subtle model differences.

A future study utilising a higher statistics sample could evaluate whether an introduction of a momentum dependence to the observable corrections would yield better results by accounting for the variation in resolution across different scales. However, in the opinion of the author, more work to improve the underlying shower energy reconstruction should be the first priority, before working further on corrective measures.

8.2 Systematic Uncertainties

Other than the observed event count, every quantity used in the cross section extraction procedure has to be estimated using Monte Carlo simulations. This introduces a number of opportunities for potential systematic bias. No simulation will perfectly match reality, and by definition we do not know exactly what the reality

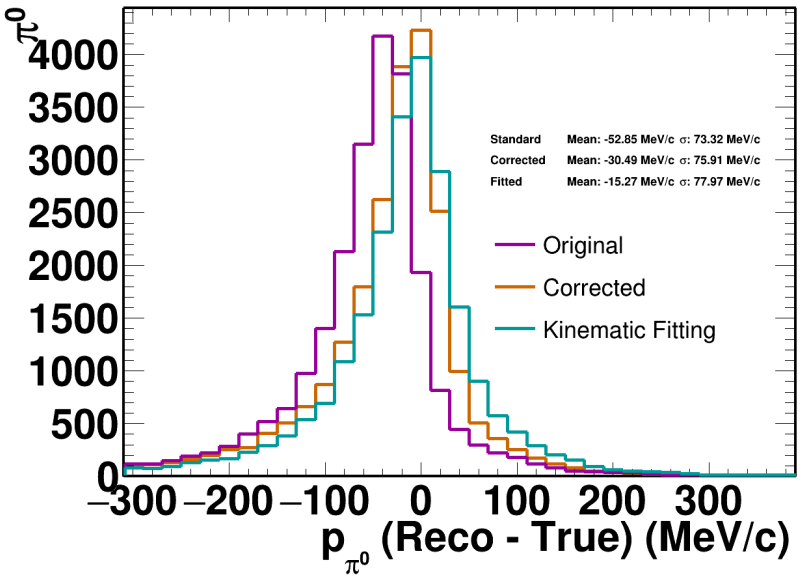


Figure 8.10: The resolution of the reconstructed π^0 momentum under the standard reconstruction, with the simple corrections applied, and finally with the kinematic fitting applied.

is. These biases need to be estimated in order to understand how precisely we can trust our simulation and thus whether or not a given prediction is in agreement with the observed data. The typical categories of systematic uncertainty for making a neutrino cross section measurement are:

- Flux modelling
- Interaction modelling
- Reinteraction modelling
- Detector modelling
- POT accounting
- Target size estimate.

Alongside these systematic uncertainties, a statistical uncertainty will be present in both the data and Monte Carlo samples. Monte Carlo samples are traditionally produced to be of a large enough size in order to make the associated statistical uncertainties negligible compared to the dominating data statistical uncertainty and/or systematic uncertainties.

The study presented in this thesis occurs in advance of SBND data taking which places some limitations on the resources and samples available to assess the measurement uncertainties. This presents itself in two ways - the absence of quantification of detector systematics and statistically limited Monte Carlo samples. The former will be discussed in section 8.2.2, whilst the latter is a consequence of SBND's huge exposure, meaning currently available samples represent less than $1\times$ the expected dataset. Typically something of the order of $10\times$ the dataset is produced to ensure the simulation statistical uncertainty remains negligible but this need not be necessary in the regime of very large datasets. Whilst the available samples provide sufficient statistics to assess the total cross section, the statistical uncertainty in some of the double differential bins can be as large as 10 % and thus larger samples would be required for pursuing a measurement in these differential spaces.

8.2.1 Reweightable Uncertainties

Systematic uncertainties, especially those arriving from complex models with many input parameters, are often evaluated by assessing the impact of changing the values of these model parameters within a defined uncertainty range. Typically we define a 'universe' in which a series of parameters take certain values and then assess how different the physics result is in such a universe. This can then be performed in 100s or 1000s of different universes to assess the overall effect of such a systematic.

Processing Monte Carlo LArTPC events through the entire simulation and recon-

struction workflow is highly computationally expensive, in no small part due to the length of the detector readout, coupled to a vast number of channels with fine grain sampling. As such, creating many 'universes' for each source of systematic uncertainty is unrealistic. Thankfully many of the systematic uncertainty sources can be approached via a process called reweighting.

Reweighting works on the principle of being able to assess how much a parameter change directly affects the chance of a given neutrino interaction occurring and forming part of the measurement. This is directly applicable to the flux, interaction and reinteraction uncertainties. If we take an input parameter A and tweak it for a given universe according to its associated uncertainty (σ_A), we get

$$A' = A + f \cdot \sigma_A, \tag{8.19}$$

where f is the scale of the parameter tweak, sampled from a Gaussian with $\mu = 0$ and $\sigma = 1$. If the parameter analytically maps to the probability, P, of an interaction occurring (normally via the cross section) then each interaction can be given a weight in universe i:

$$W_i = \frac{P(A')}{P(A)}. \tag{8.20}$$

This weight, for every interaction, is then used to generate the physics result in that universe. In the case of a cross section extraction this means that rather than each event contributing a single count to the estimation of the signal efficiency and the background rate, they will contribute W_i. A trivial example could be imagined where a universe is created in which the neutral-current cross section is reduced by 20 %. Every neutral-current interaction event in the Monte Carlo sample would now count for 0.8 and as such the efficiency of selecting NC $1\pi^0$ events would remain the same, but the relative size of the charged-current backgrounds would increase, yielding a lower selection purity and thus a lower cross section value.

8.2.1.1 Flux

As described in section 5.1.1, SBND makes use of the flux simulation developed by the MiniBooNE experiment and utilised by other BNB experiments such as MicroBooNE and ICARUS. This simulation provides tools for reweighting the flux prediction according to the 13 parameters listed in table 8.1 [181]. They can be broadly categorised into three groups:

- *Horn Magnetic Field* - the behaviour of the current pulsed through the beam's focusing horn, and the current induced on the horn's surface by the changing field. Both affect the angular and energy distributions of the different flux components.

- *Meson Production Cross Sections* - the neutrino flux results from the decays of mesons and kaons produced in the primary proton-beryllium interactions, their cross sections are critical inputs to the flux simulation.

- *Hadronic Reinteraction Cross Sections* - subsequent interactions of hadronic (meson / nucleon) primaries can modify decay pathways that result in beam neutrinos.

The primary location in which the flux uncertainties enter the cross section calculation is through the integrated flux. They do also impact the efficiency and purity predictions but in a virtually negligible manner. Figure 8.11 shows the integrated flux calculated using the path weighted method described in section 8.1.1 in 1000 universes created by varying the 13 parameters together. The resulting variation of the integrated flux value is very significant, showing a standard deviation of 6.2 % and even more significantly a bias of 12.4 % between the nominal universe used to produce the standard Monte Carlo samples and the mean of the values in each universe.

Both of these effects arise from the π^+ production uncertainty, with all other parameters contributing much smaller uncertainties. Figure 8.12 shows the energy distribution of the path weighted neutrino flux, in 1000 universes created by varying just the π^+ production rate. The uncertainty is much more significant below about 800 MeV and this is also where the bias with respect to the nominal universe enters the picture. The HARP pion production data used to constrain this parameter does not cover the low momentum and high angle regions that dominate the production of lower energy neutrinos. A Sanford-Wang parameterisation is fitted to the data and used to predict the nominal flux. In order to create the universe prediction, splines are created containing the fit predictions in a number of scenarios in which the data is varied within its error bands. The fits are unconstrained in the regions the data does not cover, and the parameterisation cannot fully accommodate the shape of the data, leading to the offset between the nominal and mean. This discrepancy will need to be addressed before SBND publishes its first results, and work is ongoing to improve the flux simulation and also to prepare to make use of upcoming data from the NA61/SHINE experiment [274] to better constrain the hadron production at these energies.

8.2.1.2 Interaction

The GENIE interaction generator is used to model the neutrino-argon interactions contributing both to the signal efficiency and the background/purity estimates. The GENIE software provides a plethora of reweightable parameters to assess the impact of uncertainties in the model predictions [115, 275]. These can broadly be separated into two categories. The first category represents a 'group' response in which a number of related parameters are thrown together to create a universe, the rest are individual parameters assessed in isolation. The former set are summarised in table 8.2 and the latter in 8.3.

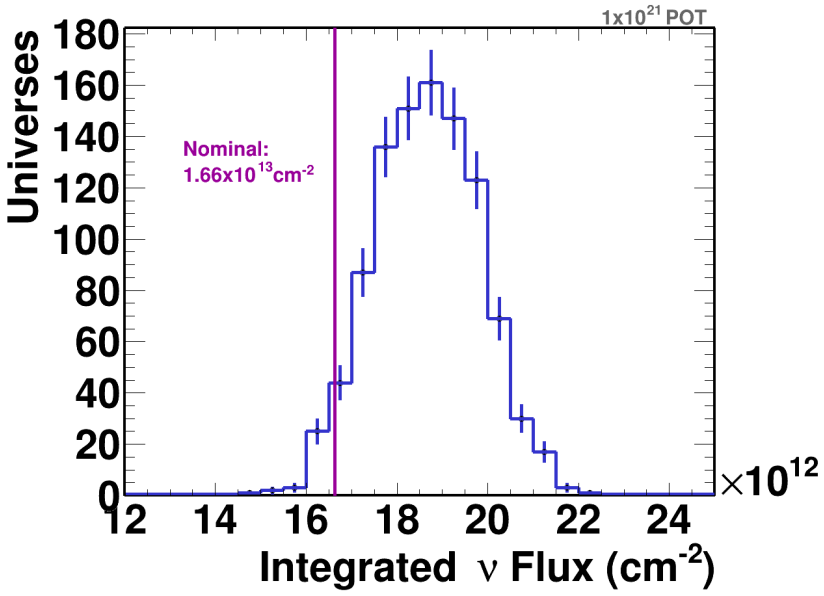


Figure 8.11: The path weighted integrated flux of all neutrino flavours from the Booster Neutrino Beam through SBND assessed in 1000 systematic universes.

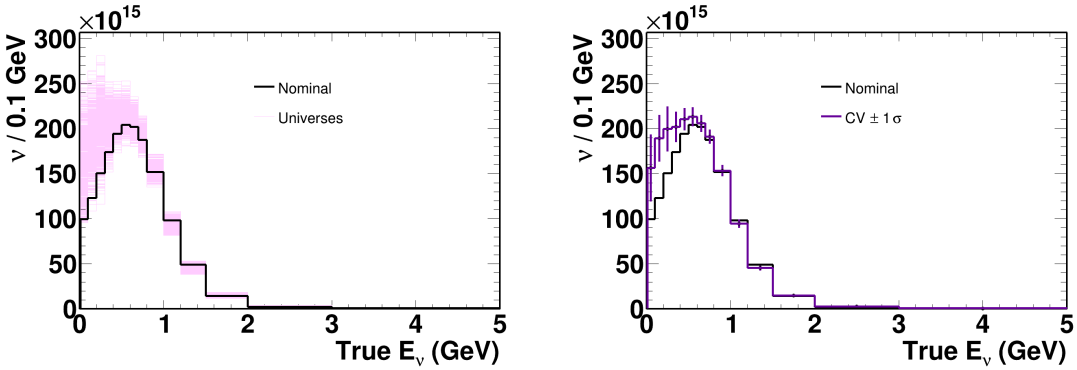


Figure 8.12: The number of neutrinos, passing through the SBND fiducial volume, weighted by path length and distributed by their true energy. The left hand panel demonstrates the distribution in each of 1000 systematic universes in which the π^+ production rate is varied according to its uncertainty. The right hand panel shows the mean and standard deviation calculated from the universes.

Parameter	Description
Horn Magnetic Field	
Skin Current	Current induced on the horn surface by the local field
Horn Current	Current pulsed through the horn
Meson Production Cross Sections	
σ_{K^-}	K^- production on Be
σ_{K^+}	K^+ production on Be
σ_{K^0}	K^0 production on Be
σ_{π^-}	π^- production on Be
σ_{π^+}	π^+ production on Be
Hadronic Reinteraction Cross Sections	
$\sigma_{N,\,\mathrm{Inel}}$	Nucleon inelastic interactions on Be / Al
$\sigma_{N,\,\mathrm{QE}}$	Nucleon quasi-elastic interactions on Be / Al
$\sigma_{N,\,\mathrm{Tot}}$	Total nucleon interactions on Be / Al
$\sigma_{\pi,\,\mathrm{Inel}}$	Pion inelastic interactions on Be / Al
$\sigma_{\pi,\,\mathrm{QE}}$	Pion quasi-elastic interactions on Be / Al
$\sigma_{\pi,\,\mathrm{Tot}}$	Total pion interactions on Be / Al

Table 8.1: The parameters available to reweight the BNB flux simulation [181].

As would be expected, the impact of the cross section uncertainties comes, in the most part, from a small subset of these parameters. This is due to the differences in the inherent uncertainties of the parameters themselves but also their level of relevance to the signal and common background interactions. The parameter that introduces the largest uncertainty to all three channels is the `NCRESVariationResponse` in which the vector and axial vector masses are varied to affect the strength of neutral-current resonant interactions. They are varied according to Gaussians with 10 % and 20 % widths respectively [275]. The other two sets of parameters that also contribute significantly to the interaction uncertainty are those relating to the charged-current resonance production and those relating to final state interactions for pions. The former is particularly important for the inclusive selection in which there is a large CC π^0 background, and the latter for the exclusive selections in which the more specific final state can be masked by the impact of FSI. Many of the parameters varied for the cross section systematic variations, including `NCRESVariationResponse`, are based on conservative placeholder uncertainties rather than constraints to existing data [276].

Figure 8.13 shows a few examples of how interaction uncertainties impact the inputs into the cross section extraction that account for the imperfect selection: efficiency, purity and background count. It is clear that the purity is subject to much more variation than the efficiency. This is due to the fact that signal events, selected or not selected, will scale in more similar ways than the signal and background with respect to each other. The top left shows the entire effect of the interaction systematics. In total the systematic uncertainty due to the interaction systematics sits around 13 % for the inclusive channel and ∼8 % for the two exclusive channels due to their higher purity. The other plots show particular parameters of interest, the charged-current resonant parameters (bottom left) have no impact on the efficiency, as the signal comes from neutral-current events only. However, this does show that it provides a very significant chunk of the purity uncertainty in this channel. The neutral-current resonant parameters (bottom right) also impact the purity significantly,

this time through the signal events, indicated by the reduced background count band, and the presence of the efficiency band. Finally, the parameter governing the modelling of the pion angular distribution (top right) shows that its impact is much more pronounced in some bins than others, indicating a shape uncertainty in the reconstructed direction space.

As is highlighted in the above evaluation, the prediction of the background rates is subject to significant systematic uncertainty. We know that there are large uncertainties in neutrino-argon interaction modelling, it is one of the motivations for carrying out more precise cross section measurements in the first place. However, for lower purity selections, the presence of a large uncertainty on the rate of neutrino induced backgrounds can reduce the precision of the result. One method for mitigating this impact is to make sideband measurements aimed at constraining these systematics with the available data. This can take the form of inverting one or more cuts to select a common background or choosing a background rich region in one of the selection variables. In the selection presented in this thesis the obvious sideband would be to invert the 'No Razzled Muons' cut in order to make a sample of the CC ν_μ $1\pi^0$ background. By constraining the rate and distribution of these events against the data, the predicted selected background due to these events can be rescaled accordingly, reducing the reliance on the GENIE prediction.

8.2.1.3 Reinteraction

As the products of neutrino interactions propagate, further interactions can occur with the argon. In a similar manner to final state interactions within the nucleus, the inelastic interactions of the hadronic products such as absorption, pion production or charge exchange, can result in a difference in the apparent final state. The reconstruction of kinematic quantities can also be affected by both elastic and inelastic interactions. The propagation of particles through the argon is handled with

Parameter	**Description**
CCRESVariationResponse	Variation of the axial and vector masses for charged-current resonant interactions
DISBYVariationResponse	Variation of four parameters defining the Bodek-Yang deep inelastic scattering model
FSI_N_VariationResponse	Variation of the inelastic, absorption and pion production interaction strengths and the mean free path for nucleons in the hA FSI model
FSI_pi_VariationResponse	Variation of the inelastic, absorption and pion production interaction strengths and the mean free path for pions in the hA FSI model
NCELVariationResponse	Variation of the axial mass and strange axial form factor for neutral-current elastic interactions
NCRESVariationResponse	Variation of the axial and vector masses for neutral-current resonant interactions
ZExpAVariationResponse	Variation of the first four coefficients of the Z expansion of the axial form factor for charged-current quasi-elastic interactions

Table 8.2: The grouped parameters available to reweight the GENIE interaction model [275].

Parameter	**Description**
CoulombCCQE	Strength of the Coulomb corrections for charged-current quasi-elastic interactions
DecayAngMEC	Variation in decay angle in meson exchange current interactions
NonRESBGvbarnCC1pi	Scale of non-resonant antineutrino-neutron charged-current interactions producing 1 π
NonRESBGvbarnCC2pi	Scale of non-resonant antineutrino-neutron charged-current interactions producing 2 π
NonRESBGvbarnNC1pi	Scale of non-resonant antineutrino-neutron neutral-current interactions producing 1 π
NonRESBGvbarnNC2pi	Scale of non-resonant antineutrino-neutron neutral-current interactions producing 2 π
NonRESBGvbarpCC1pi	Scale of non-resonant antineutrino-proton charged-current interactions producing 1 π
NonRESBGvbarpCC2pi	Scale of non-resonant antineutrino-proton charged-current interactions producing 2 π
NonRESBGvbarpNC1pi	Scale of non-resonant antineutrino-proton neutral-current interactions producing 1 π
NonRESBGvbarpNC2pi	Scale of non-resonant antineutrino-proton neutral-current interactions producing 2 π
NonRESBGvnCC1pi	Scale of non-resonant neutrino-neutron charged-current interactions producing 1 π
NonRESBGvnCC2pi	Scale of non-resonant neutrino-neutron charged-current interactions producing 2 π
NonRESBGvnNC1pi	Scale of non-resonant neutrino-neutron neutral-current interactions producing 1 π
NonRESBGvnNC2pi	Scale of non-resonant neutrino-neutron neutral-current interactions producing 2 π
NonRESBGvpCC1pi	Scale of non-resonant neutrino-proton charged-current interactions producing 1 π
NonRESBGvpCC2pi	Scale of non-resonant neutrino-proton charged-current interactions producing 2 π
NonRESBGvpNC1pi	Scale of non-resonant neutrino-proton neutral-current interactions producing 1 π
NonRESBGvpNC2pi	Scale of non-resonant neutrino-proton neutral-current interactions producing 2 π
NormCCMEC	Normalisation scale for charged-current meson exchange current interactions
NormNCMEC	Normalisation scale for neutral-current meson exchange current interactions
NormCCCOH	Normalisation scale for charged-current coherent interactions
NormNCCOH	Normalisation scale for neutral-current coherent interactions
RDecBR1eta	Branching fraction scaling factor for resonant decays with a single η
RDecBR1gamma	Branching fraction scaling factor for resonant decays with a single γ
RPA_CCQE	Variation in the random phase approximation for charged-current quasi-elastic interactions
ThetaDelta2NRad	Reweight the γ angular distribution from isotropic to $\propto \cos^2\theta$
Theta_Delta2Npi	Reweight the π angular distribution from isotropic to the Rein-Sehgal prediction
VecFFCCQEshape	Choice of the shape of the vector form factor for charged-current quasi-elastic interactions

Table 8.3: The individual parameters available to reweight the GENIE interaction model [275].

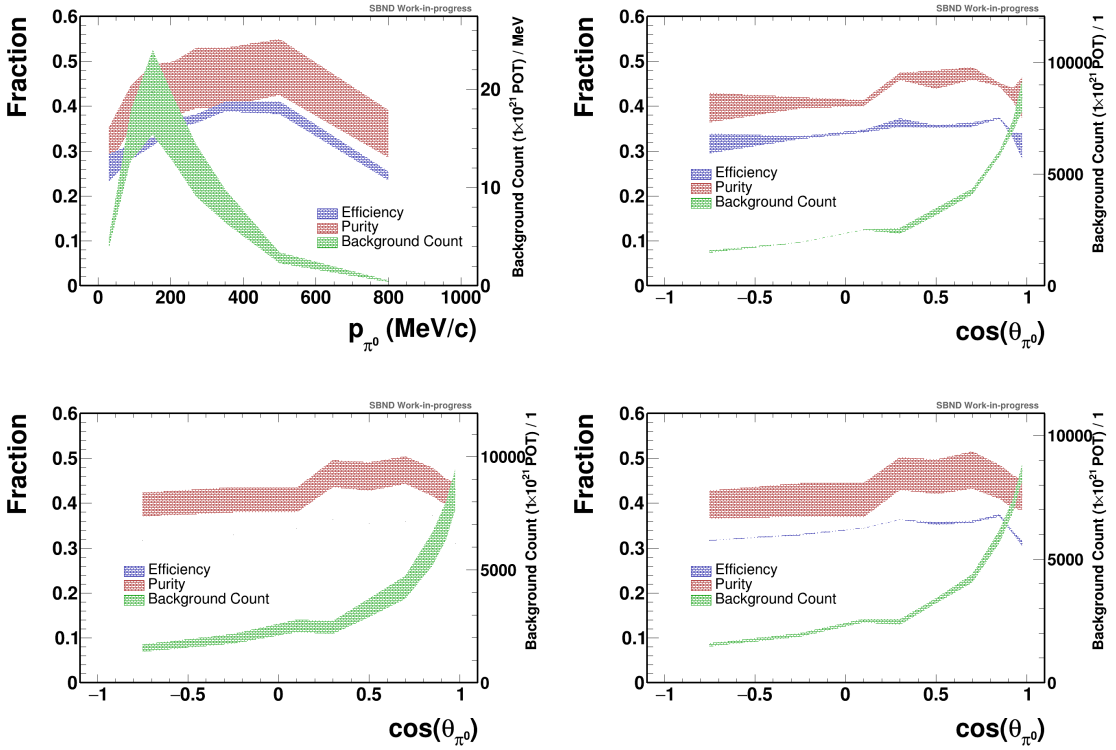


Figure 8.13: The effect of interaction systematics on the efficiency, purity and background count for the inclusive NC $1\pi^0$ selection. The top left plot shows the combination of all the GENIE interaction systematic parameters, distributed according to the reconstructed π^0 momentum. The top right, bottom left and bottom right show the `Theta_Delta2N`π, `CCRESVariationResponse` and `NCRESVariationResponse` parameters respectively, all distributed according to the reconstructed $\cos\theta_{\pi^0}$.

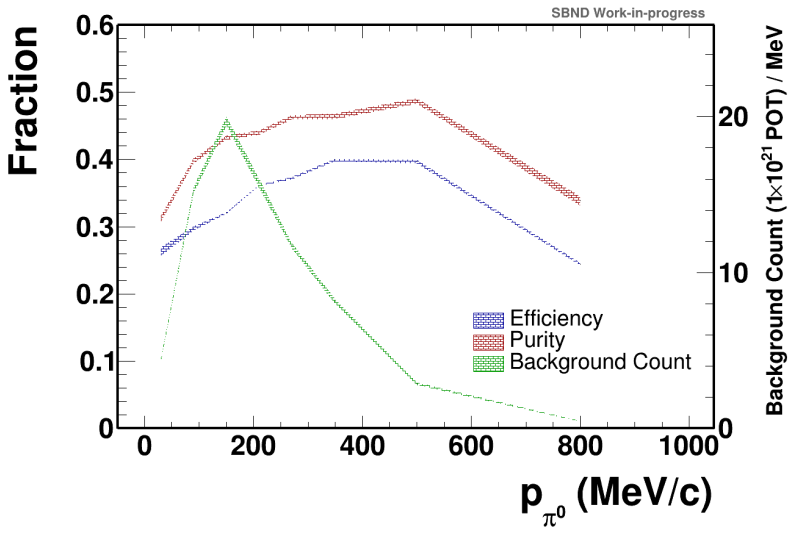


Figure 8.14: The effect of reinteraction systematics on the efficiency, purity and background count for the inclusive NC $1\pi^0$ selection in terms of the reconstructed π^0 momentum.

the Geant4 software, and a recent tool was produced by members of the neutrino physics community to reweight events in neutrino event simulations according to changes in the hadronic interaction probabilities [277]. The parameters considered relate to the reinteractions of protons and charged pions and are listed in table 8.4. These uncertainties are much smaller, typically of order 1 %, than those created by the flux and interaction modelling, as indicated by figure 8.14.

8.2.2 Detector

Systematic uncertainties relating to the modelling of the detector response cannot be approached via a reweighting procedure. A different modelling of the recombination effect, for example, will not result in a direct increase or decrease in the probability of event X occurring. It will, however, affect the number of electrons and scintillation photons produced, which itself can then affect the presence and size of hits in both

Parameter	**Description**
π^+	
PiPlusReacLow	Total inelastic scattering at low momentum
PiPlusReacHigh	Total inelastic scattering at high momentum
PiPlusAbs	Absorption
PiPlusCex	Charge exchange
PiPlusDCex	Double charge exchange
PiPlusPiProd	Pion production
PiPlusElast	Total elastic scattering
π^-	
PiMinusReacLow	Total inelastic scattering at low momentum
PiMinusReacHigh	Total inelastic scattering at high momentum
PiMinusAbs	Absorption
PiMinusCex	Charge exchange
PiMinusDCex	Double charge exchange
PiMinusPiProd	Pion production
PiMinusElast	Total elastic scattering
p	
ProtonReac	Total inelastic scattering
ProtonElast	Total elastic scattering

Table 8.4: The parameters available to reweight the Geant4 hadronic reinteraction model [277].

the TPC & PMT systems, and through them the quality of the pattern recognition and energy reconstruction, all of which can ultimately result in a different outcome during a selection procedure, or the assignment of an event to a different kinematic bin. Given the complexity of this class of systematic they typically have to be assessed using full re-simulation and reconstruction of a common set of neutrino events, under a tweaked model. The key effects that can introduce uncertainties into the detector simulation are recombination, diffusion, space charge, electron attenuation, scintillation light propagation and electronics responses. MicroBooNE recently developed a novel approach to account for some of the TPC detector systematics by parameterising variations in terms of their effect on the wire waveform rather than directly varying the underlying model for each effect [278].

SBND is currently in the process of developing the computing infrastructure required for assessing the impact of a tweaked recombination or diffusion model, but these tools are not yet available for rigorous use. Further development of these, and other, detector systematics also relies on the first round of detector calibration data to assess the performance of the SBND systems. As such, no detector systematics will be assessed in this work, although they are acknowledged to likely be one of the more significant uncertainties, alongside those relating to the flux and interaction models.

8.2.3 POT

The flux prediction (section 5.1.1) and associated uncertainties (section 8.2.1.1) are determined relative to the number of protons on target. Any uncertainty in the knowledge of this value will present as a flat normalisation uncertainty in the final neutrino event prediction. The protons arriving at the beam hall are measured by a pair of toroids upstream of the target, these toroids, and their associated calibration, are able to measure the proton rate to within 2 %. A further uncertainty is associated

with the focusing of the protons onto the beryllium target (beam optics) but this only affects the measured POT to less than 1 % and, as such, is negligible when added to the proton delivery uncertainty in quaderature [181]. The POT uncertainty used for the studies presented in this chapter is a flat 2 %.

8.2.4 Number of Targets

The number of targets on which the interaction occurs (equation 8.13) is calculated from generally well known quantities. One cause of uncertainty will be the precision of the temperature monitors, as the argon density is temperature dependent, however this will be a sub-1 % effect [279]. Another uncertainty will be the manner in which the space charge effect can change the effective fiducial volume. SBND reconstruction currently has a correction stage to account for space charge offsets, and until calibration data is taken later this year it is difficult to estimate what the size of any remaining uncertainty will be. A conservative estimate of 1 % is used for the purposes of this study. As with the POT uncertainty, this is applied as a flat normalisation uncertainty [245, 280].

8.3 Extracting the Cross Section

Having now considered the various inputs that form a cross section measurement and the systematic uncertainties that limit our current modelling, we can demonstrate the extraction of the "cross section" of our Monte Carlo sample using equation 8.5. Figure 8.15 shows this extraction for all three channels and the contributing systematic uncertainties. Detailed breakdowns of the inputs to the calculation are shown in table 8.5 and of the systematic and statistical uncertainties in table 8.6. A conscious choice was made not to include the statistical uncertainty of the sample used to produce this prediction in the plots and later model comparisons. In the

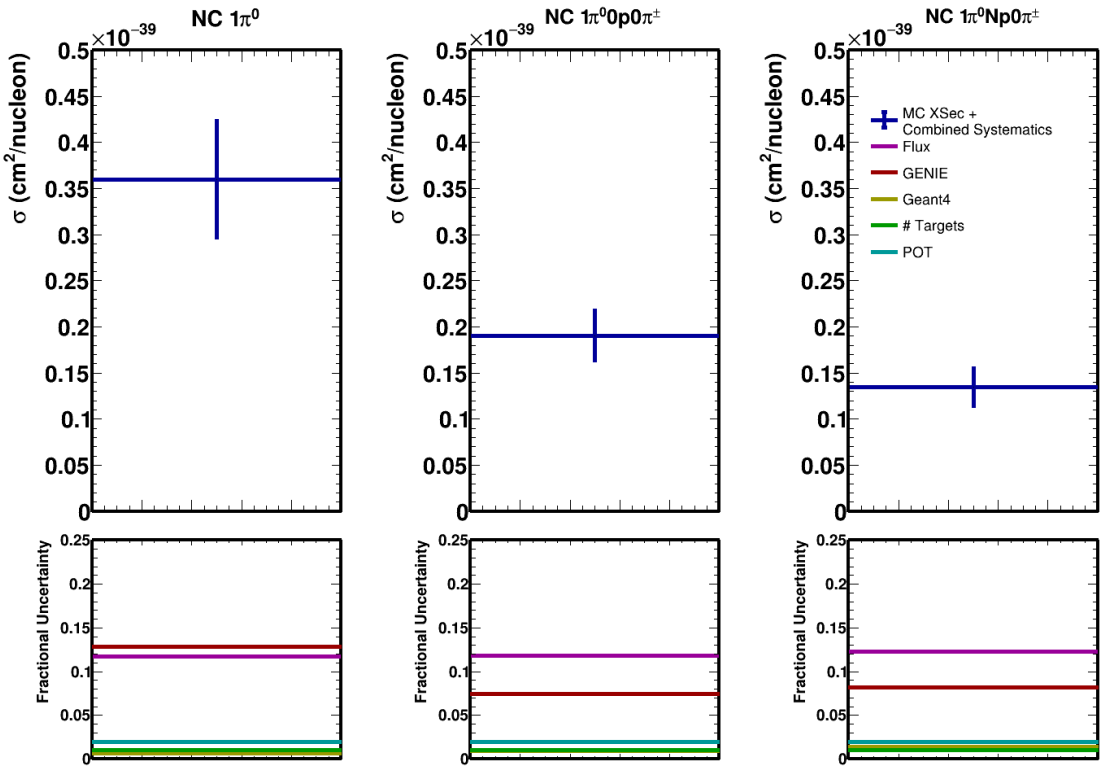


Figure 8.15: The extracted Monte Carlo cross section for each of the three NC $1\pi^0$ channels with all assessed systematic uncertainties and the breakdown of the different sources of this uncertainty.

integrated space it does not form a significant contribution when compared to the systematic uncertainty and the sample used is much smaller than the expected SBND exposure so would not faithfully reproduce the likely scale of statistical uncertainties. It must, however, be borne in mind when considering the conclusions of this study. In all three channels the flux and interaction systematics are by far the dominant ones, although it is noticeable the purer selections of the two exclusive samples result in a significant drop in the effect of interaction systematics.

For the inclusive channel the sample used does still have sufficient statistics to examine this extraction in terms of the two reconstructed observables evaluated earlier. This is shown in figure 8.16. In general, the systematic uncertainties are relatively flat with respect to these variables, although we again see the impact that

Parameter	NC $1\pi^0$	NC $1\pi^0 0p 0\pi^\pm$	NC $1\pi^0 \mathrm{N}p 0\pi^\pm$
Scaled Selected Candidates	220,519	61,229	69,945
Purity (%)	43.50	53.47	47.98
Efficiency (%)	34.38	22.16	32.06
# Targets	4.6712×10^{31}	4.6712×10^{31}	4.6712×10^{31}
Integrated Flux (cm^{-2})	1.65974×10^{13}	1.65974×10^{13}	1.65974×10^{13}
σ_{FI} (cm^2/nucleon)	3.60×10^{40}	1.91×10^{40}	1.35×10^{40}

Table 8.5: The input values used to calculate (and the result of) the integrated cross section for each of the three channels.

lower purity has on noticeably raising the interaction uncertainty in such bins.

8.4 Model Discrimination

Finally, all the pieces are in place in order to assess the power SBND will have to make comparisons between different generator predictions for the NC $1\pi^0$ cross section. To do this, we will examine the extracted cross section according to two different generator predictions:

- GENIEv3's AR23_20i_00_000 tune - the same tune used to generate the underlying sample for the cross section extraction, acting as a closure test for the method performed.
- NuWro v21.09.2 - as the alternative model to evaluate against.

The NuWro generator was chosen as the alternative due to the fact that it models the resonant contribution to single pion production in a clearly different manner to the default GENIE tune. Whilst GENIE leverages the Berger-Sehgal model,

Uncertainty Source	NC $1\pi^0$ (%)	NC $1\pi^0 0p 0\pi^\pm$ (%)	NC $1\pi^0 \mathrm{N}p 0\pi^\pm$ (%)
Flux	11.7	11.8	12.3
GENIE	12.8	7.5	8.2
Geant4	0.7	0.9	1.4
# Targets	1.0	1.0	1.0
POT	2.0	2.0	2.0
Combined Systematic	17.5	14.2	15.0
Statistical (on sample)	1.5	2.6	2.6

Table 8.6: The integrated fractional uncertainties resulting from all the considered sources of uncertainty on each of the three channels.

including 16 different nucleon resonance states, NuWro considers just the dominant $\Delta(1232)$ resonance [281]. The two predictions therefore give a significant difference with which to evaluate the separation power against.

All predictions were produced using GENIE [115] and NuWro [116] builds respectively and processed in the NUISANCE framework [282]. NUISANCE is a tool designed to allow direct comparisons to be made between multiple different generators by converting each of their outputs to a shared structure. It also provides functionality for tuning cross section parameters to published datasets. Flux spectra were created using the path traced average method described in section 8.1.1. Both generator-only samples were produced with more than $10\times$ the statistics of the full detector simulated samples, this is possible due to the significantly lower computing cost and ensures that the associated statistical uncertainties are negligible.

Firstly, we can view the integrated cross section predictions for each of the three exclusive channels and compare them to the nominal extracted cross section and its associated systematic uncertainties described above. This is shown in figure 8.17, with the values provided for comparison in table 8.7. The extracted cross section

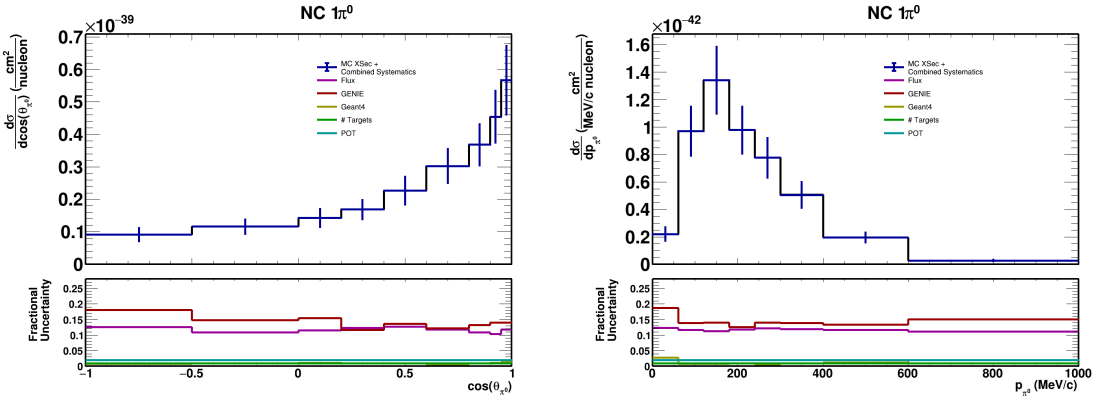


Figure 8.16: The extracted Monte Carlo cross section for the inclusive NC $1\pi^0$ channel in terms of the reconstructed $\cos\theta_{\pi^0}$ and p_{π^0} with all assessed systematic uncertainties and the breakdown of the different sources of this uncertainty.

differs around 1 % from the GENIE prediction for all three channels. This lies fully within the statistical errors of the MC sample, successfully completing the closure test. This plot clearly indicates that with the current systematic uncertainties, discriminating between different models at an integrated level is incredibly difficult. It is noticeable, however, that the sensitivity increases as we move into more exclusive channels. The most noticeable difference between the two models is that, while NuWro predicts a higher cross section for the NC $1\pi^0 0p 0\pi^\pm$ channel, the roles reverse for NC $1\pi^0 \mathrm{N}p 0\pi^\pm$ with GENIE predicting a significantly higher cross section. This can be explained by viewing figure 8.18 which shows the proton multiplicity in signal NC $1\pi^0$ interactions according to the two generator predictions. It is clear that the GENIE tune consistently predicts a large proton multiplicity and whilst that effect is greatly reduced by applying the momentum threshold used in this analysis, it is still enough to make large differences in the exclusive channel predictions. Work to identify lower momentum protons in LArTPC reconstruction will be very beneficial to understanding the multiplicity and kinematics of nucleons emitted in neutrino-nuclear interactions.

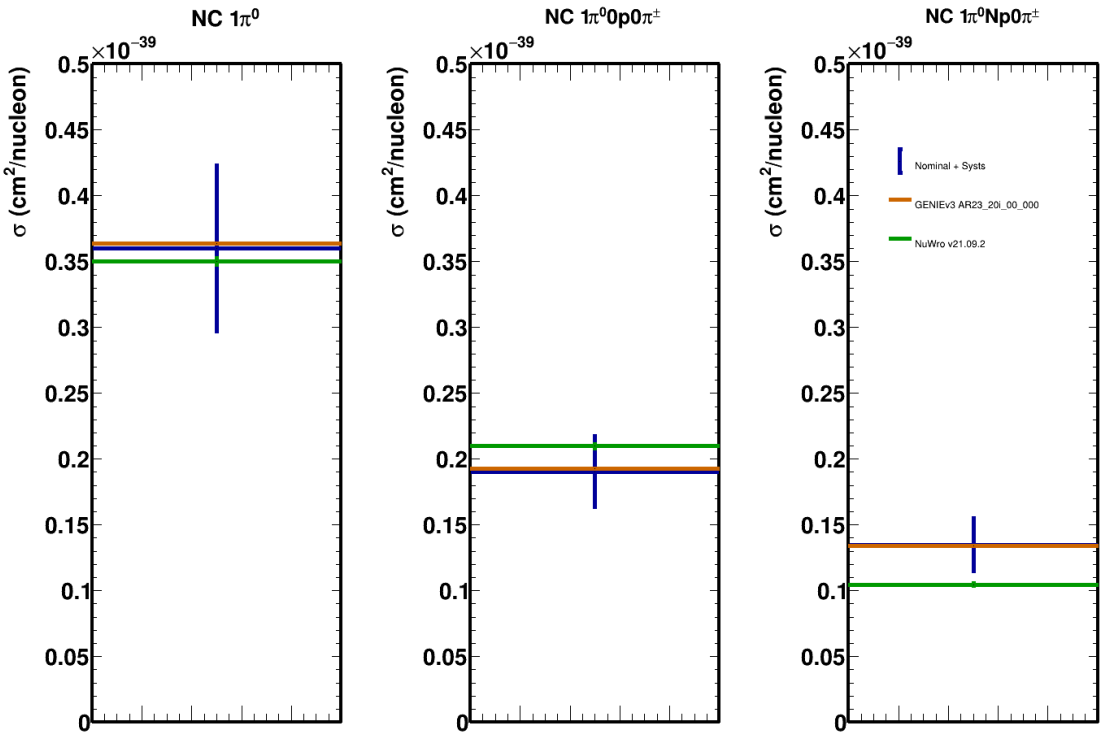


Figure 8.17: The extracted Monte Carlo cross section for each of the three NC $1\pi^0$ channels with all assessed systematic uncertainties and comparisons made to predictions from GENIEv3 AR23_20i_00_000 and NuWro v21.09.2.

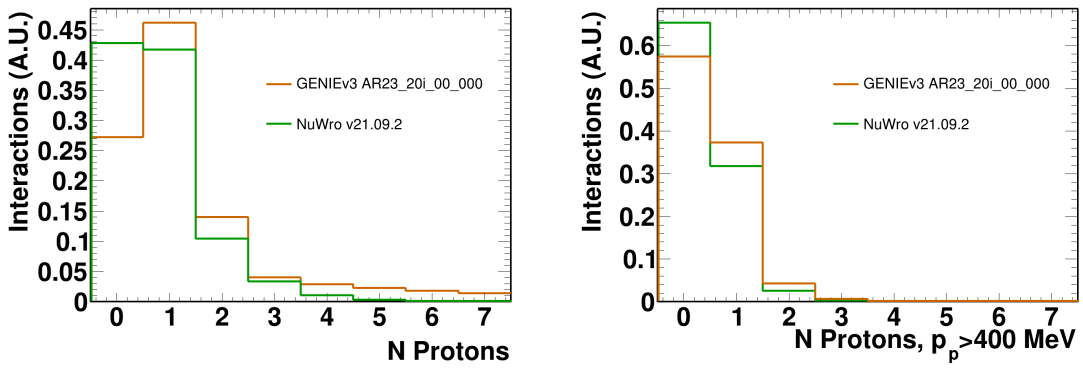


Figure 8.18: The predicted proton multiplicity in NC $1\pi^0$ events according to GENIEv3 AR23_20i_00_000 and NuWro v21.09.2 with (right) and without (left) the 400 MeV/c threshold applied in the analysis signal definition.

Generator Prediction (cm^2/nucleon)	NC $1\pi^0$	NC $1\pi^0 0p 0\pi^\pm$	NC $1\pi^0 Np 0\pi^\pm$
GENIEv3 AR23_20i_00_000	3.64×10^{40}	1.93×10^{40}	1.34×10^{40}
NuWro v21.09.2	3.50×10^{40}	2.10×10^{40}	1.05×10^{40}

Table 8.7: The integrated cross section predictions for the two generator tunes used in the comparison plots.

Evaluating the two different exclusive channels gives access to subtler model differences than just the inclusive channel. In a similar vein, we can also evaluate the generator predictions in the previously presented momentum and direction spaces to provide a finer-grain assessment. In order to assess the predictions in the reconstructed space, the forward folding technique laid out in equation 8.7 is used. The relevant matrices are shown in figure 8.19 and, as discussed in section 8.1.3, demonstrate the better reconstruction of direction than momentum, and a slight bias towards under-estimating momentum, even after the corrections made in the analysis stage. Figure 8.20 shows these forward folded comparisons for the NC $1\pi^0$ channel with all systematic uncertainties shown on the extracted cross section. The statistical errors are included on the folding matrices and later propagated into the folded predictions. The systematic uncertainties on the folding matrices can be calculated by producing the matrices in each universe but were significantly smaller than the statistical errors and excluded from here onwards. This shows another scenario in which the limited size of the available Monte Carlo samples provides a restriction to this study that should be made insignificant for full SBND measurements.

The agreement between the extracted Monte Carlo cross section and the two generator predictions can be made by assessing the χ^2 between them:

$$\chi^2 = \sum_{ij}(O_i - P_i)V_{ij}^{-1}(O_j - P_j) \tag{8.21}$$

where O are the observed data points (in this case the extracted Monte Carlo values),

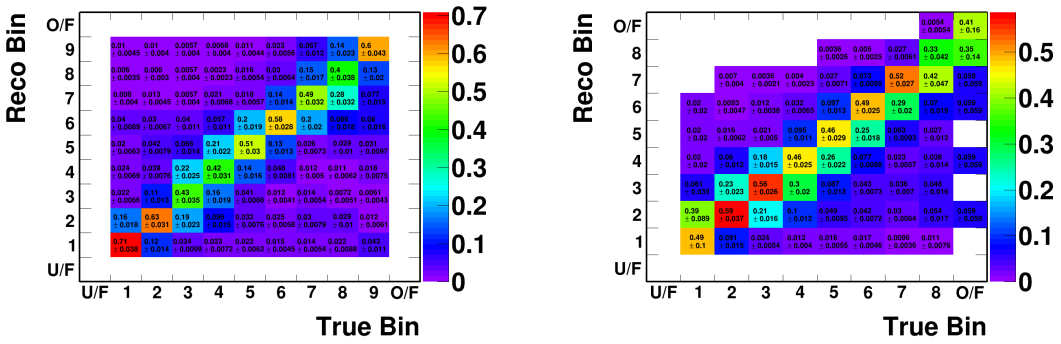


Figure 8.19: The forward folding matrices expressing the smearing relationship between the true and reconstructed directions (left) and momenta (right) of selected signal NC $1\pi^0$ interactions as a result of reconstruction imperfections.

P are the predicted values for each bin and V is the covariance matrix encoding the systematic uncertainties and their correlated effects on multiple bins. The covariance matrix can be constructed for each universe based systematic as

$$V_{ij} = \frac{1}{N}\sum_{n}^{N}(P_i^n - P_i^{CV})(P_j^n - P_j^{CV}), \tag{8.22}$$

where the N universes are indexed via n and the predicted values in each universe P^n are compared to those of the central value P^{CV}. The covariance matrices for the key systematics are found in appendix F.

The GENIE prediction in figure 8.20 again shows very good agreement in both variables, demonstrating the accuracy of the technique used to make the comparisons. The NuWro prediction shows a significant divergence in lower momentum and very forward angles. Whilst by eye the NuWro prediction still appears to be consistent with the nominal within the calculated uncertainties, the χ^2 of 20.37/8 indicates a significant discrepancy in the momentum space. This shows the importance of the correlation between the binned effects of systematic uncertainties. Figure 8.21 shows the systematic error bars decomposed into their shape and normalisation components, using a method from [283] which is described in appendix G. It is clear that the majority of the systematic uncertainty is a normalisation effect.

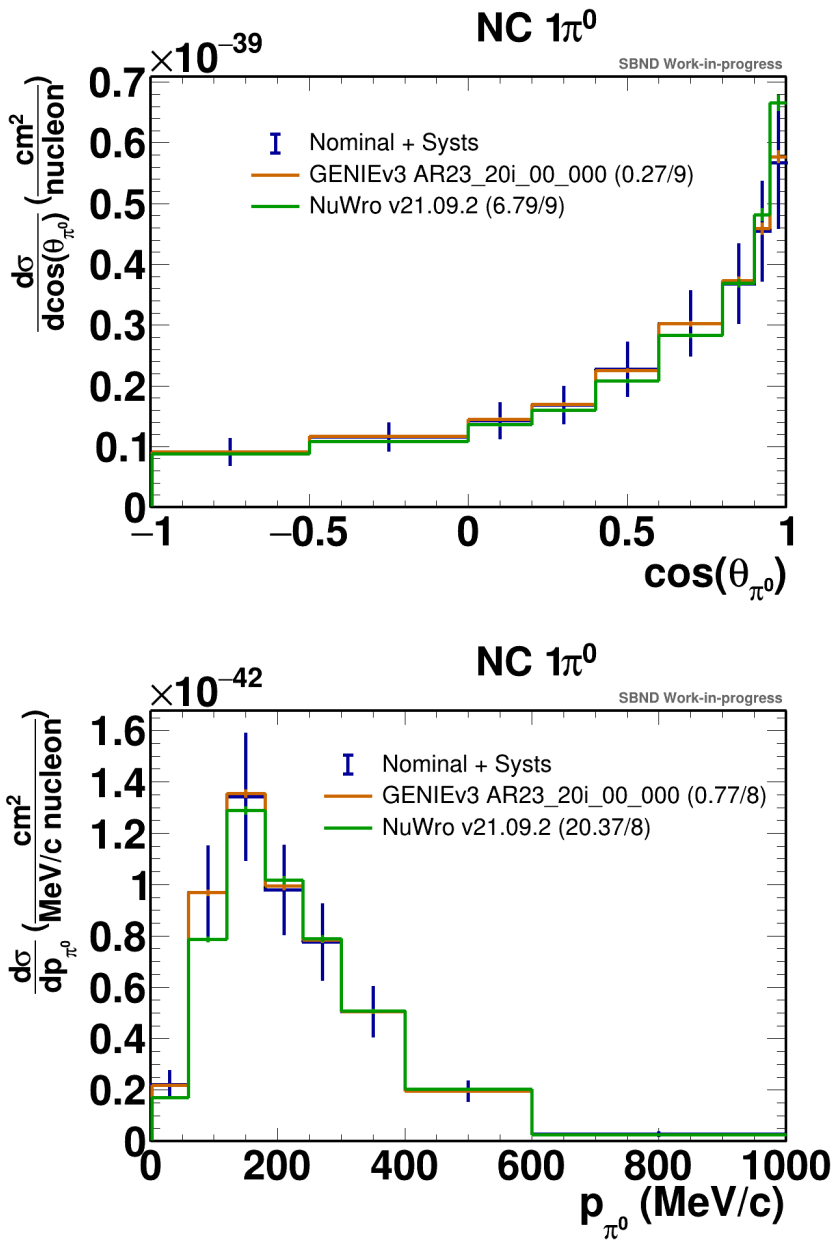


Figure 8.20: The extracted Monte Carlo cross section for the inclusive NC $1\pi^0$ channel in terms of the reconstructed $\cos\theta_{\pi^0}$ (left) and p_{π^0} (right) with all assessed systematic uncertainties. Comparisons are made to forward folded predictions from GENIEv3 AR23_20i_00_000 and NuWro v21.09.2.

If we consider the two largest contributions - flux and interaction - then this is understandable. The flux uncertainties are almost entirely normalisation based due to the large uncertainty in the integrated flux used to normalise the cross section, whilst the interaction systematics also have significant normalisation uncertainties from parameters that purely affect the strength of a given interaction class. The interaction systematics do also introduce shape uncertainties, most significantly through the FSI parameters which dominate by altering the kinematics of the observed particles.

It is clear from the shape uncertainties that no normalisation shift could bring the NuWro prediction into agreement with the nominal in the momentum space, hence the larger χ^2. The value of 20.37/8 would correspond to a rejection of the alternative model at a 97 % level. In the direction space the shape uncertainties are a little larger in the forward direction, reducing the sensitivity to differences.

The results presented in this chapter indicate that with the current SBND reconstruction and analysis tools the experiment will have some sensitivity to discriminate between different models of NC $1\pi^0$ interactions.

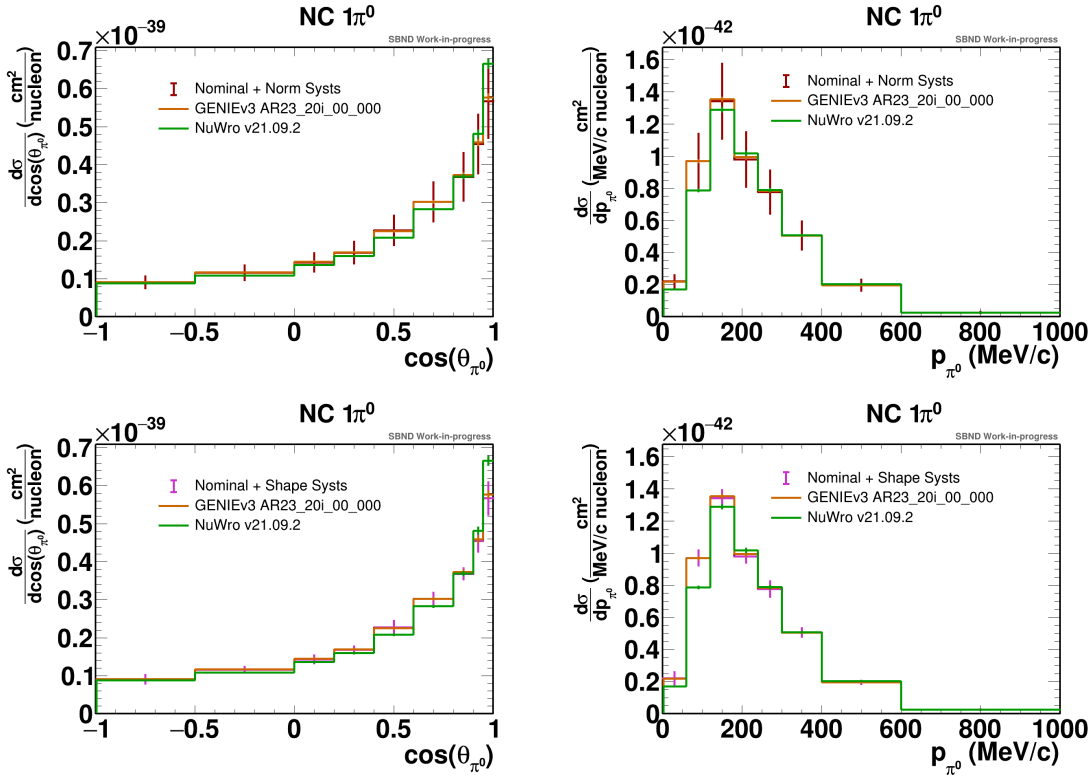


Figure 8.21: The extracted Monte Carlo cross section for the inclusive NC $1\pi^0$ channel in terms of the reconstructed $\cos\theta_{\pi^0}$ (left) and p_{π^0} (right) with uncertainties separated into normalisation (top) and shape (bottom) components. Comparisons are made to forward folded predictions from GENIEv3 AR23_20i_00_000 and NuWro v21.09.2.

Chapter 9

Summary and Outlook

This thesis has presented a detailed evaluation of the status of the Short-Baseline Near Detector as it moves towards taking first physics data. A number of significant improvements to the reconstruction for both the TPC and CRT systems, and the development of a new cosmic rejection tool, harnessing all three of the detector's sub-systems, provide a stronger foundation for many physics measurements to build upon. In particular, the potential model discrimination power of SBND's cross section program has been evaluated in the context of a trio of NC $1\pi^0$ channels. Powerful selections reduced backgrounds in these challenging channels by $> 99\,\%$ and correction techniques improved the resolution of the neutral pion momentum and direction reconstruction, allowing analysis of the inclusive channel in a binned kinematic space. Finally, a number of the systematic uncertainties that limit the power of the measurement have been assessed and the extracted cross section compared to a pair of models. The results indicate that in a scenario of significant difference between the underlying model used to make the corrections to the measurement and the true description of these interactions, a significant measurement can be made in the momentum space, and a strong indication could be found in the direction space.

Whilst this assessment uses the best available tools at the time of writing, a number of factors will modify this assessment before SBND data is used to make a measurement of the NC $1\pi^0$ cross section. Foremost among these are:

- The production and use of much larger Monte Carlo samples for the nominal and systematic universe predictions, such that their finite statistics are no longer a relevant limit to the conclusions.
- The inclusion of detector systematics based on a growing understanding of the detector's performance during commissioning and calibration.
- Attempts to reduce the systematic uncertainties associated with the neutrino flux modelling, partially through the external collection of more relevant pion production data.
- The ability to develop sideband studies to constrain the largest background categories. In this scenario the obvious example would be a CC ν_μ π^0 sideband.

Some of these changes will reduce the measurement's sensitivity, particularly that of the detector systematics, whilst others have the ability to make improvements, most significantly the opportunity to incorporate modern pion production data into the flux prediction. In summary, this thesis shows that SBND presents a generational opportunity to improve our measurements of neutrino-argon interactions, but that work to better constrain our flux, interaction and detector models will be critical to exploiting this opportunity in a systematic dominated landscape.

Appendix A

Matrix Least Squares Approach

Consider a set of line equations

$$\vec{x}_i = a_i \cdot \vec{u}_i + \vec{x}_{0,i} \tag{A.1}$$

where $\vec{x}_i$ is the general point on that line, a_i is a running coefficient, $\vec{u}_i$ is the unit direction vector of the line and $\vec{x}_{0,i}$ is a particular 'zero-point' on that line.

Following the approach in [252, 253], such a set of line equations can be assembled into a matrix equation

$$G \cdot m = d, \tag{A.2}$$

where

$$G = \begin{pmatrix} 1 & 0 & 0 & -u_{x,1} & 0 & 0 & ... & 0 \\ 0 & 1 & 0 & -u_{y,1} & 0 & 0 & ... & 0 \\ 0 & 0 & 1 & -u_{z,1} & 0 & 0 & ... & 0 \\ 1 & 0 & 0 & 0 & -u_{x,2} & 0 & ... & 0 \\ 0 & 1 & 0 & 0 & -u_{y,2} & 0 & ... & 0 \\ 0 & 0 & 1 & 0 & -u_{z,2} & 0 & ... & 0 \\ & & & & ... & & & \\ 1 & 0 & 0 & 0 & 0 & 0 & ... & -u_{x,n} \\ 0 & 1 & 0 & 0 & 0 & 0 & ... & -u_{y,n} \\ 0 & 0 & 1 & 0 & 0 & 0 & ... & -u_{z,n} \end{pmatrix}, \tag{A.3}$$

$$d = \begin{pmatrix} x_{0,1} \\ y_{0,1} \\ z_{0,1} \\ x_{0,2} \\ y_{0,2} \\ z_{0,2} \\ ... \\ x_{0,n} \\ y_{0,n} \\ z_{0,n} \end{pmatrix}, \qquad m = \begin{pmatrix} x \\ y \\ z \\ a_1 \\ a_2 \\ ... \\ a_n \end{pmatrix}.$$

Consider the example where n = 2:

$$G = \begin{pmatrix} 1 & 0 & 0 & -u_{x,1} & 0 \\ 0 & 1 & 0 & -u_{y,1} & 0 \\ 0 & 0 & 1 & -u_{z,1} & 0 \\ 1 & 0 & 0 & 0 & -u_{x,2} \\ 0 & 1 & 0 & 0 & -u_{y,2} \\ 0 & 0 & 1 & 0 & -u_{z,2} \end{pmatrix}, \qquad d = \begin{pmatrix} x_{0,1} \\ y_{0,1} \\ z_{0,1} \\ x_{0,2} \\ y_{0,2} \\ z_{0,2} \end{pmatrix}, \qquad m = \begin{pmatrix} x \\ y \\ z \\ a_1 \\ a_2 \end{pmatrix}. \tag{A.4}$$

Then

$$G \cdot m - d = \begin{pmatrix} x - a_1 \cdot u_{x,1} - x_{0,1} \\ y - a_1 \cdot u_{y,1} - y_{0,1} \\ z - a_1 \cdot u_{z,1} - z_{0,1} \\ x - a_2 \cdot u_{x,2} - x_{0,2} \\ y - a_2 \cdot u_{y,2} - y_{0,2} \\ z - a_2 \cdot u_{z,2} - z_{0,2} \end{pmatrix}, \tag{A.5}$$

where we can see the line equations being recovered. The equation is setup with a shared general point ($\vec{x} = (x, y, z)$). If the two lines intersect then $G \cdot m - d = 0$ has an exact solution. If not, then we want to find the least squares solution which is the point, $\vec{x'}$ which satisfies

$$||G \cdot m' - d|| \leq ||G \cdot m - d||, \tag{A.6}$$

where m contains the general point, $\vec{x}$, and m' contains the point $\vec{x'}$. Considering

equation A.5 we see

$$G \cdot m - d = \begin{pmatrix} x - a_1 \cdot u_{x,1} - x_{0,1} \\ y - a_1 \cdot u_{y,1} - y_{0,1} \\ z - a_1 \cdot u_{z,1} - z_{0,1} \\ x - a_2 \cdot u_{x,2} - x_{0,2} \\ y - a_2 \cdot u_{y,2} - y_{0,2} \\ z - a_2 \cdot u_{z,2} - z_{0,2} \end{pmatrix}$$

$$= \begin{pmatrix} l_{x,1} \\ l_{y,1} \\ l_{z,1} \\ l_{x,2} \\ l_{y,2} \\ l_{z,2} \end{pmatrix} \tag{A.7}$$

$$||G \cdot m - d|| = \sqrt{l_{x,1}^2 + l_{y,1}^2 + l_{z,1}^2 + l_{x,2}^2 + l_{y,2}^2 + l_{z,2}^2} \,,$$

where $l_{i,n} = i - a_n \cdot u_{i,n} - i_{0,n}$ is the distance in dimension i between the point $\vec{x}$ and the n^{th} line. Therefore, the requirement in equation A.6 corresponds to minimising the sum of the squared distances from the point in question to the nearest point on each line.

Following [284] we find that such a point corresponds to the exact solutions of the equation

$$G^T \cdot G \cdot m = G^T \cdot d \tag{A.8}$$

and therefore

$$m = (G^T \cdot G)^{-1} \cdot G^T \cdot d. \tag{A.9}$$

Note that a weighting, w_i can be applied to each line such that the setup in equation A.4 becomes

$$G = \begin{pmatrix} w_1 & 0 & 0 & -w_1 \cdot u_{x,1} & 0 \\ 0 & w_1 & 0 & -w_1 \cdot u_{y,1} & 0 \\ 0 & 0 & w_1 & -w_1 \cdot u_{z,1} & 0 \\ w_2 & 0 & 0 & 0 & -w_2 \cdot u_{x,2} \\ 0 & w_2 & 0 & 0 & -w_2 \cdot u_{y,2} \\ 0 & 0 & w_2 & 0 & -w_2 \cdot u_{z,2} \end{pmatrix}, \tag{A.10}$$

$$d = \begin{pmatrix} w_1 \cdot x_{0,1} \\ w_1 \cdot y_{0,1} \\ w_1 \cdot z_{0,1} \\ w_2 \cdot x_{0,2} \\ w_2 \cdot y_{0,2} \\ w_2 \cdot z_{0,2} \end{pmatrix}, \qquad m = \begin{pmatrix} x \\ y \\ z \\ a_1 \\ a_2 \end{pmatrix},$$

so

$$G \cdot m - d = \begin{pmatrix} w_1 \cdot x - w_1 \cdot a_1 \cdot u_{x,1} - w_1 \cdot x_{0,1} \\ w_1 \cdot y - w_1 \cdot a_1 \cdot u_{y,1} - w_1 \cdot y_{0,1} \\ w_1 \cdot z - w_1 \cdot a_1 \cdot u_{z,1} - w_1 \cdot z_{0,1} \\ w_2 \cdot x - w_2 \cdot a_2 \cdot u_{x,2} - w_2 \cdot x_{0,2} \\ w_2 \cdot y - w_2 \cdot a_2 \cdot u_{y,2} - w_2 \cdot y_{0,2} \\ w_2 \cdot z - w_2 \cdot a_2 \cdot u_{z,2} - w_2 \cdot z_{0,2} \end{pmatrix} \tag{A.11}$$

and

$$||G \cdot m - d|| = \sqrt{w_1^2 \cdot (l_{x,1}^2 + l_{y,1}^2 + l_{z,1}^2) + w_2^2 \cdot (l_{x,2}^2 + l_{y,2}^2 + l_{z,2}^2)} \; . \tag{A.12}$$

This results in the solution point being closer to lines with a greater weighting.

Appendix B

CRUMBS Inputs

B.1 TPC

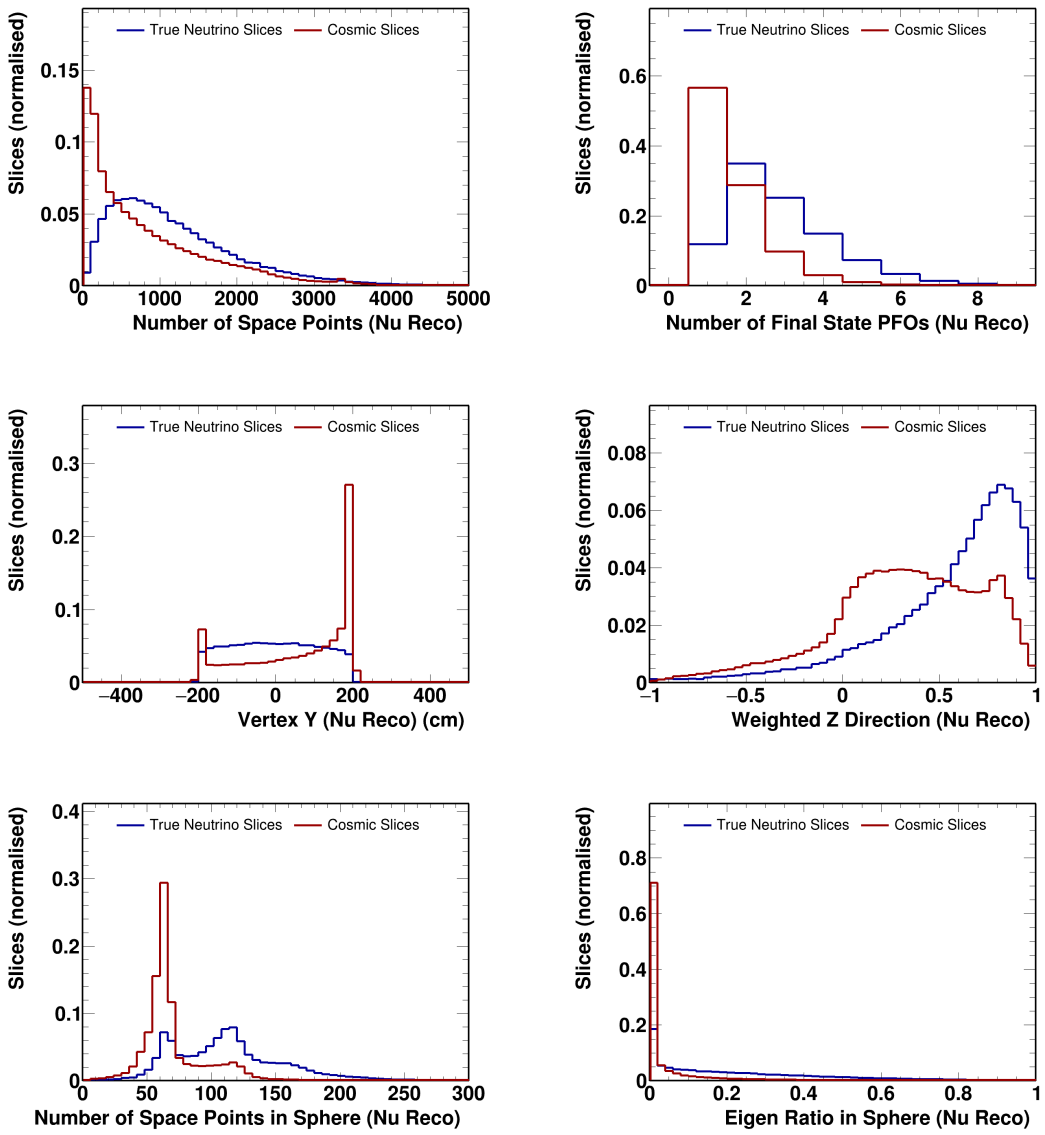


Figure B.1: The inputs to CRUMBS from the neutrino hypothesis Pandora TPC reconstruction.

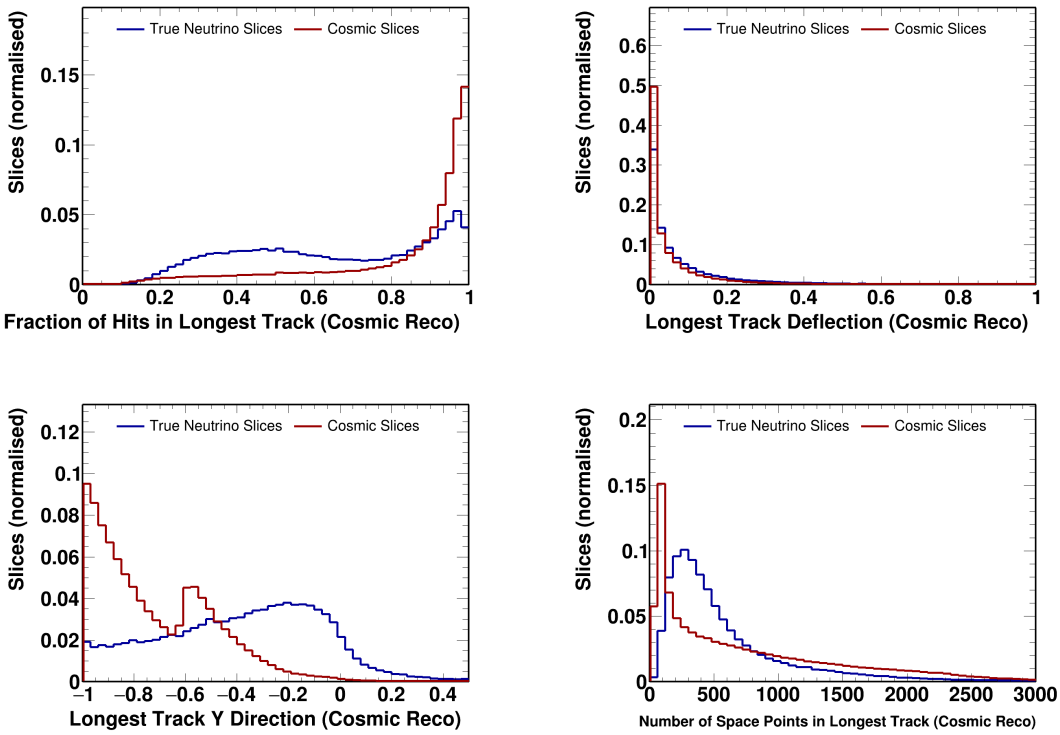


Figure B.2: The inputs to CRUMBS from the cosmic hypothesis Pandora TPC reconstruction.

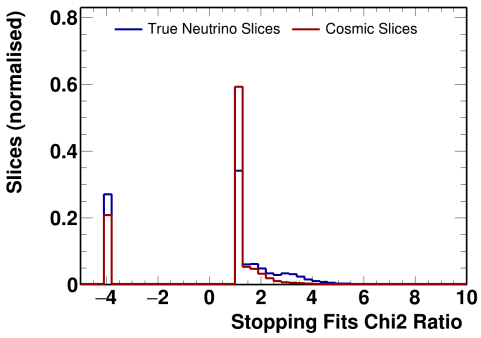


Figure B.3: The ratio between the 0^{th}-order polynomial and exponential fits to the $\mathrm{d}E/\mathrm{d}x$ - residual range distribution for the longest track in the slice.

B.2 PDS

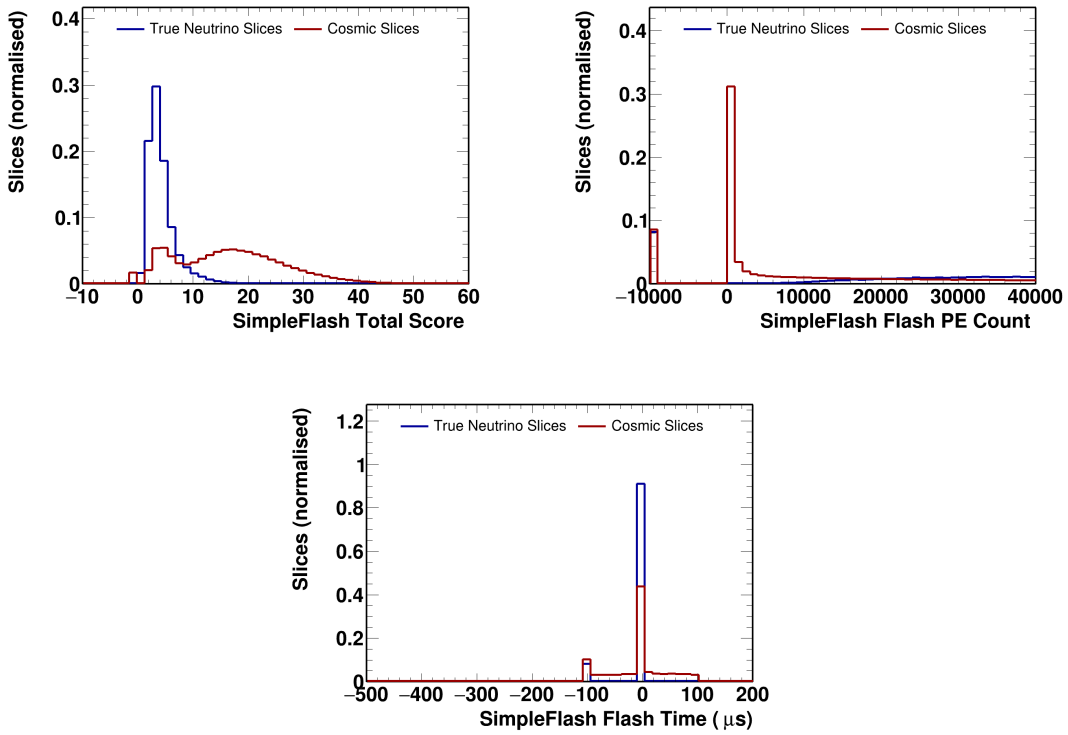


Figure B.4: The inputs to CRUMBS from the SimpleFlash flash matching algorithm.

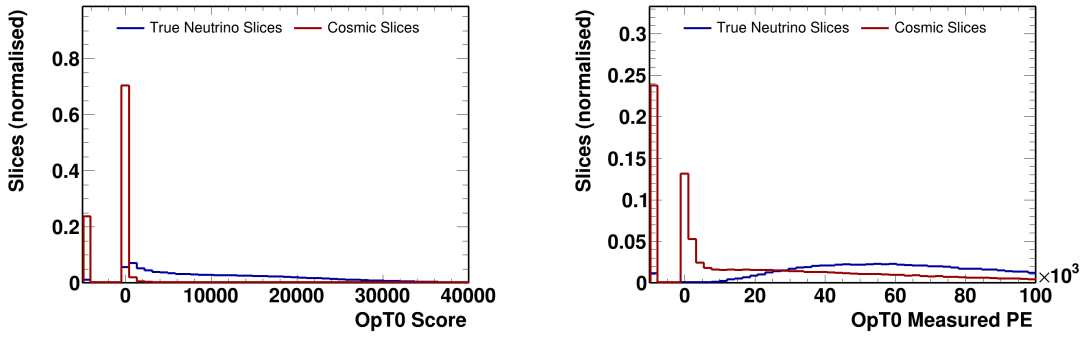


Figure B.5: The inputs to CRUMBS from the OpT0Finder flash matching algorithm.

B.3 CRT

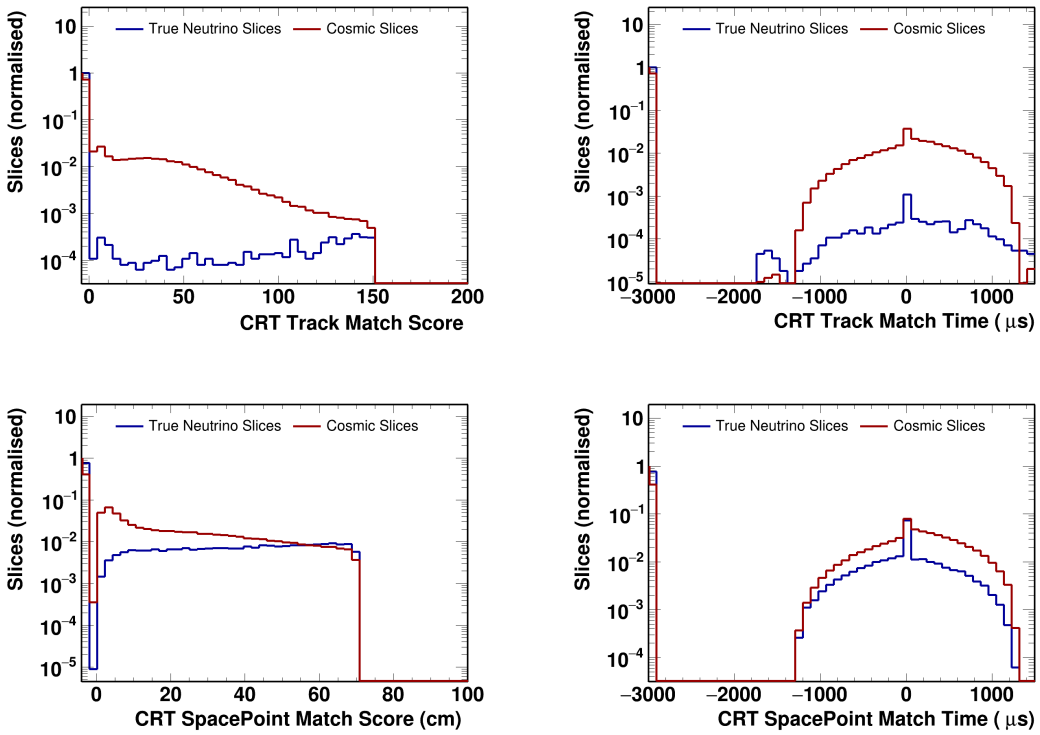


Figure B.6: The inputs to CRUMBS from the CRT-TPC matching algorithms.

Appendix C

Selection Optimisation

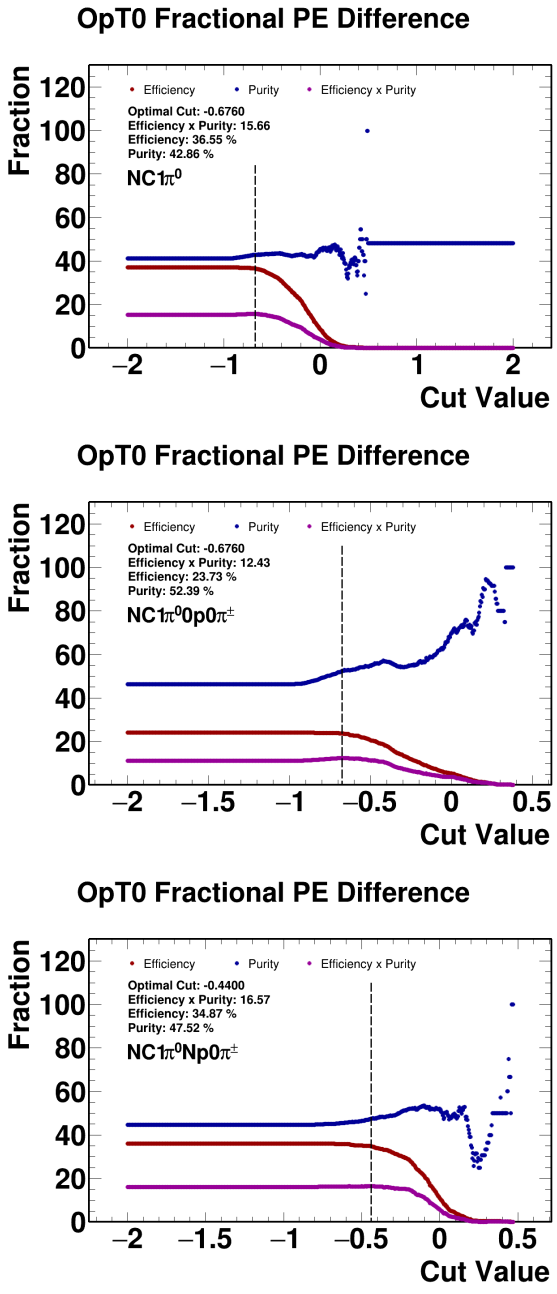


Figure C.1: Efficiency, purity and efficiency × purity of the selections (top: NC $1\pi^0$, centre: NC $1\pi^0 0p 0\pi^\pm$, bottom: NC $1\pi^0 \mathrm{N}p 0\pi^\pm$) as a function of the minimum cut on the fractional difference between the hypothesised and measured PE count from the OpT0Finder flash matching tool.

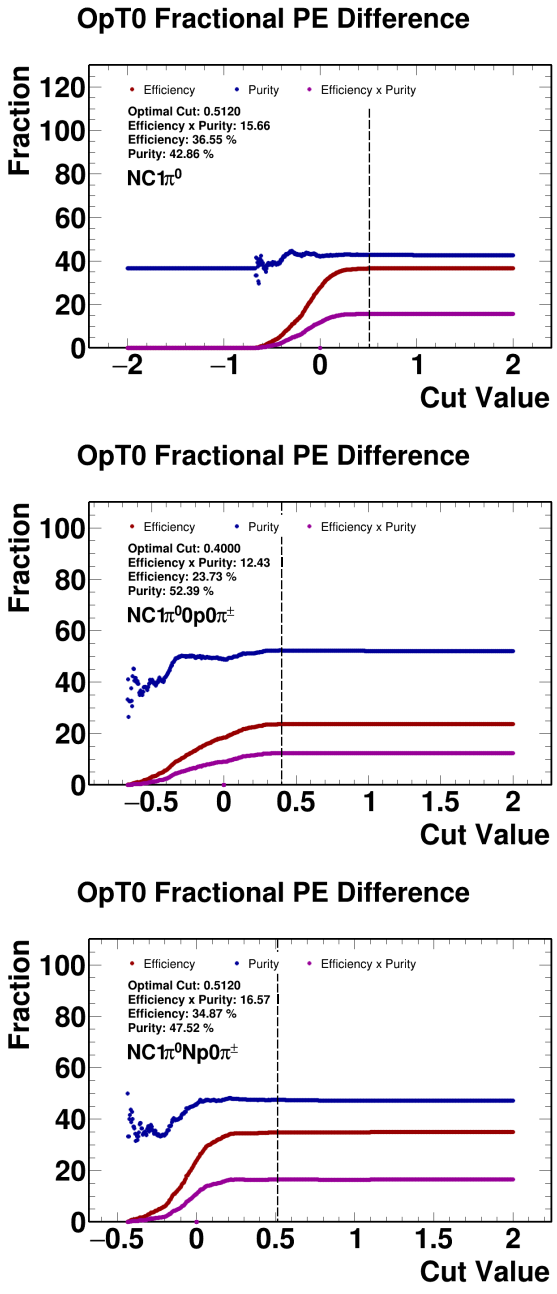


Figure C.2: Efficiency, purity and efficiency × purity of the selections (top: NC $1\pi^0$, centre: NC $1\pi^0 0p 0\pi^\pm$, bottom: NC $1\pi^0 \mathrm{N}p 0\pi^\pm$) as a function of the maximum cut on the fractional difference between the hypothesised and measured PE count from the OpT0Finder flash matching tool.

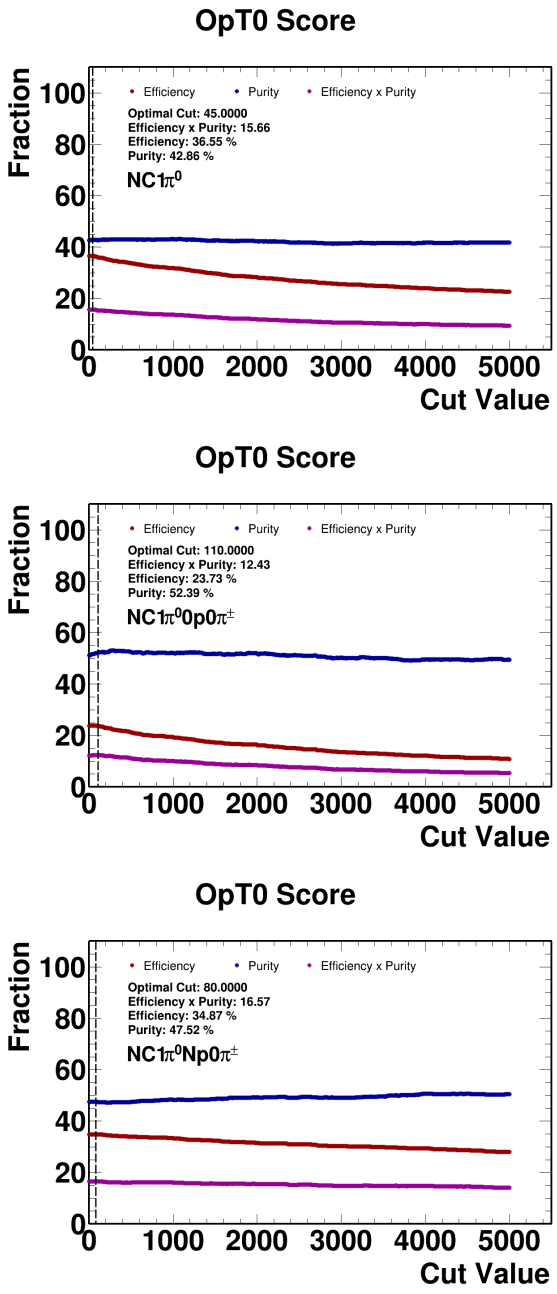


Figure C.3: Efficiency, purity and efficiency × purity of the selections (top: NC $1\pi^0$, centre: NC $1\pi^0 0p 0\pi^\pm$, bottom: NC $1\pi^0 \mathrm{N}p 0\pi^\pm$) as a function of the OpT0Finder matching score cut value.

Appendix D

Selection Tables

Cut Name	Signal (NC $1\pi^0$)	Other NC	CC $\nu_\mu\pi^0$	Other CC ν_μ	CC ν_e	Dirt	Non-FV ν	Cosmic	Mis-Reco Signal
No Cut	262167	721315	550644	4062132	35862	3472831	2128591	185168280	67847
Not Clear Cosmic	261280	614568	514493	3785069	33269	3166755	1908626	19468175	47920
FV	247383	580146	493436	3671452	31784	424126	132480	7059973	39409
CRUMBS Cut	235835	515247	445626	3248477	30366	286127	102712	256246	27994
No Razzled Muons	226415	471537	203430	969784	28326	145468	69066	207463	27573
Has Two Razzled Photons	115346	26597	104419	29745	4366	12877	8134	24148	1906
Good PiZero Kinematics	108763	23539	99276	21788	3856	11082	7225	20924	1374
Good OpT0 Frac High	108165	23428	98656	21633	3834	11038	7137	17502	1374
Good OpT0 Frac Low	105771	22763	95553	20480	3657	8688	6317	13732	421
OpT0 Score	104530	22696	94910	20413	3635	8533	6272	11389	421
All Other Tracks Contained	95930	20280	65586	14141	3125	5940	5097	10017	398

Table D.1: Number of slices per category, scaled to 1×10^{21} POT , passing each cut in the NC $1\pi^0$ selection.

Cut Name	Signal (NC $1\pi^0 0p 0\pi^\pm$)	Other NC $1\pi^0$	Other NC	CC ν_μ	CC ν_e	Dirt	Non-FV ν	Cosmic	Mis-Reco Signal
No Cut	137068	153426	721315	4612776	35862	3472831	2128591	185168280	39520
Not Clear Cosmic	136602	143496	614568	4299563	33269	3166755	1908626	19468175	29102
FV	126163	136802	580146	4164888	31784	424126	132480	7059973	23827
CRUMBS Cut	117452	128955	515247	3694104	30366	286127	102712	256246	17421
No Razzled Muons	115080	121707	471537	1173214	28326	145468	69066	207463	17200
Has Two Razzled Photons	54326	61773	26597	134164	4366	12877	8134	24148	1152
Good PiZero Kinematics	49982	59424	23539	121065	3856	11082	7225	20924	731
No Razzled Pions	46147	46391	15448	24292	2526	8467	5363	16570	687
No Razzled Protons	34555	5519	5829	8821	687	6228	4011	10350	620
Good OpT0 Frac High	34400	5452	5785	8777	687	6228	3967	8606	620
Good OpT0 Frac Low	33269	4987	5297	6272	531	4521	3346	5711	88
OpT0 Score	32737	4876	5297	6184	509	4344	3302	3887	88

Table D.2: Number of slices per category, scaled to 1×10^{21} POT, passing each cut in the NC $1\pi^0 0p 0\pi^\pm$ selection.

Cut Name	Signal (NC $1\pi^0 Np0\pi^{\pm}$)	Other NC $1\pi^0$	Other NC	CC ν_μ	CC ν_e	Dirt	Non-FV ν	Cosmic	Mis-Reco Signal
No Cut	99254	210877	721315	4612776	35862	3472831	2128591	185168280	19881
Not Clear Cosmic	99010	196093	614568	4299563	33269	3166755	1908626	19468175	14096
FV	96373	178583	580146	4164888	31784	424126	132480	7059973	11836
CRUMBS Cut	86133	125542	371706	2516833	27817	124367	59845	58468	3213
No Razzled Muons	83251	120444	336021	939218	26199	73255	41138	48583	3147
Has Two Razzled Photons	46147	57828	24292	119247	4100	8733	6095	7339	332
Good PiZero Kinematics	44640	53550	21854	108364	3657	7669	5474	6530	332
No Razzled Pions	39409	43266	14296	21455	2393	5674	3945	4564	265
Has One Razzled Proton	35663	12678	9087	14451	1795	1773	1130	2168	155
Good OpT0 Frac High	35419	12634	9021	14274	1795	1751	1130	2013	155
Good OpT0 Frac Low	33624	10705	8267	12456	1640	1086	864	1446	66
OpT0 Score	33557	10639	8267	12456	1640	1086	864	1365	66

Table D.3: Number of slices per category, scaled to 1×10^{21} POT , passing each cut in the NC $1\pi^0$N$p0\pi^{\pm}$ selection.

Appendix E

True Observables

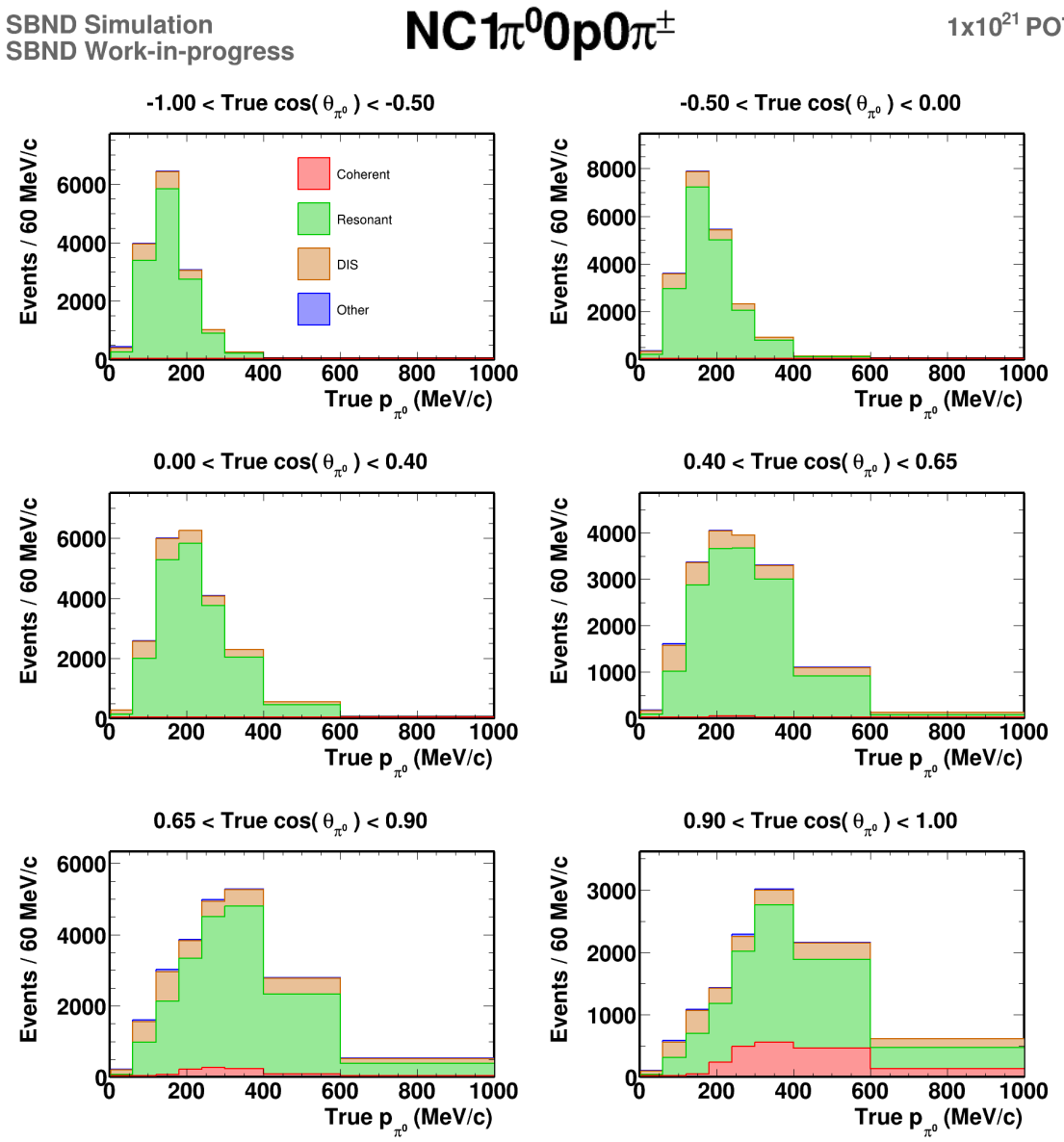


Figure E.1: Two-dimensional distribution of the neutral pion momentum for true signal events in the NC $1\pi^0 0p 0\pi^\pm$ signal definition.

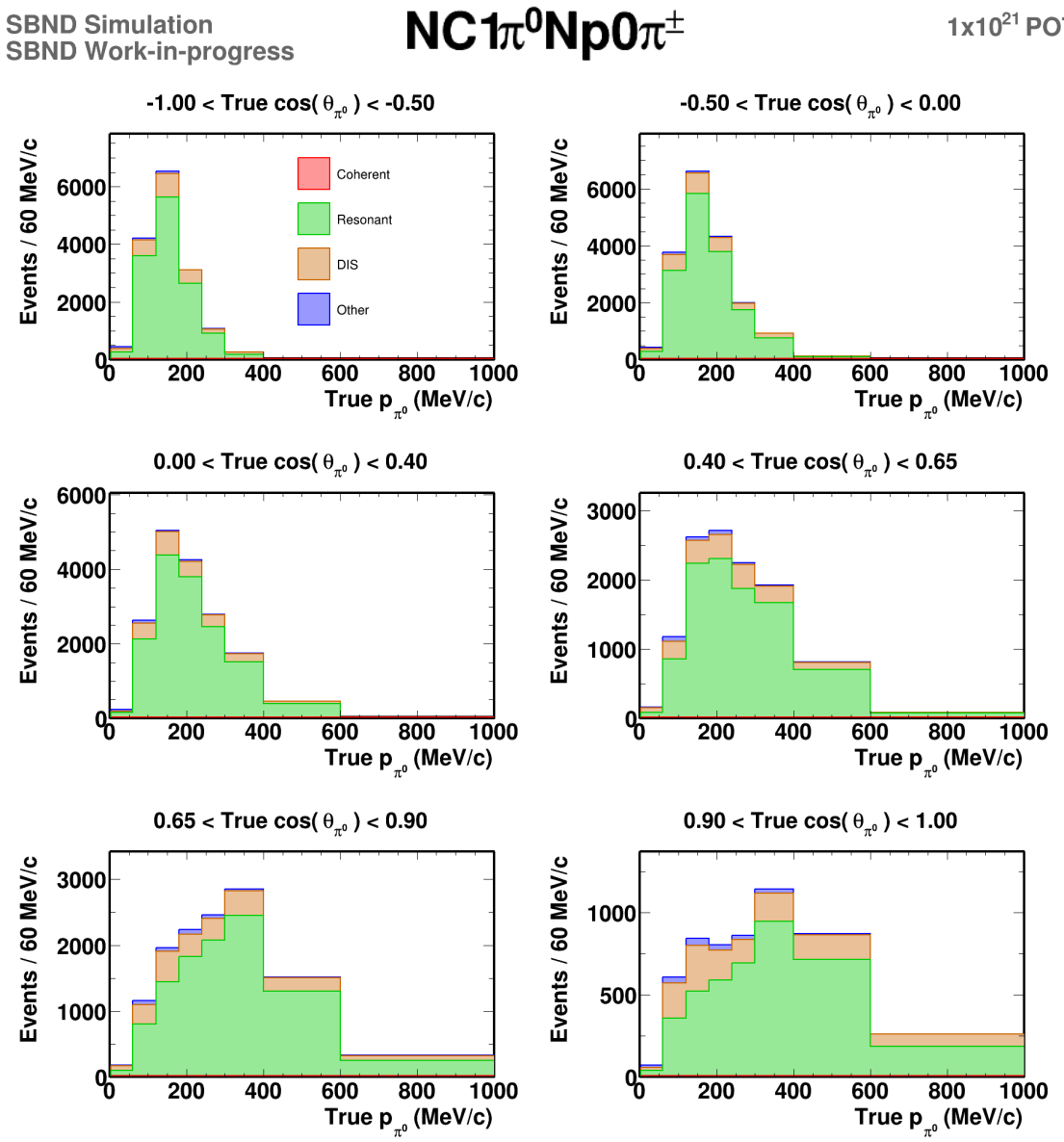


Figure E.2: Two-dimensional distribution of the neutral pion momentum for true signal events in the NC $1\pi^0$N$p0\pi^{\pm}$ signal definition.

Appendix F

Covariance Matrices

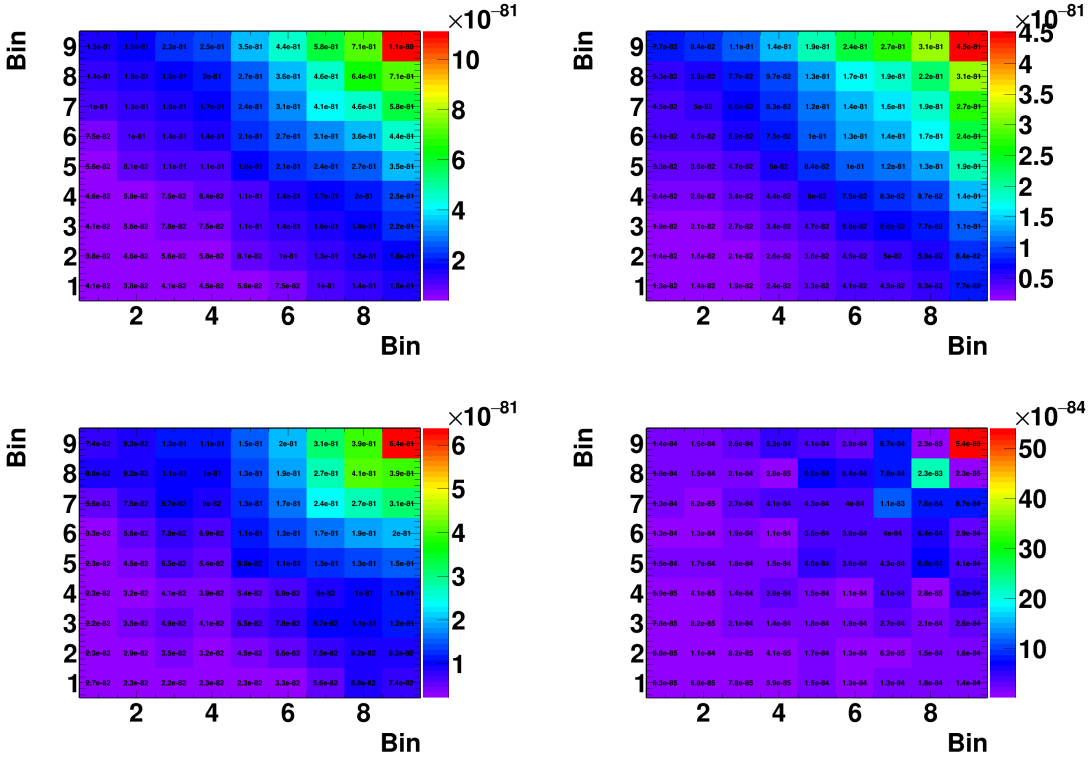


Figure F.1: Covariance matrices describing the correlated uncertainties in the binned $\cos\theta_{\pi^0}$ space for the NC $1\pi^0$ channel. From the top left, clockwise they show all systematics, combined flux systematics, combined Geant4 reinteraction systematics and GENIE interaction systematics respectively.

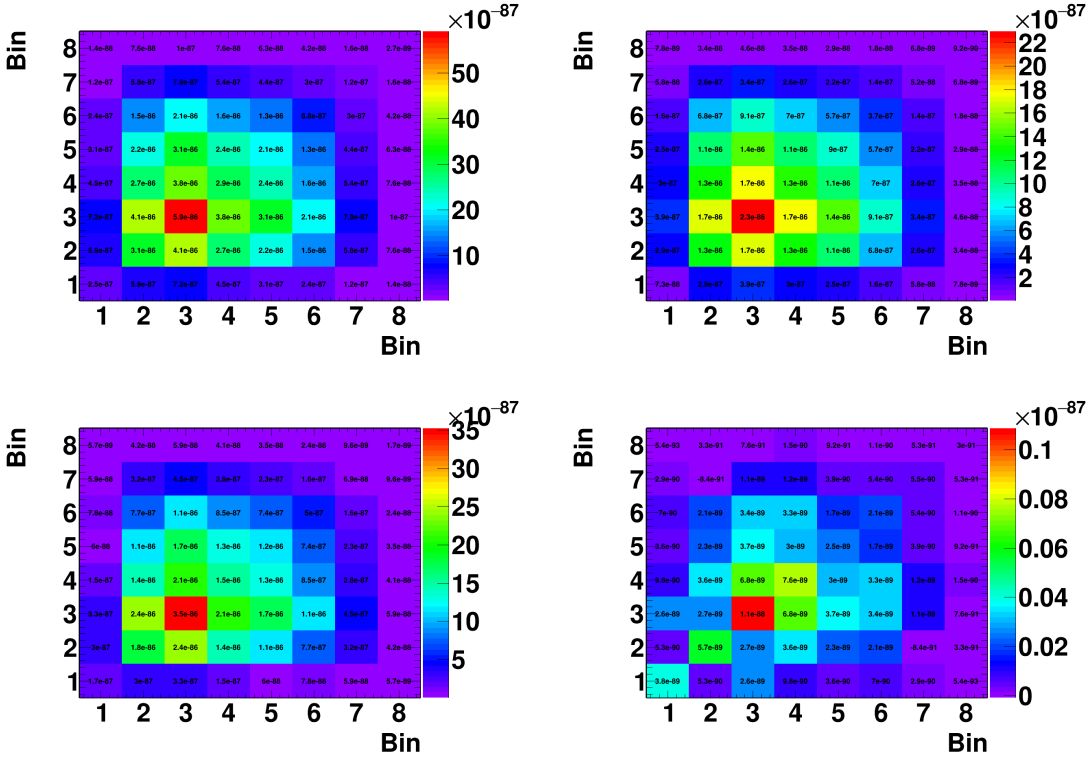


Figure F.2: Covariance matrices describing the correlated uncertainties in the binned p_{π^0} space for the NC $1\pi^0$ channel. From the top left, clockwise they show all systematics, combined flux systematics, combined Geant4 reinteraction systematics and GENIE interaction systematics respectively.

Appendix G

Shape and Normalisation Decomposition

Following [283], any given covariance matrix V_{ij} can be decomposed into a shape and normalisation components by calculating the element common to every bin:

$$V_{ij}^{\text{Norm}} = P_i P_j \frac{\sum_l \sum_k V_{kl}}{\left(\sum_k P_k\right)^2} \tag{G.1}$$

where P is the nominal predicted value. The shape component can then be calculated by normalising all universes to the same number of entries:

$$V_{ij}^{\text{Shape}} = M_{ij} - \frac{P_j}{\sum_k P_k} \sum_k M_{ik} - \frac{P_i}{\sum_k P_k} \sum_k M_{kj} + \frac{P_i P_j}{\left(\sum_k P_k\right)^2} \sum_k \sum_l M_{kl}. \tag{G.2}$$

References

[1] G. L. Kane. *Modern elementary particle physics*. 2nd Ed. Cambridge: Cambridge University Press, 2017. ISBN: 1-316-73273-8.

[2] M. Thomson. *Modern particle physics*. Cambridge: Cambridge University Press, 2013. ISBN: 1-139-52536-0.

[3] M. Veltman. *Facts and mysteries in elementary particle physics*. River Edge, NJ: World Scientific Pub., 2003. ISBN: 1-281-87256-3.

[4] D. J. Griffiths. *Introduction to elementary particles*. 2nd, Rev. Ed. Physics textbook. Weinheim: Wiley-VCH, 2008. ISBN: 9783527406012.

[5] X. Fan et al. "Measurement of the Electron Magnetic Moment". *Phys. Rev. Lett.* 130-7 (Feb. 2023), p. 071801. DOI: 10.1103/PhysRevLett.130.071801.

[6] Q. R. Ahmad et al. "Direct evidence for neutrino flavor transformation from neutral-current interactions in the Sudbury Neutrino Observatory". *Phys. Rev. Lett.* 89.1 (2002). ISSN: 0031-9007.

[7] Y. Fukuda et al. "Evidence for oscillation of atmospheric neutrinos". *Phys. Rev. Lett.* 81 (1998), pp. 1562–1567. DOI: 10.1103/PhysRevLett.81.1562. arXiv: hep-ex/9807003.

[8] W. Pauli. "Dear radioactive ladies and gentlemen". *Phys. Today* 31N9 (1978), p. 27.

[9] J. Chadwick. "The intensity distribution in the magnetic spectrum of beta particles from radium (B + C)". *Verh. Phys. Gesell.* 16 (1914), pp. 383–391.

[10] J. Chadwick. "The Existence of a Neutron". *Proc. Roy. Soc. Lond. A* 136.830 (1932), pp. 692–708. DOI: 10.1098/rspa.1932.0112.

[11] J. Formaggio. *Neutrinos in the Spotlight, MIT Physics Annual 2016*. https://physics.mit.edu/wp-content/uploads/2021/01/physicsatmit_16_formaggio.pdf. [Accessed: 8-1-2024].

[12] C. Cowan and F. Reines. *Pauli letter collection: telegram from Frederick Reines and Clyde Cowan*. http://cds.cern.ch/record/828311. [Accessed: 8-1-2024].

[13] C. L. Cowan et al. "Detection of the Free Neutrino: a Confirmation". *Science* 124.3212 (1956), pp. 103–104. ISSN: 0036-8075. DOI: 10.1126/science.124.3212.103.

[14] S. H. Neddermeyer and C. D. Anderson. "Note on the nature of cosmic-ray particles". *Phys. Rev.* 51.10 (1937), pp. 884–886. ISSN: 0031-899X.

[15] J. Oppenheimer and R. Serber. "Note on the nature of cosmic-ray particles [15]". *Phys. Rev.* 51.12 (1937), pp. 1113–1113. ISSN: 0031-899X.

[16] M. Conversi and O. Piccioni. "On the disintegration of slow mesons". *Phys. Rev.* 70.11-12 (1946), pp. 874–881. ISSN: 0031-899X.

[17] M. Conversi and O. Piccioni. "On the mean life of slow mesons". *Phys. Rev.* 70.11-12 (1946), pp. 859–873. ISSN: 0031-899X.

[18] M. Conversi, E. Pancini, and O. Piccioni. "On the Disintegration of Negative Mesons". *Phys. Rev.* 71 - 3 (Feb. 1947), pp. 209–210. DOI: 10.1103/PhysRev.71.209.

[19] B. Pontecorvo. "Electron and Muon Neutrinos". *Zh. Eksp. Teor. Fiz.* 37 (1959), pp. 1751–1757.

[20] R. Davis. "Bull. Am. Phys. Soc." (1959).

[21] R. Davis. "Attempt to Detect the Antineutrinos from a Nuclear Reactor by the $Cl^{37}(\overline{\nu}, e^-)A^{37}$ Reaction". *Phys. Rev.* 97 (3 Feb. 1955), pp. 766–769. DOI: 10.1103/PhysRev.97.766.

[22] G. Danby et al. "Observation of high-energy neutrino reactions and the existence of two kinds of neutrinos". *Phys. Rev. Lett.* 9.1 (1962), pp. 36–44. ISSN: 0031-9007.

[23] M. L. Perl et al. "Evidence for Anomalous Lepton Production in $e^+ - e^-$ Annihilation". *Phys. Rev. Lett.* 35 - 22 (Dec. 1975), pp. 1489–1492. DOI: 10.1103/PhysRevLett.35.1489.

[24] M. L. Perl. "Review of Heavy Lepton Production in e+ e- Annihilation". *SLAC-PUB-2022*. Oct. 1977, pp. 197–216.

[25] M. Cavalli Sforza and G. Goggi. "Inclusive Muon Production At Spear And The Hypothesis Of Heavy Leptons". *11th Rencontres de Moriond: International Meeting on Storage Ring Physics*. Orsay, 1976, p. 301.

[26] H. Harari. "Three Generations of Quarks and Leptons". *5th International Conference on Meson Spectroscopy*. July 1977, p. 0170.

[27] K. Kodama et al. "Observation of tau neutrino interactions". *Phys. Lett. B* 504 (2001), pp. 218–224. DOI: 10.1016/S0370-2693(01)00307-0. arXiv: hep-ex/0012035.

[28] K. Kodama et al. "Final tau-neutrino results from the DONuT experiment". *Phys. Rev. D* 78 (5 Sept. 2008), p. 052002. DOI: 10.1103/PhysRevD.78.052002.

[29] G. Arnison et al. "Experimental Observation of Lepton Pairs of Invariant Mass Around $95\,\mathrm{GeV/c^2}$ at the CERN SPS Collider". *Phys. Lett. B* 126 (1983), pp. 398–410. DOI: 10.1016/0370-2693(83)90188-0.

[30] P. Bagnaia et al. "Evidence for $Z^0 \to e^+e^-$ at the CERN $\bar{p}p$ Collider". *Phys. Lett. B* 129 (1983), pp. 130–140. DOI: 10.1016/0370-2693(83)90744-X.

[31] D. Décamp et al. "Determination of the number of light neutrino species". *Phys. Lett. B* 231 (1989), pp. 519–529. DOI: 10.1016/0370-2693(89)90704-1.

[32] R. L. Workman et al. "Review of Particle Physics". *PTEP* 2022 (2022), p. 083C01. DOI: 10.1093/ptep/ptac097.

[33] S. Schael et al. "Precision electroweak measurements on the Z resonance". *Phys. Rept.* 427 (2006), pp. 257–454. DOI: 10.1016/j.physrep.2005.12.006. arXiv: hep-ex/0509008.

[34] K. Kainulainen and K. A. Olive. "Astrophysical and cosmological constraints on neutrino masses". *Springer Tracts Mod. Phys.* 190 (2003), pp. 53–74. DOI: 10.1007/978-3-540-44901-0_3. arXiv: hep-ph/0206163.

[35] P. F. de Salas and S. Pastor. "Relic neutrino decoupling with flavour oscillations revisited". *JCAP* 07 (2016), p. 051. DOI: 10.1088/1475-7516/2016/07/051. arXiv: 1606.06986 [hep-ph].

[36] G. Aad et al. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". *Phys. Lett. B* 716 (2012), pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214 [hep-ex].

[37] S. Chatrchyan et al. "Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC". *Phys. Lett. B* 716 (2012), pp. 30–61. DOI: 10.1016/j.physletb.2012.08.021. arXiv: 1207.7235 [hep-ex].

[38] J. N. Bahcall. "Solar Neutrinos. I. Theoretical". *Phys. Rev. Lett.* 12 (11 Mar. 1964), pp. 300–302. DOI: 10.1103/PhysRevLett.12.300.

[39] J. N. Bahcall. "Solar Neutrinos". *Phys. Rev. Lett.* 17 (7 Aug. 1966), pp. 398–401. DOI: 10.1103/PhysRevLett.17.398.

[40] R. Davis, D. S. Harmer, and K. C. Hoffman. "Search for Neutrinos from the Sun". *Phys. Rev. Lett.* 20 (21 May 1968), pp. 1205–1209. DOI: 10.1103/PhysRevLett.20.1205.

[41] J. N. Bahcall, N. A. Bahcall, and G. Shaviv. "Present Status of the Theoretical Predictions for the ^{37}Cl Solar-Neutrino Experiment". *Phys. Rev. Lett.* 20 (21 May 1968), pp. 1209–1212. DOI: 10.1103/PhysRevLett.20.1209.

[42] J. N. Bahcall and R. Davis. "Solar Neutrinos: A Scientific Puzzle". *Science* 191.4224 (1976), pp. 264–267. ISSN: 00368075, 10959203. URL: http://www.jstor.org/stable/1741261 (visited on 09/25/2023).

[43] J. Abdurashitov et al. "Results from SAGE". *Nuclear Physics B - Proceedings Supplements* 48.1 (1996). Proceedings of the Fourth International Workshop on Theoretical and Phenomenological Aspects of Underground Physics, pp. 299–303. ISSN: 0920-5632. DOI: https://doi.org/10.1016/0920-5632(96)00264-2.

[44] T. Kirsten. "GALLEX solar neutrino results". *Progress in Particle and Nuclear Physics* 40 (1998). Neutrinos in Astro, Particle and Nuclear Physics, pp. 85–99. ISSN: 0146-6410. DOI: https://doi.org/10.1016/S0146-6410(98)00013-1.

[45] Y. Suzuki. "Kamiokande solar neutrino results". *Nuclear Physics B - Proceedings Supplements* 38.1 (1995). Neutrino 94, pp. 54–59. ISSN: 0920-5632. DOI: https://doi.org/10.1016/0920-5632(94)00733-C.

[46] M. Velho. "ATLAS Colouring Books: a guide for Parents and Teachers" (2022). URL: https://cds.cern.ch/record/2801358.

[47] T. KAJITA. "Atmospheric neutrinos and discovery of neutrino oscillations". *Proceedings of the Japan Academy, Series B* 86.4 (2010), pp. 303–321. DOI: 10.2183/pjab.86.303.

[48] T. J. Haines et al. "Calculation of Atmospheric Neutrino-Induced Backgrounds in a Nucleon-Decay Search". *Phys. Rev. Lett.* 57 (16 Oct. 1986), pp. 1986–1989. DOI: 10.1103/PhysRevLett.57.1986.

[49] M. Nakahata et al. "Atmospheric Neutrino Background and Pion Nuclear Effect for KAMIOKA Nucleon Decay Experiment". *Journal of the Physical Society of Japan* 55.11 (1986), pp. 3786–3805. DOI: 10.1143/JPSJ.55.3786. eprint: https://doi.org/10.1143/JPSJ.55.3786.

[50] D. Casper et al. "Measurement of atmospheric neutrino composition with the IMB-3 detector". *Phys. Rev. Lett.* 66 (20 May 1991), pp. 2561–2564. DOI: 10.1103/PhysRevLett.66.2561.

[51] K. Hirata et al. "Experimental study of the atmospheric neutrino flux". *Physics Letters B* 205.2 (1988), pp. 416–420. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(88)91690-5.

[52] M. Ambrosio et al. "Measurement of the atmospheric neutrino induced upgoing muon flux using MACRO". *Phys. Lett. B* 434 (1998), pp. 451–457. DOI: 10.1016/S0370-2693(98)00885-5. arXiv: hep-ex/9807005.

[53] W. W. M. Allison et al. "Measurement of the atmospheric neutrino flavor composition in Soudan-2". *Phys. Lett. B* 391 (1997), pp. 491–500. DOI: 10.1016/S0370-2693(96)01609-7. arXiv: hep-ex/9611007.

[54] H. H. Chen. "Direct Approach to Resolve the Solar-Neutrino Problem". *Phys. Rev. Lett.* 55 (14 Sept. 1985), pp. 1534–1536. DOI: 10.1103/PhysRevLett.55.1534.

[55] K. Eguchi et al. "First results from KamLAND: Evidence for reactor anti-neutrino disappearance". *Phys. Rev. Lett.* 90 (2003), p. 021802. DOI: 10.1103/PhysRevLett.90.021802. arXiv: hep-ex/0212021.

[56] M. H. Ahn et al. "Indications of neutrino oscillation in a 250 km long baseline experiment". *Phys. Rev. Lett.* 90 (2003), p. 041801. DOI: 10.1103/PhysRevLett.90.041801. arXiv: hep-ex/0212007.

[57] M. H. Ahn et al. "Measurement of Neutrino Oscillation by the K2K Experiment". *Phys. Rev. D* 74 (2006), p. 072003. DOI: 10.1103/PhysRevD.74.072003. arXiv: hep-ex/0606032.

[58] B. Pontecorvo. "Inverse beta processes and nonconservation of lepton charge". *Zh. Eksp. Teor. Fiz.* 34 (1957), p. 247.

[59] B. Pontecorvo. "Mesonium and anti-mesonium". *Sov. Phys. JETP* 6 (1957), p. 429.

[60] B. Pontecorvo. "Neutrino Experiments and the Problem of Conservation of Leptonic Charge". *Zh. Eksp. Teor. Fiz.* 53 (1967), pp. 1717–1725.

[61] C. Giunti and C. W. Kim. *Fundamentals of Neutrino Physics and Astrophysics*. 2007. ISBN: 978-0-19-850871-7.

[62] L. Wolfenstein. "Neutrino oscillations in matter". *Phys. Rev. D* 17 (9 May 1978), pp. 2369–2374. DOI: 10.1103/PhysRevD.17.2369.

[63] S. P. Mikheev and A. Y. Smirnov. "Resonant amplification of neutrino oscillations in matter and solar neutrino spectroscopy". *Nuovo Cim. C* 9 (1986), pp. 17–26. DOI: 10.1007/BF02508049.

[64] L. D. Kolupaeva et al. "Neutrino oscillations: status and prospects for the determination of neutrino mass ordering and the leptonic CP-violation phase". *Uspekhi Fizicheskikh Nauk* 193.8 (2023), pp. 801–824.

[65] B. Aharmim et al. "Combined Analysis of all Three Phases of Solar Neutrino Data from the Sudbury Neutrino Observatory". *Phys. Rev. C* 88 (2013), p. 025501. DOI: 10.1103/PhysRevC.88.025501. arXiv: 1109.0763 [nucl-ex].

[66] K. Abe et al. "Solar neutrino measurements in Super-Kamiokande-IV". *Phys. Rev. D* 94 (5 Sept. 2016), p. 052010. DOI: 10.1103/PhysRevD.94.052010.

[67] A. Gando et al. "Reactor on-off antineutrino measurement with KamLAND". *Phys. Rev. D* 88 (3 Aug. 2013), p. 033001. DOI: 10.1103/PhysRevD.88.033001.

[68] K. Abe et al. "Indication of Electron Neutrino Appearance from an Accelerator-Produced Off-Axis Muon Neutrino Beam". *Phys. Rev. Lett.* 107 (4 July 2011), p. 041801. DOI: 10.1103/PhysRevLett.107.041801.

[69] P. Adamson et al. "Improved Search for Muon-Neutrino to Electron-Neutrino Oscillations in MINOS". *Phys. Rev. Lett.* 107 (18 Oct. 2011), p. 181802. DOI: 10.1103/PhysRevLett.107.181802.

[70] F. P. An et al. "Precision Measurement of Reactor Antineutrino Oscillation at Kilometer-Scale Baselines by Daya Bay". *Phys. Rev. Lett.* 130.16 (2023), p. 161802. DOI: 10.1103/PhysRevLett.130.161802. arXiv: 2211.14988 [hep-ex].

[71] C. D. Shin et al. "Observation of reactor antineutrino disappearance using delayed neutron capture on hydrogen at RENO". *JHEP* 04 (2020), p. 029. DOI: 10.1007/JHEP04(2020)029. arXiv: 1911.04601 [hep-ex].

[72] K. Abe et al. "Atmospheric neutrino oscillation analysis with external constraints in Super-Kamiokande I-IV". *Phys. Rev. D* 97.7 (2018), p. 072001. DOI: 10.1103/PhysRevD.97.072001. arXiv: 1710.09126 [hep-ex].

[73] K. Abe et al. "Updated T2K measurements of muon neutrino and antineutrino disappearance using 3.6×1021 protons on target". *Phys. Rev. D* 108.7 (2023), p. 072011. DOI: 10.1103/PhysRevD.108.072011. arXiv: 2305.09916 [hep-ex].

[74] M. A. Acero et al. "Improved measurement of neutrino oscillation parameters by the NOvA experiment". *Phys. Rev. D* 106.3 (2022), p. 032004. DOI: 10.1103/PhysRevD.106.032004. arXiv: 2108.08219 [hep-ex].

[75] S. K. Agarwalla et al. "A close look on 2-3 mixing angle with DUNE in light of current neutrino oscillation data". *JHEP* 03 (2022), p. 206. DOI: 10.1007/JHEP03(2022)206. arXiv: 2111.11748 [hep-ph].

[76] I. Esteban et al. "The fate of hints: updated global analysis of three-flavor neutrino oscillations". *JHEP* 09 (2020), p. 178. DOI: 10.1007/JHEP09(2020)178. arXiv: 2007.14792 [hep-ph].

[77] P. F. de Salas et al. "2020 global reassessment of the neutrino oscillation picture". *JHEP* 02 (2021), p. 071. DOI: 10.1007/JHEP02(2021)071. arXiv: 2006.11237 [hep-ph].

[78] F. Capozzi et al. "Unfinished fabric of the three neutrino paradigm". *Phys. Rev. D* 104.8 (2021), p. 083031. DOI: 10.1103/PhysRevD.104.083031. arXiv: 2107.00532 [hep-ph].

[79] STEREO. *About the Stereo Experiment*. https://www.stereo-experiment.org/scientific.php. [Accessed: 7-5-2024].

[80] C. S. Wu et al. "Experimental Test of Parity Conservation in Beta Decay". *Phys. Rev.* 105 (4 Feb. 1957), pp. 1413–1415. DOI: 10.1103/PhysRev.105.1413.

[81] J. H. Christenson et al. "Evidence for the 2π Decay of the K_2^0 Meson". *Phys. Rev. Lett.* 13 (4 July 1964), pp. 138–140. DOI: 10.1103/PhysRevLett.13.138.

[82] A. D. Sakharov. "Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe". *Pisma Zh. Eksp. Teor. Fiz.* 5 (1967), pp. 32–35. DOI: 10.1070/PU1991v034n05ABEH002497.

[83] K. Abe et al. "Constraint on the matter–antimatter symmetry-violating phase in neutrino oscillations". *Nature* 580.7803 (2020). [Erratum: Nature 583, E16 (2020)], pp. 339–344. DOI: 10.1038/s41586-020-2177-0. arXiv: 1910.03887 [hep-ex].

[84] K. Abe et al. "Measurements of neutrino oscillation parameters from the T2K experiment using 3.6×10^{21} protons on target". *Eur. Phys. J. C* 83.9 (2023), p. 782. DOI: 10.1140/epjc/s10052-023-11819-x. arXiv: 2303.03222 [hep-ex].

[85] Z.-z. Xing and Z.-h. Zhao. "A review of μ-τ flavor symmetry in neutrino physics". *Rept. Prog. Phys.* 79.7 (2016), p. 076201. DOI: 10.1088/0034-4885/79/7/076201. arXiv: 1512.04207 [hep-ph].

[86] S. F. King and C. C. Nishi. "Mu-tau symmetry and the Littlest Seesaw". *Physics Letters B* 785 (2018), pp. 391–398. ISSN: 0370-2693. DOI: https://doi.org/10.1016/j.physletb.2018.08.056. URL: https://www.sciencedirect.com/science/article/pii/S0370269318306713.

[87] G. Mention et al. "Reactor antineutrino anomaly". *Phys. Rev. D* 83 (7 Apr. 2011), p. 073006. DOI: 10.1103/PhysRevD.83.073006.

[88] C. Giunti et al. "Gallium Anomaly: critical view from the global picture of ν_e and $\overline{\nu}_e$ disappearance". *JHEP* 10 (2022), p. 164. DOI: 10.1007/JHEP10(2022)164. arXiv: 2209.00916 [hep-ph].

[89] A. Aguilar et al. "Evidence for neutrino oscillations from the observation of $\bar{\nu}_e$ appearance in a $\bar{\nu}_\mu$ beam". *Phys. Rev. D* 64 (2001), p. 112007. DOI: 10.1103/PhysRevD.64.112007. arXiv: hep-ex/0104049.

[90] A. A. Aguilar-Arevalo et al. "Updated MiniBooNE neutrino oscillation results with increased data and new background studies". *Phys. Rev. D* 103 (5 Mar. 2021), p. 052002. DOI: 10.1103/PhysRevD.103.052002.

[91] P. A. Machado, O. Palamara, and D. W. Schmitz. "The Short-Baseline Neutrino Program at Fermilab". *Ann. Rev. Nucl. Part. Sci.* 69 (2019), pp. 363–387. DOI: 10.1146/annurev-nucl-101917-020949. arXiv: 1903.04608 [hep-ex].

[92] C. Giunti and M. Laveder. "3+1 and 3+2 Sterile Neutrino Fits". *Phys. Rev. D* 84 (2011), p. 073008. DOI: 10.1103/PhysRevD.84.073008. arXiv: 1107.1452 [hep-ph].

[93] C. Giunti et al. "Reactor antineutrino anomaly in light of recent flux model refinements". *Physics Letters B* 829 (2022), p. 137054. ISSN: 0370-2693. DOI: https://doi.org/10.1016/j.physletb.2022.137054.

[94] J. M. Berryman, V. Brdar, and P. Huber. "Particle physics origin of the 5 MeV bump in the reactor antineutrino spectrum?" *Phys. Rev. D* 99.5 (2019), p. 055045. DOI: 10.1103/PhysRevD.99.055045. arXiv: 1803.08506 [hep-ph].

[95] M. Danilov and N. Skrobova. "New results from the DANSS experiment". *PoS* EPS-HEP2021 (2022), p. 241. DOI: 10.22323/1.398.0241.

[96] Z. Atif et al. "Search for sterile neutrino oscillations using RENO and NEOS data". *Phys. Rev. D* 105.11 (2022), p. L111101. DOI: 10.1103/PhysRevD.105.L111101. arXiv: 2011.00896 [hep-ex].

[97] M. Andriamirado et al. "Improved short-baseline neutrino oscillation search and energy spectrum measurement with the PROSPECT experiment at HFIR". *Phys. Rev. D* 103 (3 Feb. 2021), p. 032001. DOI: 10.1103/PhysRevD.103.032001.

[98] H. Almazán et al. "Improved sterile neutrino constraints from the STEREO experiment with 179 days of reactor-on data". *Phys. Rev. D* 102 (5 Sept. 2020), p. 052002. DOI: 10.1103/PhysRevD.102.052002.

[99] A. P. Serebrov et al. "Search for sterile neutrinos with the Neutrino-4 experiment and measurement results". *Phys. Rev. D* 104 (3 Aug. 2021), p. 032003. DOI: 10.1103/PhysRevD.104.032003.

[100] M. Danilov. "Review of sterile neutrino searches at very short-baseline reactor experiments". *Phys. Scripta* 97.9 (2022), p. 094001. DOI: 10.1088/1402-4896/ac81fd. arXiv: 2203.03042 [hep-ex].

[101] V. V. Barinov et al. "Results from the Baksan Experiment on Sterile Transitions (BEST)". *Phys. Rev. Lett.* 128.23 (2022), p. 232501. DOI: 10.1103/PhysRevLett.128.232501. arXiv: 2109.11482 [nucl-ex].

[102] V. V. Barinov et al. "Search for electron-neutrino transitions to sterile states in the BEST experiment". *Phys. Rev. C* 105.6 (2022), p. 065502. DOI: 10.1103/PhysRevC.105.065502. arXiv: 2201.07364 [nucl-ex].

[103] P. Abratenko et al. "Search for an Excess of Electron Neutrino Interactions in MicroBooNE Using Multiple Final-State Topologies". *Phys. Rev. Lett.* 128.24 (2022), p. 241801. DOI: 10.1103/PhysRevLett.128.241801. arXiv: 2110.14054 [hep-ex].

[104] P. Abratenko 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". *Phys. Rev. Lett.* 128 (2022), p. 111801. DOI: 10.1103/PhysRevLett.128.111801. arXiv: 2110.00409 [hep-ex].

[105] P. Abratenko et al. "First Constraints on Light Sterile Neutrino Oscillations from Combined Appearance and Disappearance Searches with the MicroBooNE Detector". *Phys. Rev. Lett.* 130.1 (2023), p. 011801. DOI: 10.1103/PhysRevLett.130.011801. arXiv: 2210.10216 [hep-ex].

[106] P. Adamson et al. "Search for Sterile Neutrinos Mixing with Muon Neutrinos in MINOS". *Phys. Rev. Lett.* 117.15 (2016), p. 151803. DOI: 10.1103/PhysRevLett.117.151803. arXiv: 1607.01176 [hep-ex].

[107] P. Adamson et al. "Search for active-sterile neutrino mixing using neutral-current interactions in NOvA". *Phys. Rev. D* 96.7 (2017), p. 072006. DOI: 10.1103/PhysRevD.96.072006. arXiv: 1706.04592 [hep-ex].

[108] M. A. Acero et al. "Search for Active-Sterile Antineutrino Mixing Using Neutral-Current Interactions with the NOvA Experiment". *Phys. Rev. Lett.* 127.20 (2021), p. 201801. DOI: 10.1103/PhysRevLett.127.201801. arXiv: 2106.04673 [hep-ex].

[109] M. G. Aartsen et al. "Searching for eV-scale sterile neutrinos with eight years of atmospheric neutrinos at the IceCube Neutrino Telescope". *Phys. Rev. D* 102.5 (2020), p. 052009. DOI: 10.1103/PhysRevD.102.052009. arXiv: 2005.12943 [hep-ex].

[110] M. Dentler et al. "Updated Global Analysis of Neutrino Oscillations in the Presence of eV-Scale Sterile Neutrinos". *JHEP* 08 (2018), p. 010. DOI: 10.1007/JHEP08(2018)010. arXiv: 1803.10661 [hep-ph].

[111] A. Diaz et al. "Where Are We With Light Sterile Neutrinos?" *Phys. Rept.* 884 (2020), pp. 1–59. DOI: 10.1016/j.physrep.2020.08.005. arXiv: 1906.00045 [hep-ex].

[112] B. Batell, J. Berger, and A. Ismail. "Probing the Higgs portal at the Fermilab short-baseline neutrino experiments". *Phys. Rev. D* 100 (11 Nov. 2019), p. 115039. DOI: 10.1103/PhysRevD.100.115039.

[113] *Neutrino Non-Standard Interactions: A Status Report*. Vol. 2. 2019, p. 001. DOI: 10.21468/SciPostPhysProc.2.001. arXiv: 1907.00991 [hep-ph].

[114] P. Ballett, S. Pascoli, and M. Ross-Lonergan. "MeV-scale sterile neutrino decays at the Fermilab Short-Baseline Neutrino program". *JHEP* 04 (2017), p. 102. DOI: 10.1007/JHEP04(2017)102. arXiv: 1610.08512 [hep-ph].

[115] 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 (2010), pp. 87–104. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2009.12.009.

[116] T. Golan, J. Sobczyk, and J. Żmuda. "NuWro: the Wrocław Monte Carlo Generator of Neutrino Interactions". *Nuclear Physics B - Proceedings Supplements* 229-232 (2012). Neutrino 2010, p. 499. ISSN: 0920-5632. DOI: https://doi.org/10.1016/j.nuclphysbps.2012.09.136.

[117] Y. Hayato and L. Pickering. "The NEUT neutrino interaction simulation program library". *Eur. Phys. J. ST* 230.24 (2021), pp. 4469–4481. DOI: 10.1140/epjs/s11734-021-00287-7. arXiv: 2106.15809 [hep-ph].

[118] O. Buss et al. "Transport-theoretical Description of Nuclear Reactions". *Phys. Rept.* 512 (2012), pp. 1–124. DOI: 10.1016/j.physrep.2011.12.001. arXiv: 1106.1344 [hep-ph].

[119] L. d. Broglie. "XXXV. A tentative theory of light quanta". *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 47.278 (1924), pp. 446–458.

[120] J. A. Formaggio and G. P. Zeller. "From eV to EeV: Neutrino cross sections across energy scales". *Rev. Mod. Phys.* 84 (3 Sept. 2012), pp. 1307–1341. DOI: 10.1103/RevModPhys.84.1307.

[121] A. Schukraft. *Private Communication.*

[122] C. H. Llewellyn Smith. "Neutrino Reactions at Accelerator Energies". *Phys. Rept.* 3 (1972), pp. 261–379. DOI: 10.1016/0370-1573(72)90010-5.

[123] G. Eichmann and C. S. Fischer. "Nucleon axial and pseudoscalar form factors from the covariant Faddeev equation". *Eur. Phys. J. A* 48 (2012), p. 9. DOI: 10.1140/epja/i2012-12009-6. arXiv: 1111.2614 [hep-ph].

[124] L. Alvarez-Ruso. "Neutrino interactions: challenges in the current theoretical picture". *Nucl. Phys. B Proc. Suppl.* 229-232 (2012). Ed. by G. S. Tzanakos, pp. 167–173. DOI: 10.1016/j.nuclphysbps.2012.09.027. arXiv: 1012.3871 [nucl-th].

[125] V. Bernard, L. Elouadrhiri, and U.-G. Meissner. "Axial structure of the nucleon: Topical Review". *J. Phys. G* 28 (2002), R1–R35. DOI: 10.1088/0954-3899/28/1/201. arXiv: hep-ph/0107088.

[126] T. De Forest. "Off-shell electron-nucleon cross sections: The impulse approximation". *Nuclear Physics A* 392.2 (1983), pp. 232–248. ISSN: 0375-9474. DOI: https://doi.org/10.1016/0375-9474(83)90124-0.

[127] R. A. Smith and E. J. Moniz. "NEUTRINO REACTIONS ON NUCLEAR TARGETS". *Nucl. Phys. B* 43 (1972). [Erratum: Nucl.Phys.B 101, 547 (1975)], p. 605. DOI: 10.1016/0550-3213(75)90612-4.

[128] T. Leitner et al. "Electron- and neutrino-nucleus scattering from the quasielastic to the resonance region". *Phys. Rev. C* 79 (3 Mar. 2009), p. 034601. DOI: 10.1103/PhysRevC.79.034601.

[129] S. Boyd et al. "Comparison of Models of Neutrino-Nucleus Interactions". *AIP Conference Proceedings*. Vol. 1189. 1. American Institute of Physics. 2009, pp. 60–73.

[130] L. B. Auerbach et al. "Measurements of charged current reactions of ν_μ on ^{12}C". *Phys. Rev. C* 66 (1 July 2002), p. 015501. DOI: 10.1103/PhysRevC.66.015501.

[131] V. Lyubushkin et al. "A Study of quasi-elastic muon neutrino and antineutrino scattering in the NOMAD experiment". *Eur. Phys. J. C* 63 (2009), pp. 355–381. DOI: 10.1140/epjc/s10052-009-1113-0. arXiv: 0812.4543 [hep-ex].

[132] L. Alvarez-Ruso. "2p2h or not 2p2h?" *AIP Conference Proceedings*. Vol. 1405. 1. American Institute of Physics. 2011, pp. 71–76.

[133] A. A. Aguilar-Arevalo et al. "First measurement of the muon neutrino charged current quasielastic double differential cross section". *Phys. Rev. D* 81 (9 May 2010), p. 092005. DOI: 10.1103/PhysRevD.81.092005.

[134] M. Martini et al. "A Unified approach for nucleon knock-out, coherent and incoherent pion production in neutrino interactions with nuclei". *Phys. Rev. C* 80 (2009), p. 065501. DOI: 10.1103/PhysRevC.80.065501. arXiv: 0910.2622 [nucl-th].

[135] M. Martini, M. Ericson, and G. Chanfray. "Neutrino quasielastic interaction and nuclear dynamics". *Phys. Rev. C* 84 (5 Nov. 2011), p. 055502. DOI: 10.1103/PhysRevC.84.055502.

[136] D. Rein and L. M. Sehgal. "Neutrino-excitation of baryon resonances and single pion production". *Annals of Physics* 133.1 (1981), pp. 79–153. ISSN: 0003-4916. DOI: https://doi.org/10.1016/0003-4916(81)90242-6.

[137] C. Berger and L. M. Sehgal. "Lepton mass effects in single pion production by neutrinos". *Phys. Rev. D* 76 (11 Nov. 2007), p. 113004. DOI: 10.1103/PhysRevD.76.113004.

[138] D. Rein and L. M. Sehgal. "Coherent π^0 production in neutrino reactions". *Nuclear Physics B* 223.1 (1983), pp. 29–44. ISSN: 0550-3213. DOI: https://doi.org/10.1016/0550-3213(83)90090-1.

[139] S. L. Adler. "Tests of the Conserved Vector Current and Partially Conserved Axial-Vector Current Hypotheses in High-Energy Neutrino Reactions". *Phys. Rev.* 135 (4B Aug. 1964), B963–B966. DOI: 10.1103/PhysRev.135.B963.

[140] L. Alvarez-Ruso. "Review of weak coherent pion production". *AIP Conference Proceedings*. Vol. 1405. 1. American Institute of Physics. 2011, pp. 140–145.

[141] L. Alvarez-Ruso et al. "NuSTEC White Paper: Status and challenges of neutrino–nucleus scattering". *Prog. Part. Nucl. Phys.* 100 (2018), pp. 1–68. DOI: 10.1016/j.ppnp.2018.01.006. arXiv: 1706.03621 [hep-ph].

[142] Y. Kurimoto et al. "Improved measurement of neutral current coherent π^0 production on carbon in a few-GeV neutrino beam". *Phys. Rev. D* 81 (11 June 2010), p. 111102. DOI: 10.1103/PhysRevD.81.111102.

[143] K. Abe et al. "Measurements of the ν_μ and $\overline{\nu}_\mu$-induced coherent charged pion production cross sections on ^{12}C by the T2K experiment". *Phys. Rev. D* 108 (9 Nov. 2023), p. 092009. DOI: 10.1103/PhysRevD.108.092009.

[144] J. M. Conrad, M. H. Shaevitz, and T. Bolton. "Precision measurements with high-energy neutrino beams". *Rev. Mod. Phys.* 70 (4 Oct. 1998), pp. 1341–1392. DOI: 10.1103/RevModPhys.70.1341.

[145] A. Bodek and U. K. Yang. "Higher twist, xi(omega) scaling, and effective LO PDFs for lepton scattering in the few GeV region". *J. Phys. G* 29 (2003). Ed. by K. Long and R. Edgecock, pp. 1899–1906. DOI: 10.1088/0954-3899/29/8/369. arXiv: hep-ex/0210024.

[146] T. Yang et al. "A Hadronization Model for Few-GeV Neutrino Interactions". *Eur. Phys. J. C* 63 (2009), pp. 1–10. DOI: 10.1140/epjc/s10052-009-1094-z. arXiv: 0904.4043 [hep-ph].

[147] S. Dytman et al. "Comparison of validation methods of simulations for final state interactions in hadron production experiments". *Phys. Rev. D* 104.5 (2021), p. 053006. DOI: 10.1103/PhysRevD.104.053006. arXiv: 2103.07535 [hep-ph].

[148] R. S. Jones. "Muon-neutrino disappearance with multiple liquid argon time projection chambers in the Fermilab Booster neutrino beam." PhD thesis. University of Liverpool, 2021. DOI: 10.17638/03143192.

[149] X. G. Lu et al. "Measurement of final-state correlations in neutrino muon-proton mesonless production on hydrocarbon at $<E_\nu> = 3$ GeV". *Phys. Rev. Lett.* 121.2 (2018), p. 022504. DOI: 10.1103/PhysRevLett.121.022504. arXiv: 1805.05486 [hep-ex].

[150] T. Cai et al. "Nucleon binding energy and transverse momentum imbalance in neutrino-nucleus reactions". *Phys. Rev. D* 101.9 (2020), p. 092001. DOI: 10.1103/PhysRevD.101.092001. arXiv: 1910.08658 [hep-ex].

[151] D. Coplowe et al. "Probing nuclear effects with neutrino-induced charged-current neutral pion production". *Phys. Rev. D* 102.7 (2020). [Erratum: Phys.Rev.D 110, 059903 (2024)], p. 072007. DOI: 10.1103/PhysRevD.102.072007. arXiv: 2002.05812 [hep-ex].

[152] K. Abe et al. "Characterization of nuclear effects in muon-neutrino scattering on hydrocarbon with a measurement of final-state kinematics and correlations in charged-current pionless interactions at T2K". *Phys. Rev. D* 98.3 (2018), p. 032003. DOI: 10.1103/PhysRevD.98.032003. arXiv: 1802.05078 [hep-ex].

[153] K. Abe et al. "First T2K measurement of transverse kinematic imbalance in the muon-neutrino charged-current single-π^+ production channel containing at least one proton". *Phys. Rev. D* 103 (11 June 2021), p. 112009. DOI: 10.1103/PhysRevD.103.112009.

[154] P. Abratenko et al. "First Double-Differential Measurement of Kinematic Imbalance in Neutrino Interactions with the MicroBooNE Detector". *Phys. Rev. Lett.* 131 (10 Sept. 2023), p. 101802. DOI: 10.1103/PhysRevLett.131.101802.

[155] X. .-. Lu et al. "Reconstruction of Energy Spectra of Neutrino Beams Independent of Nuclear Effects". *Phys. Rev. D* 92.5 (2015), p. 051302. DOI: 10.1103/PhysRevD.92.051302. arXiv: 1507.00967 [hep-ex].

[156] X. .-. Lu et al. "Measurement of nuclear effects in neutrino interactions with minimal dependence on neutrino energy". *Phys. Rev. C* 94.1 (2016), p. 015503. DOI: 10.1103/PhysRevC.94.015503. arXiv: 1512.05748 [nucl-th].

[157] X. Lu and J. T. Sobczyk. "Identification of nuclear effects in neutrino and antineutrino interactions on nuclei using generalized final-state correlations". *Phys. Rev. C* 99.5 (2019), p. 055504. DOI: 10.1103/PhysRevC.99.055504. arXiv: 1901.06411 [hep-ph].

[158] P. Abratenko et al. "Measurement of nuclear effects in neutrino-argon interactions using generalized kinematic imbalance variables with the MicroBooNE detector" (Oct. 2023). arXiv: 2310.06082 [nucl-ex].

[159] W. Lee et al. "Single-Pion Production in Neutrino and Antineutrino Reactions". *Phys. Rev. Lett.* 38 (5 Jan. 1977), pp. 202–204. DOI: 10.1103/PhysRevLett.38.202.

[160] W. Krenz et al. "Experimental Study of Exclusive One Pion Production in All Neutrino Induced Neutral Current Channels". *Nucl. Phys. B* 135 (1978), pp. 45–65. DOI: 10.1016/0550-3213(78)90213-4.

[161] E. Isiksal, D. Rein, and J. G. Morfin. "Evidence for Neutrino- and Antineutrino-Induced Coherent π^0 Production". *Phys. Rev. Lett.* 52 (13 Mar. 1984), pp. 1096–1099. DOI: 10.1103/PhysRevLett.52.1096.

[162] M. Derrick et al. "Study of single-pion production by weak neutral currents in low-energy νd interactions". *Phys. Rev. D* 23 (3 Feb. 1981), pp. 569–575. DOI: 10.1103/PhysRevD.23.569.

[163] H. Faissner et al. "Observation of Neutrino and Anti-neutrino Induced Coherent Neutral Pion Production Off ^{27}Al". *Phys. Lett. B* 125 (1983), pp. 230–236. DOI: 10.1016/0370-2693(83)91274-1.

[164] F. Bergsma et al. "Measurement of the cross section of coherent π^0 production by muon-neutrino and antineutrino neutral-current interactions on nuclei". *Physics Letters B* 157.5 (1985), pp. 469–474. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(85)90402-2.

[165] H. J. Grabosch et al. "COHERENT PION PRODUCTION IN NEUTRINO AND ANTI-NEUTRINO INTERACTIONS ON NUCLEI OF HEAVY FREON MOLECULES". *Z. Phys. C* 31 (1986), p. 203. DOI: 10.1007/BF01479528.

[166] C. Baltay et al. "Evidence for Coherent Neutral-Pion Production by High-Energy Neutrinos". *Phys. Rev. Lett.* 57 (21 Nov. 1986), pp. 2629–2632. DOI: 10.1103/PhysRevLett.57.2629.

[167] S. J. Barish et al. "Observation of Single-Pion Production by a Weak Neutral Current". *Phys. Rev. Lett.* 33 (7 Aug. 1974), pp. 448–451. DOI: 10.1103/PhysRevLett.33.448.

[168] S. Nakayama et al. "Measurement of single pi0 production in neutral current neutrino interactions with water by a 1.3-GeV wide band muon neutrino beam". *Phys. Lett. B* 619 (2005), pp. 255–262. DOI: 10.1016/j.physletb.2005.05.044. arXiv: hep-ex/0408134.

[169] A. A. Aguilar-Arevalo et al. "First Observation of Coherent π^0 Production in Neutrino Nucleus Interactions with $E_\nu < 2$ GeV". *Phys. Lett. B* 664 (2008), pp. 41–46. DOI: 10.1016/j.physletb.2008.05.006. arXiv: 0803.3423 [hep-ex].

[170] A. A. Aguilar-Arevalo et al. "Measurement of ν_μ and $\bar{\nu}_\mu$ induced neutral current single π^0 production cross sections on mineral oil at $E_\nu \sim \mathcal{O}(1\text{GeV})$". *Phys. Rev. D* 81 (2010), p. 013005. DOI: 10.1103/PhysRevD.81.013005. arXiv: 0911.2063 [hep-ex].

[171] C. T. Kullenberg et al. "A Measurement of Coherent Neutral Pion Production in Neutrino Neutral Current Interactions in NOMAD". *Phys. Lett. B* 682 (2009), pp. 177–184. DOI: 10.1016/j.physletb.2009.10.083. arXiv: 0910.0062 [hep-ex].

[172] Y. Kurimoto et al. "Measurement of Inclusive Neutral Current Neutral π^0 Production on Carbon in a Few-GeV Neutrino Beam". *Phys. Rev. D* 81 (2010), p. 033004. DOI: 10.1103/PhysRevD.81.033004. arXiv: 0910.5768 [hep-ex].

[173] Y. Kurimoto et al. "Improved Measurement of Neutral Current Coherent π^0 Production on Carbon in a Few-GeV Neutrino Beam". *Phys. Rev. D* 81 (2010), p. 111102. DOI: 10.1103/PhysRevD.81.111102. arXiv: 1005.0059 [hep-ex].

[174] R. Acciarri et al. "Measurement of ν_μ and $\bar{\nu}_\mu$ neutral current $\pi^0 \to \gamma\gamma$ production in the ArgoNeuT detector". *Phys. Rev. D* 96.1 (2017), p. 012006. DOI: 10.1103/PhysRevD.96.012006. arXiv: 1511.00941 [hep-ex].

[175] P. Adamson et al. "Measurement of single π^0 production by coherent neutral-current ν Fe interactions in the MINOS Near Detector". *Phys. Rev. D* 94.7 (2016), p. 072006. DOI: 10.1103/PhysRevD.94.072006. arXiv: 1608.05702 [hep-ex].

[176] J. Wolcott et al. "Evidence for Neutral-Current Diffractive π^0 Production from Hydrogen in Neutrino Interactions on Hydrocarbon". *Phys. Rev. Lett.* 117.11 (2016), p. 111801. DOI: 10.1103/PhysRevLett.117.111801. arXiv: 1604.01728 [hep-ex].

[177] M. A. Acero et al. "Measurement of neutrino-induced neutral-current coherent π^0 production in the NOvA near detector". *Phys. Rev. D* 102.1 (2020), p. 012004. DOI: 10.1103/PhysRevD.102.012004. arXiv: 1902.00558 [hep-ex].

[178] P. Abratenko et al. "Measurement of neutral current single $\pi 0$ production on argon with the MicroBooNE detector". *Phys. Rev. D* 107.1 (2023), p. 012004. DOI: 10.1103/PhysRevD.107.012004. arXiv: 2205.07943 [hep-ex].

[179] P. Abratenko et al. "First double-differential cross section measurement of neutral-current π^0 production in neutrino-argon scattering in the MicroBooNE detector" (Apr. 2024). arXiv: 2404.10948 [hep-ex].

[180] 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 [physics.ins-det].

[181] A. A. Aguilar-Arevalo et al. "Neutrino flux prediction at MiniBooNE". *Phys. Rev. D* 79 (7 Apr. 2009), p. 072002. DOI: 10.1103/PhysRevD.79.072002.

[182] C. Rubbia. *The liquid-argon time projection chamber: a new concept for neutrino detectors*. Tech. rep. Geneva: CERN, 1977. URL: https://cds.cern.ch/record/117852.

[183] B. Baller et al. "Liquid Argon Time Projection Chamber Research and Development in the United States". *JINST* 9 (2014), T05005. DOI: 10.1088/1748-0221/9/05/T05005. arXiv: 1307.8166 [physics.ins-det].

[184] C. Anderson et al. "The ArgoNeuT Detector in the NuMI Low-Energy beam line at Fermilab". *JINST* 7 (2012), P10019. DOI: 10.1088/1748-0221/7/10/P10019. arXiv: 1205.6747 [physics.ins-det].

[185] W. Foreman et al. "Calorimetry for low-energy electrons using charge and light in liquid argon". *Phys. Rev. D* 101.1 (2020), p. 012010. DOI: 10.1103/PhysRevD.101.012010. arXiv: 1909.07920 [physics.ins-det].

[186] MicroBooNE Website. https://microboone-exp.fnal.gov/public/approved_plots/Event_Displays.html. [Accessed: 8-1-2024].

[187] P. Green. "Light and Dark in Liquid Argon Time Projection Chamber Neutrino Detectors". PhD thesis. University of Manchester, 31 Dec 2022.

[188] S. Kubota et al. "Dynamical behavior of free electrons in the recombination process in liquid argon, krypton, and xenon". *Phys. Rev. B* 20 (8 Oct. 1979), pp. 3486–3496. DOI: 10.1103/PhysRevB.20.3486.

[189] R. Acciarri et al. "A Study of Electron Recombination Using Highly Ionizing Particles in the ArgoNeuT Liquid Argon TPC". *JINST* 8 (2013), P08005. DOI: 10.1088/1748-0221/8/08/P08005. arXiv: 1306.1712 [physics.ins-det].

[190] S. Amoruso et al. "Study of electron recombination in liquid argon with the ICARUS TPC". *Nucl. Instrum. Meth. A* 523 (2004), pp. 275–286. DOI: 10.1016/j.nima.2003.11.423.

[191] J. B. Birks. "Scintillations from Organic Crystals: Specific Fluorescence and Relative Response to Different Radiations". *Proceedings of the Physical Society. Section A* 64.10 (Oct. 1951), p. 874. DOI: 10.1088/0370-1298/64/10/303.

[192] J. Thomas and D. A. Imel. "Recombination of electron-ion pairs in liquid argon and liquid xenon". *Phys. Rev. A* 36 (2 July 1987), pp. 614–616. DOI: 10.1103/PhysRevA.36.614.

[193] R. Acciarri et al. "Effects of Nitrogen contamination in liquid Argon". *JINST* 5 (2010), P06003. DOI: 10.1088/1748-0221/5/06/P06003. arXiv: 0804.1217 [nucl-ex].

[194] R. Acciarri et al. "Oxygen contamination in liquid Argon: Combined effects on ionization electron charge and scintillation light". *JINST* 5 (2010), P05003. DOI: 10.1088/1748-0221/5/05/P05003. arXiv: 0804.1222 [nucl-ex].

[195] T. Doke et al. "Absolute Scintillation Yields in Liquid Argon and Xenon for Various Particles". *Japanese Journal of Applied Physics* 41.3R (Mar. 2002), p. 1538. DOI: 10.1143/JJAP.41.1538.

[196] M. Miyajima et al. "Average energy expended per ion pair in liquid argon". *Phys. Rev. A* 9 (1974), pp. 1438–1443. DOI: 10.1103/PhysRevA.9.1438.

[197] A. Lister and M. Stancari. "Investigations on a fuzzy process: effect of diffusion on calibration and particle identification in Liquid Argon Time Projection Chambers". *JINST* 17.07 (2022), P07016. DOI: 10.1088/1748-0221/17/07/P07016. arXiv: 2201.09773 [physics.ins-det].

[198] P. Abratenko et al. "Measurement of the longitudinal diffusion of ionization electrons in the MicroBooNE detector". *Journal of Instrumentation* 16.09 (Sept. 2021), P09025. DOI: 10.1088/1748-0221/16/09/P09025.

[199] M. Torti and on behalf of the–ICARUS Collaboration. "Electron diffusion measurements in the ICARUS T600 detector". *Journal of Physics: Conference Series* 888.1 (Sept. 2017), p. 012060. DOI: 10.1088/1742-6596/888/1/012060.

[200] Y. Li et al. "Measurement of Longitudinal Electron Diffusion in Liquid Argon". *Nucl. Instrum. Meth. A* 816 (2016), pp. 160–170. DOI: 10.1016/j.nima.2016.01.094. arXiv: 1508.07059 [physics.ins-det].

[201] "Measurement of the Electronegative Contaminants and Drift Electron Lifetime in the MicroBooNE Experiment" (May 2016). DOI: 10.2172/1573040.

[202] MicroBooNE. "A Measurement of the Attenuation of Drifting Electrons in the MicroBooNE LArTPC" (Aug. 2017). DOI: 10.2172/1573054.

[203] M. Babicz et al. "A measurement of the group velocity of scintillation light in liquid argon". *JINST* 15.09 (2020), P09009. DOI: 10.1088/1748-0221/15/09/P09009. arXiv: 2002.09346 [physics.ins-det].

[204] A. Bettini et al. "The ICARUS liquid argon TPC: a complete imaging device for particle physics". *Nucl. Instrum. Methods Phys. Res., A* 315 (1992), pp. 223–228. DOI: 10.1016/0168-9002(92)90707-B.

[205] S. Amerio et al. "Design, construction and tests of the ICARUS T600 detector". *Nucl. Instrum. Meth. A* 527 (2004), pp. 329–410. DOI: 10.1016/j.nima.2004.02.044.

[206] M. Antonello et al. "Search for anomalies in the ν_e appearance from a ν_μ beam". *Eur. Phys. J. C* 73 (2013), p. 2599. DOI: 10.1140/epjc/s10052-013-2599-z. arXiv: 1307.4699 [hep-ex].

[207] M. Antonello et al. "Measurement of the neutrino velocity with the ICARUS detector at the CNGS beam". *Phys. Lett. B* 713 (2012), pp. 17–22. DOI: 10.1016/j.physletb.2012.05.033. arXiv: 1203.3433 [hep-ex].

[208] R. Acciarri et al. "Measurements of Inclusive Muon Neutrino and Antineutrino Charged Current Differential Cross Sections on Argon in the NuMI Antineutrino Beam". *Phys. Rev. D* 89.11 (2014), p. 112003. DOI: 10.1103/PhysRevD.89.112003. arXiv: 1404.4809 [hep-ex].

[209] R. Acciarri et al. "First measurement of electron neutrino scattering cross section on argon". *Phys. Rev. D* 102.1 (2020), p. 011101. DOI: 10.1103/PhysRevD.102.011101. arXiv: 2004.01956 [hep-ex].

[210] R. Acciarri et al. "The Liquid Argon In A Testbeam (LArIAT) Experiment". *JINST* 15.04 (2020), P04026. DOI: 10.1088/1748-0221/15/04/P04026. arXiv: 1911.10379 [physics.ins-det].

[211] E. Gramellini et al. "Measurement of the π–Ar total hadronic cross section at the LArIAT experiment". *Phys. Rev. D* 106.5 (2022), p. 052009. DOI: 10.1103/PhysRevD.106.052009. arXiv: 2108.00040 [hep-ex].

[212] P. Abratenko et al. "ICARUS at the Fermilab Short-Baseline Neutrino program: initial operation". *Eur. Phys. J. C* 83.6 (2023), p. 467. DOI: 10.1140/epjc/s10052-023-11610-y. arXiv: 2301.08634 [hep-ex].

[213] B. Abi et al. "Deep Underground Neutrino Experiment (DUNE), Far Detector Technical Design Report, Volume I Introduction to DUNE". *JINST* 15.08 (2020), T08008. DOI: 10.1088/1748-0221/15/08/T08008. arXiv: 2002.02967 [physics.ins-det].

[214] R. Acciarri et al. "Construction of precision wire readout planes for the Short-Baseline Near Detector (SBND)". *JINST* 15.06 (2020), P06033. DOI: 10.1088/1748-0221/15/06/P06033. arXiv: 2002.08424 [physics.ins-det].

[215] R. Acciarri et al. "Noise Characterization and Filtering in the MicroBooNE Liquid Argon TPC". *JINST* 12.08 (2017), P08003. DOI: 10.1088/1748-0221/12/08/P08003. arXiv: 1705.07341 [physics.ins-det].

[216] M. Kuźniak and A. M. Szelc. "Wavelength Shifters for Applications in Liquid Argon Detectors". *Instruments* 5.1 (2020), p. 4. DOI: 10.3390/instruments5010004. arXiv: 2012.15626 [physics.ins-det].

[217] A. Machado and E. Segreto. "ARAPUCA a new device for liquid argon scintillation light detection". *Journal of Instrumentation* 11.02 (Feb. 2016), p. C02004. DOI: 10.1088/1748-0221/11/02/C02004.

[218] A. A. Machado et al. "The X-ARAPUCA: An improvement of the ARAPUCA device". Apr. 2018. DOI: 10.1088/1748-0221/13/04/C04026. arXiv: 1804.01407 [physics.ins-det].

[219] V. Pandey. *Private Communication.*

[220] H. Lay and K. Lin. *Internal Documentation.* https://sbn-docdb.fnal.gov/cgi-bin/sso/ShowDocument?docid=32827.

[221] J. Serrano et al. "The White Rabbit Project". *Proceedings of the 12th International Conference on Accelerator and Large Experimental Physics Control Systems.* Vol. 1. 2009, p. 93. ISBN: 978-4-9905391-0-8. URL: https://accelconf.web.cern.ch/icalepcs2009/papers/tuc004.pdf.

[222] D. Brailsford. "Physics Program of the Short-Baseline Near Detector". *J. Phys. Conf. Ser.* 888.1 (2017), p. 012186. DOI: 10.1088/1742-6596/888/1/012186.

[223] P. Ballett, M. Hostert, and S. Pascoli. "Dark neutrinos and a three-portal connection to the standard model". *Phys. Rev. D* 101 (11 June 2020), p. 115025. DOI: 10.1103/PhysRevD.101.115025.

[224] I. D. I. Astiz. "Dark neutrino search with the short-baseline near detector" (Aug. 2023). URL: https://sussex.figshare.com/articles/thesis/Dark_neutrino_search_with_the_short-baseline_near_detector/24018792.

[225] K. J. Kelly and P. A. N. Machado. "MicroBooNE experiment, NuMI absorber, and heavy neutral leptons". *Phys. Rev. D* 104 (5 Sept. 2021), p. 055015. DOI: 10.1103/PhysRevD.104.055015.

[226] P. Abratenko et al. "Search for a Higgs Portal Scalar Decaying to Electron-Positron Pairs in the MicroBooNE Detector". *Phys. Rev. Lett.* 127.15 (2021), p. 151803. DOI: 10.1103/PhysRevLett.127.151803. arXiv: 2106.00568 [hep-ex].

[227] S. Bhadra et al. "Letter of Intent to Construct a nuPRISM Detector in the J-PARC Neutrino Beamline" (Dec. 2014). arXiv: 1412.3086 [physics.ins-det].

[228] R. Postel. *Private Communication.*

[229] E. L. Snider and G. Petrillo. "LArSoft: Toolkit for Simulation, Reconstruction and Analysis of Liquid Argon TPC Neutrino Detectors". *J. Phys. Conf. Ser.* 898.4 (2017). Ed. by R. Mount and C. Tull, p. 042057. DOI: 10.1088/1742-6596/898/4/042057.

[230] S. Agostinelli et al. "GEANT4–a simulation toolkit". *Nucl. Instrum. Meth. A* 506 (2003), pp. 250–303. DOI: 10.1016/S0168-9002(03)01368-8.

[231] M. G. Catanesi et al. "Measurement of the production cross-section of positive pions in the collision of 8.9-GeV/c protons on beryllium". *Eur. Phys. J. C* 52 (2007), pp. 29–53. DOI: 10.1140/epjc/s10052-007-0382-8. arXiv: hep-ex/0702024.

[232] I. Chemakin et al. "Pion production by protons on a thin beryllium target at 6.4, 12.3, and 17.5 GeV/c incident proton momenta". *Phys. Rev. C* 77 (1 Jan. 2008), p. 015209. DOI: 10.1103/PhysRevC.77.015209.

[233] J. Tena-Vidal et al. "Neutrino-nucleon cross-section model tuning in GENIE v3". *Phys. Rev. D* 104 (7 Oct. 2021), p. 072009. DOI: 10.1103/PhysRevD.104.072009.

[234] K. S. Egiyan et al. "Measurement of 2- and 3-nucleon short range correlation probabilities in nuclei". *Phys. Rev. Lett.* 96 (2006), p. 082501. DOI: 10.1103/PhysRevLett.96.082501. arXiv: nucl-ex/0508026.

[235] J. Nieves, I. R. Simo, and M. J. V. Vacas. "Inclusive charged-current neutrino-nucleus reactions". *Phys. Rev. C* 83 (4 Apr. 2011), p. 045501. DOI: 10.1103/PhysRevC.83.045501.

[236] G. D. Megias et al. "Meson-exchange currents and quasielastic predictions for charged-current neutrino-^{12}C scattering in the superscaling approach". *Phys. Rev. D* 91.7 (2015), p. 073004. DOI: 10.1103/PhysRevD.91.073004. arXiv: 1412.1822 [nucl-th].

[237] L. A. Soplin et al. "Monte Carlo Simulation Development and Implementation of the GiBUU Model for Neutrino Experiments" (Nov. 2023). arXiv: 2311.14286 [hep-ex].

[238] K. Gallmeister, U. Mosel, and J. Weil. "Neutrino-Induced Reactions on Nuclei". *Phys. Rev. C* 94.3 (2016), p. 035502. DOI: 10.1103/PhysRevC.94.035502. arXiv: 1605.09391 [nucl-th].

[239] D. Heck et al. *CORSIKA: a Monte Carlo code to simulate extensive air showers.* 1998.

[240] C. Adams et al. "Ionization electron signal processing in single phase LArTPCs. Part I. Algorithm Description and quantitative evaluation with MicroBooNE simulation". *JINST* 13.07 (2018), P07006. DOI: 10.1088/1748-0221/13/07/P07006. arXiv: 1802.08709 [physics.ins-det].

[241] C. Adams et al. "Ionization electron signal processing in single phase LArTPCs. Part II. Data/simulation comparison and performance in MicroBooNE". *JINST* 13.07 (2018), P07007. DOI: 10.1088/1748-0221/13/07/P07007. arXiv: 1804.02583 [physics.ins-det].

[242] D. Garcia-Gamez, P. Green, and A. M. Szelc. "Predicting Transport Effects of Scintillation Light Signals in Large-Scale Liquid Argon Detectors". *Eur. Phys. J. C* 81.4 (2021), p. 349. DOI: 10.1140/epjc/s10052-021-09119-3. arXiv: 2010.00324 [physics.ins-det].

[243] L. Tung and M. J. Jung. *Signal Processing with WireCell in SBND*. FERMILAB-POSTER-23-121-V. DOI: 10.2172/1989875.

[244] T. Usher. *GausHitFinder Hit Finding and Fitting*. https://indico.fnal.gov/event/18523/contributions/47921/attachments/29850/36753/GaussHitFinder121118.pdf. [Accessed: 18-12-2024].

[245] E. Tyley. "Reconstructing and Selecting Electron Neutrino and Anti-Neutrino Interactions on Argon in the Short-Baseline Near Detector". PhD thesis. University of Sheffield, Feb. 2023. URL: https://etheses.whiterose.ac.uk/32440/.

[246] J. S. Marshall and M. A. Thomson. "The Pandora Software Development Kit for Pattern Recognition". *Eur. Phys. J. C* 75.9 (2015), p. 439. DOI: 10.1140/epjc/s10052-015-3659-3. arXiv: 1506.05348 [physics.data-an].

[247] R. Acciarri et al. "The Pandora multi-algorithm approach to automated pattern recognition of cosmic-ray muon and neutrino events in the MicroBooNE detector". *Eur. Phys. J. C* 78.1 (2018), p. 82. DOI: 10.1140/epjc/s10052-017-5481-6. arXiv: 1708.03135 [hep-ex].

[248] R. E. Schapire. "The boosting approach to machine learning: An overview". *Nonlinear estimation and classification* (2003), pp. 149–171.

[249] L. Breiman et al. "Classification and Regression Trees (1st ed.)" (1984). DOI: https://doi.org/10.1201/9781315139470.

[250] Y. Coadou. "Boosted decision trees" (Mar. 2022). DOI: 10.1142/9789811234033_0002. arXiv: 2206.09645 [physics.data-an].

[251] M. Greenacre et al. "Principal component analysis". *Nature Reviews Methods Primers* 2 (Dec. 2022), p. 100. DOI: 10.1038/s43586-022-00184-w.

[252] L. Han and J. C. Bancroft. "Nearest approaches to multiple lines in n-dimensional space". *CREWES Res. Rep* 22 (2010), pp. 1–17.

[253] H. Seo et al. "Overthere: A Simple and Intuitive Object Registration Method for an Absolute Mid-air Pointing Interface". *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5.3 (Sept. 2021). DOI: 10.1145/3478128. URL: https://doi.org/10.1145/3478128.

[254] P. Abratenko et al. "Scintillation Light in SBND: Simulation, Reconstruction, and Expected Performance of the Photon Detection System" (June 2024). arXiv: 2406.07514 [physics.ins-det].

[255] I. Kreslo, M. Auger, and D. Lorca. *Internal Documentation*. https://sbn-docdb.fnal.gov/cgi-bin/sso/ShowDocument?docid=685.

[256] J. Autran et al. "Characterization of atmospheric muons at sea level using a cosmic ray telescope". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 903 (2018), pp. 77–84. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2018.06.038.

[257] W. Bragg and R. Kleeman. "LXXIV. On the ionization curves of radium". *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 8.48 (1904), pp. 726–738. DOI: 10.1080/14786440409463246. eprint: https://doi.org/10.1080/14786440409463246.

[258] A. Hocker et al. "TMVA - Toolkit for Multivariate Data Analysis" (Mar. 2007). arXiv: physics/0703039.

[259] J. W. Pratt and J. D. Gibbons. "Kolmogorov-Smirnov Two-Sample Tests". *Concepts of Nonparametric Theory*. New York, NY: Springer New York, 1981, pp. 318–344. ISBN: 978-1-4612-5931-2. DOI: 10.1007/978-1-4612-5931-2_7.

[260] G. Putnam. *Internal Documentation*. https://sbn-docdb.fnal.gov/cgi-bin/sso/ShowDocument?docid=21438.

[261] C. Adams et al. "Reconstruction and Measurement of $\mathcal{O}(100)$ MeV Energy Electromagnetic Activity from $\pi^0 \to \gamma\gamma$ Decays in the MicroBooNE LArTPC". *JINST* 15.02 (2020), P02007. DOI: 10.1088/1748-0221/15/02/P02007. arXiv: 1910.02166 [hep-ex].

[262] M. Lüthi. "A Novel UV Laser System for Electric Field Calibration in Liquid Argon Time Projection Chambers". PhD thesis. Bern U., 2019. DOI: 10.7892/boris.141465.

[263] G. R. Lynch and O. I. Dahl. "Approximations to multiple Coulomb scattering". *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* 58.1 (1991), pp. 6–10. ISSN: 0168-583X. DOI: https://doi.org/10.1016/0168-583X(91)95671-Y. URL: https://www.sciencedirect.com/science/article/pii/0168583X9195671Y.

[264] L. Koch. "A response-matrix-centred approach to presenting cross-section measurements". *Journal of Instrumentation* 14.09 (2019), P09013.

[265] M. B. Avanzini et al. "Comparisons and challenges of modern neutrino-scattering experiments". *Phys. Rev. D* 105 (9 May 2022), p. 092004. DOI: 10.1103/PhysRevD.105.092004.

[266] S. Gardiner. "Mathematical methods for neutrino cross-section extraction" (Jan. 2024). arXiv: 2401.04065 [hep-ex].

[267] G. D'Agostini. "A multidimensional unfolding method based on Bayes' theorem". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 362.2 (1995), pp. 487–498. ISSN: 0168-9002. DOI: https://doi.org/10.1016/0168-9002(95)00274-X.

[268] W. Tang et al. "Data Unfolding with Wiener-SVD Method". *JINST* 12.10 (2017), P10002. DOI: 10.1088/1748-0221/12/10/P10002. arXiv: 1705.03568 [physics.data-an].

[269] BNL LAr Group. *Liquid Argon Properties (Tables and Calculators)*. https://lar.bnl.gov/properties/. [Accessed: 4-4-2024].

[270] C. Tegeler, R. Span, and W. 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 (May 1999), pp. 779–850. ISSN: 0047-2689. DOI: 10.1063/1.556037. eprint: https://pubs.aip.org/aip/jpr/article-pdf/28/3/779/8183487/779_1_online.pdf.

[271] K. Yang. "Measurement of the pion charge exchange differential cross section on Argon with the ProtoDUNE-SP detector". PhD thesis. Oxford University, 2023.

[272] R. Brun et al. *ROOT Project*. Version v6-18-02. June 2020. DOI: 10.5281/zenodo.3895860.

[273] F. James and M. Roos. "Minuit - a system for function minimization and analysis of the parameter errors and correlations". *Computer Physics Communications* 10.6 (1975), pp. 343–367. ISSN: 0010-4655. DOI: https://doi.org/10.1016/0010-4655(75)90039-9.

[274] N. Abgrall et al. "NA61/SHINE facility at the CERN SPS: beams and detector system". *Journal of Instrumentation* 9.06 (June 2014), P06005. DOI: 10.1088/1748-0221/9/06/P06005.

[275] C. Andreopoulos et al. "The GENIE Neutrino Monte Carlo Generator: Physics and User Manual" (Oct. 2015). arXiv: 1510.05494 [hep-ph].

[276] A. Papadopoulou. *Private Communication.*

[277] J. Calcutt et al. "Geant4Reweight: a framework for evaluating and propagating hadronic interaction uncertainties in Geant4". *Journal of Instrumentation* 16.08 (Aug. 2021), P08042. DOI: 10.1088/1748-0221/16/08/P08042.

[278] P. Abratenko et al. "Novel approach for evaluating detector-related uncertainties in a LArTPC using MicroBooNE data". *Eur. Phys. J. C* 82.5 (2022), p. 454. DOI: 10.1140/epjc/s10052-022-10270-8. arXiv: 2111.03556 [hep-ex].

[279] MicroBooNE Collaboration. *MICROBOONE-NOTE-1045-PUB. First Muon-Neutrino Charged-Current Inclusive Differential Cross Section Measurement for MicroBooNE Run 1 Data.* https://microboone.fnal.gov/wp-content/uploads/MICROBOONE-NOTE-1045-PUB.pdf. [Accessed 30-03-2024].

[280] MicroBooNE Collaboration. *MICROBOONE-NOTE-1111-PUB. Measurement of Neutral-Current π^0 Cross Section in MicroBooNE using Wire-Cell Reconstruction.* https://microboone.fnal.gov/wp-content/uploads/MICROBOONE-NOTE-1111-PUB.pdf. [Accessed 30-03-2024].

[281] J. Żmuda. "Pion Production in Wroclaw Neutrino Event Generator and its Comparison with Other Generators". *JPS Conf. Proc.* 12 (2016), p. 010023. DOI: 10.7566/JPSCP.12.010023.

[282] P. Stowell et al. "NUISANCE: a neutrino cross-section generator tuning and comparison framework". *Journal of Instrumentation* 12.01 (Jan. 2017), P01016. DOI: 10.1088/1748-0221/12/01/P01016.

[283] K. B. McConnel Mahn. "A search for muon neutrino and antineutrino disappearance in the Booster Neutrino Beam". PhD thesis. Columbia University, 2009. DOI: 10.2172/959057.

[284] D. Margalit, J. Rabinoff, and L. Rolen. "Interactive Linear Algebra (Chapter 6.5)". *Georgia Institute of Technology* (2017).