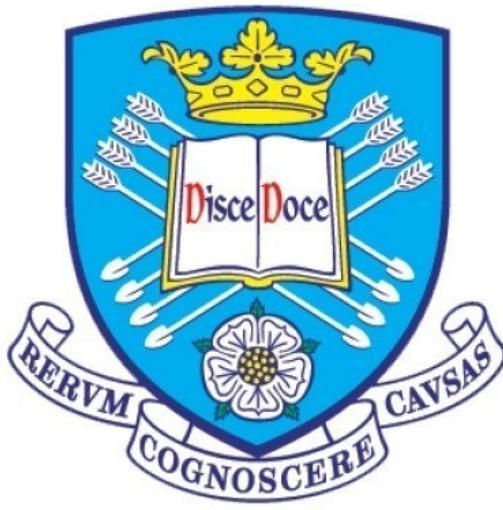


# **Developments Towards a $\nu_e$ CC Sterile Appearance Sensitivity in the Short-Baseline Neutrino Programme**



**Dominic Barker**

Supervisor: Prof. Neil Spooner

Department of Physics & Astronomy  
University of Sheffield

This dissertation is submitted for the degree of

*Doctor of Philosophy*

July 2020



I would like to dedicate this thesis to my loving parents ...



## **Declaration**

The author hereby declares that except where specific reference is made to the work of others, the content of this thesis is original and have not been submitted for any other degree or qualification in this, or any other university.

Chapters 2 and 3 reference and review the current status of neutrino theory and experiment as well as describing the functionality of liquid argon time projection chambers. As such, the author has not contributed to the work within these chapters and wherever possible, the author has referenced the work done by others.

Chapter 4 discusses the design and physics capabilities of the SBN programme. The author arrived after the design concepts and proposal for the programme were submitted. As such, the author supported the collaboration during installation, software development and the R&D research period of the detector. This predominately involved creating CRT modules. The author was also the documentation manager for the SBND software websites. This involved creating and maintaining a Monte Carlo production website as well as software wiki website for new beginners.

Chapter 5 goes into detail of electronics tests that were performed for SBND prior to installation. The author was involved in the development of the online monitoring and data acquisition, in particular the implementation of a hit finding algorithm, the maintenance of the file transfer system and the associated Postgres server, as well as the electron lifetime study within the Vertical Slices Test (VST). The author would like to note that there involved was minor in places and author would like to thank and acknowledge all involved. The author would especially like to thank and acknowledge Gray Putnam and Andy Mastbaum, who wrote the framework to the online monitoring software to which the author added small developments. The author would also like to acknowledge and thank Randy Johnson as well as Scientific Computing Division team, who initially set up the file transfer system for LArIAT and did the majority of the changes for the VST. The author would especially like to thank Bill Badgett for his support. The author would also like to thank Tom Brooks who, wrote the majority of the lifetime algorithm for the VST.

The author was the sole contributor to the signal to noise in Chapter 5 measurement but would like to thank and acknowledge, Minerba Betancourt, Brian Rebel and Shanshan

Gao for their guidance and collaboration. The author would like to acknowledge and thank Michelle Stancari for her guidance and help when extrapolating the analysis to the size of SBND. The signal to analysis of the ICARUS-SBND hybrid test stand was the original work of the author. However, the author did not contribute to the data acquisition and would therefore like to acknowledge and thank Francesco Pietropaolo, Angela Fava and others for this. The author would especially like to thank Michelle Stancari, Angela Fava and Palamara Ornella Palamara for their support during the analysis.

For the electron lifetime studies in SBND, discussed in Chapter 5 the author would like to note that Tom Brooks was the primary author for this work and that the authors contribution was small in comparison. The author would like to thank Tom Brooks and Michelle Stancari as the author would still be lost without the knowledge that was gained during the analysis.

For Chapter 6, the author would like to acknowledge the fact that the reconstruction and simulation is the result of many collaborators within the LArSoft framework. Where possible references have been provided. The hit Monte Carlo hit finding study presented in the chapter was the work of the author and the author would like to thank Michelle Stancari for her discussion and guidance. The author also implemented the Blurred Clustering and EMShower module in the SBND reconstruction and the author would like to thank and acknowledge Mike Wallbank for the original development of the software and the discussions provided during the implementation. The author was also involved in the implementation of machine learning in the Pandora Pattern recognition. The author would like to thank the Pandora Team for the pattern recognition software and especially Jack Anthony, who wrote the software for the machine learning. The author would like to note that their contribution to the machine learning was minimal and the worked was passed onto Ed Tyley very swiftly. Therefore, the author would like to thank and acknowledge Ed Tyley.

The author is the lead author on the Tool-based Reconstruction Algorithm for Characterising Showers (TRACS), the validation analysis and contributed to the majority of the framework and the tools used, discussed in 6. The author would like to note that some of the tools are recreations of existing shower reconstruction software and therefore would like to thank and acknowledge Mike Wallbank and the Pandora team again for this. The author would also like to acknowledge and thank Dom Brailsford and Ed Tyley for their collaboration within the shower reconstruction team. The author would also like to acknowledge and thank Ed Tyley for migration of TRACS into the Pandora framework where it will be known as the PandoraModularShowerCreation and will be used by several other Liquid Argon experiments. The author would also like to acknowledge and thank Bruce Howard for his incorporation of the TRACS software into ICARUS and Dom Brailsford for his incorporation into the DUNE far detector software. Finally, the author would again like

to thank Dom Brailsford for his guidance when created the software and for performing the Monte Carlo production of the events used in the shower validation.

For Chapter 7, the author would like to acknowledge the fact that the reconstruction and simulation is the result of many collaborators within the LArSoft framework. Therefore the results of the Chapter are as a result of many peoples collaborative efforts. The author is the sole contributor to the selection work discussed in the Chapter. However, the work was based on the proposal selection to which author would like to acknowledge Corey Adams, Joseph Zennamo and others. The author created a small framework for the reconstruction selection and developed the current infrastructure used in SBN analyses to make grid submission easier. The author would also to acknowledge Gray Putnam for generating the  $\nu_\mu$  samples which were used as a template for generating other events. For the reconstruction selection, the author would like to thank Dom Brailsford, Chris Farnese, Ornella Palamara, Michelle Stancari and others for there discussion and guidance for improving the selection. The author would also like to acknowledge the SBN production team: Dom Brailsford, Ed Tyley and Gray Putnam for the providing the reconstruction samples. Finally, the author would like to acknowledge Bruce Howard for performing the selection in ICARUS, noticing bugs and validating the performabce in ICARUS. Finally, the  $dE/dx$  truth study was perfomred fully by Ed Tyley.

For Chapter 8, the author simulated the oscillated signal and the dirt, cosmic and intrinsic backgrounds, reconstructed events in the three experiments, performed the selection and performed all parts the  $\nu_e$  appearance analysis described in the Chapter. The author would like to acknowledge Gray Putnam and Andy Mastbaum for the SBN selection infrastructure. The author would like to thank and acknowledge the two aforementioned people again for implemented the event reweighting framework. For the sensitivity study, the author would like to acknowledge the VALOR fitting group for their initial developments in the DUNE near detector such that an SBN sensitivity was easy to implement. The author would like to acknowledge and thank especially Steve Dennis for his development and guidance in the SBN studies and Rhiannon Jones for her developments in reading in the SBN infrastructure output and support. The author implemented software such that the VALOR fitting framework could be easily used on the Fermigrid. Also, the author improved the processing speed of the code and also changed the software to use the true baseline of events rather than a fixed baseline in oscillation calculations. The author also added infrastructure to include cosmic and dirt events in the fit. Finally, the author would like to thank Dom Brailsford for his guidance during the analysis.

For Chapter 9 the author would like to acknowledge and thank Ed Tyley, Anthony Ezeribe, Andrew Scarff and Ala Zglam for their contributions to the experimental setup

and the massive contribution of hours to get the Sheffield rig setup. The author would also like to acknowledge and thank the University of Bern, particularly: Damian Goeldi, Camilla Tognina, Igor Kreslo, James Sinclair and Roman Berner who provided the electronic designs for their pixel board and LArASIC chip, provided software to run the chip and also provided a large amount of support. The author would like thank BNL for supply LArASIC chips. The author would like to especially acknowledge and thank Matt Wright for his development of the cryogenics of the rig and continuing support in maintaining and using the rig infrastructure. His support is invaluable to the group. The author would also like to acknowledge and thank Matt Thiesse and Nicola McConkey for their work prior to the setup and support. The Keysight DAQ readout and analysis as well as the high voltage I/O was originally written by Matt Thiesse. The setup of the rig and the slow controls readout was done by Nicola McConkey. The author was one the lead contributors to the experimental setup and was directly involved in the majority of the design of the experiment (not the cryogenic system), installation and testing. The author's main contributions include the installation of the experiment, writing the CAEN DAQ readout software and ASIC API, the Postgres online monitoring sever and display, the software infrastructure, advancements of existing code, calibrations and testing of the setup. The author also provided support in the installation and data acquisition of the PixLAr experiment where the pixel readout was tested in LArIAT. The author would also like to acknowledge and thank Neil Spooner for the advice and facilitation of the R&D activities at Sheffield.

Dominic Barker  
July 2020

## Acknowledgements

Firstly I would like to thank my supervisor, Prof. Neil Spooner, for his support and guidance throughout the PhD. His regular attention has kept me on track and I could not thank him enough for the opportunity.

There have been many people during this PhD who have provided tremendous support. For the particulars, I implore the reader to have a quick read of the declaration which does not do justice to the support I have received. I would especially like to thank Dom Brailsford, Ornella Palamara and Michelle Stancari for their support and guidance over several stages of my PhD. I would also like to thank Tom Brooks, Nicola McConkey and Mike Wallbank for initially helping me in my first few months where everything was alien. I would like to thank Dom Brailsford and Ed Tyley for their contributions, discussion and support within the shower reconstruction group.

I would also like to give a dedicated thank you to the people that helped with Sheffield liquid argon rig. These people sank many hours with me working on the project. Without the knowledge and support of these people I would have never left the room: Ed Tyley, Anthony Ezeribe, Andrew Scarff, Ala Zglam, Matt Thiesse, Nicola McConkey and last but by no means least Matt Wright.

I also met many people during my PhD and had some unforgettable experiences with them. For that and their support throughout I would like to thank Tom, Rhiannon, Andy, Pete, Warren, Dom, Ed, Andy, Anthony, Steve, Nic, Matt and Matt. I could go on so thank you to all who were not mentioned.

Finally, I would like to thank my family for the unconditional support. My mum and dad have been terrific. They have been there from the start to the finish of this PhD and beyond and I can not thank them enough. I would also like to thank my partner Naomi for her unwavering support and her extensive efforts to help make this work readable. Her support, especially in the last few months of the PhD, where I must have been intolerable has been fantastic.



## Abstract

The Short Baseline Neutrino (SBN) programme is an upcoming neutrino experiment situated on the Booster Neutrino Beam (BNB) at the Fermilab National Laboratory. One of the primary objectives of the SBN programme is to confirm or refute the low energy electron neutrino excess observed in previous neutrino experiments: LSND and MiniBooNE. It was postulated that this observed low energy electron neutrino excess was caused by the existence of one or more sterile neutrinos. If this is confirmed, it will alter our current understanding of physics as well as the standard model and the prescription of neutrino oscillations.

To achieve this primary objective, the SBN programme will perform studies which are sensitive to electron neutrino appearances. These are carried out assuming several sterile models, in particular the  $3 + 1$  model. To undertake the physics goals of the SBN programme, three Liquid Argon Time Projection Chambers (LArTPCs) are positioned at various points along the BNB beamline. These LArTPCs are known as The Short Baseline Near Detector (SBND) (110 m), Micro Booster Neutrino Experiment (MicroBooNE) (470 m), and the Imaging Cosmic And Rare Underground Signals (ICARUS) (600 m) detector. LArTPCs provide sophisticated calorimetric and topological information to identify the energy and flavour of charged particles in neutrino interactions. For an electron neutrino excess search, it is important to reconstruct and identify the resultant electron from neutrino Charge Current (CC) events. A new framework with new methods was developed to characterise electromagnetic showers to help identify electrons from background photon showers. The new methods were then employed in an oscillated electron neutrino selection upon simulated events in the SBN detectors. The resultant event distributions were then used to perform an electron neutrino appearance sensitivity analysis using the  $3 + 1$  sterile model in the VALencia-Oxford-Rutherford (VALOR) neutrino oscillation fitting framework.

The single-phase wire near detector of the SBN programme, SBND, is also viewed as a prototype for the upcoming Deep Underground Neutrino Experiment (DUNE) far detector. Due to the high rate of events at the location of the DUNE near detector, single-phase wire LArTPCs are not feasible. Therefore, alternative readout methods are being considered, such as a pixelated readout. To test these alternative readout methods, a research rig at the University of Sheffield has been developed.



# Table of contents

<b>List of figures</b>	<b>xvii</b>
<b>List of tables</b>	<b>xlix</b>
<b>Nomenclature</b>	<b>li</b>
<b>1 Introduction</b>	<b>1</b>
<b>2 The Case for the Sterile Neutrino: An Overview of Neutrino Physics</b>	<b>3</b>
2.1 Historical Overview of Neutrino Physics Theory . . . . .	4
2.2 Beyond the Standard Model in Neutrino Physics . . . . .	10
2.2.1 Neutrino Oscillations . . . . .	10
2.2.2 Experimental Search for Neutrino Mass . . . . .	19
2.2.3 The Seesaw Mechanism . . . . .	22
2.2.4 Calculating the Number of Neutrinos . . . . .	23
2.2.5 Sterile Neutrino Theory . . . . .	25
2.2.6 Anomalies in Neutrino Experiments . . . . .	26
2.3 Concluding Remarks . . . . .	32
<b>3 Liquid Argon Time Projection Chamber Theory</b>	<b>33</b>
3.1 LArTPC Design . . . . .	34
3.2 Ionisation, Recombination and Scintillation . . . . .	36
3.3 Transportation of Ionised Electrons, Diffusion and Impurities . . . . .	38
3.4 Particle Calorimetry and Identification . . . . .	40
3.5 Other LArTPC Developments . . . . .	42
3.6 Neutrino Interactions and Cross-Sections . . . . .	42
3.7 Concluding Remarks . . . . .	45

<b>4 Overview of The Short-Baseline Neutrino Programme (SBN)</b>	<b>47</b>
4.1 The Booster Neutrino Beam (BNB) . . . . .	48
4.2 Neutrinos at the Main Injector (NuMI) Beam . . . . .	50
4.3 The Short-Baseline Near Detector (SBND) . . . . .	51
4.3.1 Detector Design . . . . .	51
4.3.2 Electronics Design . . . . .	55
4.4 The MicroBooNE Detector . . . . .	57
4.5 The ICARUS Detector . . . . .	60
4.6 Physics Capabilities . . . . .	62
4.6.1 Sensitivity of Sterile Neutrino Searches . . . . .	62
4.6.2 Other Physics Searches . . . . .	67
4.7 Software Used in the SBN Programme Simulation . . . . .	67
<b>5 Studies on the SBND Electronics Chain</b>	<b>69</b>
5.1 VST Overview . . . . .	70
5.2 Signal to Noise Ratio Measurement in the VST . . . . .	75
5.2.1 Raw Hit Finding . . . . .	77
5.2.2 Electron Lifetime Analysis . . . . .	78
5.2.3 Signal to Noise Ratio Analysis . . . . .	80
5.3 VST Signal to Noise Ratio Extrapolation to SBND . . . . .	85
5.4 Warm Icarus Electronics Alternative chain for SBND . . . . .	88
5.4.1 Test Stand Electronic Design . . . . .	89
5.4.2 Event Selection . . . . .	89
5.4.3 Noise Characterisation . . . . .	91
5.4.4 Signal to Noise Ratio Measurement . . . . .	93
5.4.5 ENC Measurement . . . . .	95
5.5 Electron Lifetime Studies in SBND . . . . .	97
5.5.1 Methods to Calculate the Lifetime . . . . .	97
5.5.2 Efficiency and Accuracy of the Methods . . . . .	100
5.5.3 Comparison of the Methods . . . . .	101
5.6 Concluding Remarks . . . . .	102
<b>6 Reconstruction of Particle Interactions in the SBND Detector</b>	<b>103</b>
6.1 Reconstruction Chain in SBND . . . . .	104
6.1.1 Charge Reconstruction . . . . .	104
6.1.2 Hit Finding . . . . .	105
6.1.3 Pandora Pattern Recognition . . . . .	111

---

6.1.4	Alternative Shower Clustering . . . . .	113
6.1.5	Comparison of Shower Clustering Methods . . . . .	114
6.1.6	Track Reconstruction . . . . .	120
6.1.7	Calorimetry . . . . .	121
6.1.8	Particle Identification . . . . .	122
6.2	Shower Characterisation in Liquid Argon TPCs . . . . .	124
6.2.1	Overview of Pre-Existing Shower Reconstruction Software in SBND	126
6.2.2	The TRACS Reconstruction Chain and Performance Evaluation . .	127
6.3	Concluding Remarks . . . . .	153
<b>7</b>	<b>Producing and Selecting Charge Current <math>\nu_e</math> events in the SBN Programme</b>	<b>155</b>
7.1	Sample Production . . . . .	156
7.1.1	Producing the Modern Samples . . . . .	156
7.1.2	Producing the Proposal Samples . . . . .	158
7.2	Truth-Based Proposal Like Selection . . . . .	160
7.2.1	Beam Induced Active Volume Signal and Background Events . .	160
7.2.2	Dirt Events . . . . .	166
7.2.3	Cosmic Removal . . . . .	167
7.3	Comparisons Between the Truth-Based Analyses . . . . .	169
7.4	Evaluation of Reconstructed Selections in SBND . . . . .	174
7.4.1	Reconstruction Samples Used . . . . .	175
7.4.2	Pre-selection Cuts . . . . .	176
7.4.3	Evaluating the Proposal Selection Cuts on Reconstructed Events .	177
7.4.4	Including Additional Metrics in the $\nu_e$ Selection . . . . .	183
7.4.5	Performing a Multivariate Analysis to improve the selections . .	188
7.5	Comparing The Selection . . . . .	200
7.6	Concluding Remarks . . . . .	203
<b>8</b>	<b>SBN Sensitivity to Sterile <math>\nu_e</math> Appearance</b>	<b>205</b>
8.1	The VALOR Framework for the SBN Analysis . . . . .	206
8.2	The $\nu_e$ CC Selection Spectra . . . . .	208
8.3	Flux and Interaction Systematics . . . . .	210
8.4	Results of the Truth-Based Analyses . . . . .	213
8.5	Results of the TPC Cut Based Reconstruction Analyses . . . . .	214
8.6	Concluding Remarks . . . . .	217

<b>9 R&amp;D In New Liquid Argon Readouts</b>	<b>219</b>
9.1 The Sheffield Liquid Argon Rig Design . . . . .	220
9.2 TPC Design and Electronic Readout . . . . .	222
9.3 Sub-detector Setup and Electronic Readout . . . . .	225
9.4 Data Acquisition and Readout Monitoring Software . . . . .	227
9.5 Noise Analysis . . . . .	230
9.6 Outlook . . . . .	232
<b>10 Conclusion</b>	<b>233</b>
<b>References</b>	<b>237</b>
<b>Appendix A Full Hit Finding Efficiencies</b>	<b>259</b>
<b>Appendix B Details of the TRACS framework</b>	<b>265</b>
<b>Appendix C Differences Between GENIE version v2_12 and v2_8</b>	<b>269</b>
<b>Appendix D Including a SVM for PFO Characterisation</b>	<b>275</b>
<b>Appendix E Further Analysis on the Pre-selection Cuts</b>	<b>277</b>
<b>Appendix F Neutrino Energy Reconstruction and Correctness</b>	<b>281</b>
<b>Appendix G Discussion of a Truth-Based <math>dE/dx</math> study</b>	<b>285</b>
<b>Appendix H <math>\nu_e</math> CC Selection as a Function of Interaction Variables</b>	<b>287</b>
<b>Appendix I More Detail on the Additional Selection Metrics</b>	<b>295</b>
<b>Appendix J <math>\nu_e</math> Selection Spectra</b>	<b>307</b>
<b>Appendix K Systematical Covariance Matrices for the <math>\nu_e</math> Analysis</b>	<b>311</b>

# List of figures

2.1	Feynman diagram for the Fermi transition for beta decay. . . . .	4
2.2	The asymmetry in the electron count rate for the change in spin orientation of Co(60) in the Wu et al. experiment. From Reference [1] . . . . .	6
2.3	A Muon track (a) and an electron shower (b) in the Brookhaven spark chamber detector. From Reference [2] . . . . .	8
2.4	Feynman diagrams of the elastic scattering interactions between $\nu_e$ and $e^-$ . . . . .	14
2.5	The solar deficit observed in Kamiokande as a function of the angle from the sun. The solid line indicates the Monte Carlo prediction. From Reference [3].	14
2.6	Electron and muon rates in Super-K as a function of the zenith angle. The hatched squares correspond the Monte Carlo assuming no neutrino oscillations while the black line corresponds to the best fit of the data. From Reference [4]. . . . .	15
2.7	The flux, $\phi_{\mu\tau}$ , of the $\nu_\mu$ and $\nu_\tau$ neutrinos from $B^8$ interactions in the sun against the flux, $\phi_e$ , of $\nu_e$ for the three types of interactions in SNO. The contours indicate the best fit and the dashed lines indicated the predicated flux from the SSM. From Reference [5]. . . . .	16
2.8	Pictorial explanation of the normal and inverted mass hierarchy, from Reference [6]. The contributions of the flavour eigenstates to the mass eigenstates are shown in the coloured bars. For the first mass eigenstate, the main contribution is the $\nu_e$ . For the third mass eigenstate, the $\nu_\mu$ and $\nu_\tau$ contribute more. . . . .	20
2.9	The oscillation probability in reactor experiments with a baseline of 60 km, from Reference [7]. The figure shows the difference in the probability for the normal ordering (black line) and the inverse ordering (dashed line). . . . .	20

---

2.11 The CMB power spectrum and the effect of increasing $N_{\text{eff}}$ . The black line corresponds to the best fit of WMAP7 (black points) (The Wilkinson Microwave Anisotropy Probe 7-year result) and SPT K11 (blue points) (South Pole Telescope, Keisler et al, 2011) data. The top graph fixes the matter-radiation equality redshift ( $z_{eq}$ ), and the angular size of the sound horizon $\theta_S$ as well as baryon density. The bottom graph also keeps the fraction of baryonic mass in helium, $Y_P$ , constant in order to fix the angular size of the photon diffusion, $\theta_D$ . As the difference in the different $N_{\text{eff}}$ is small, when $\theta_D$ is fixed, it shows that the impact is related to the diffusion of photons, i.e. Silk damping. From Reference [8] . . . . .	24
2.12 The LSND excess results in terms of distance travelled in meters divided by the neutrino energy in MeV. The neutrino energy is determined from the measured positron energy and angle with respect to the neutrino beam, from Reference [9]. . . . .	28
2.13 The MiniBooNE excess from both neutrino and anti-neutrino mode data in terms of neutrino energy assuming quasi-elastic scattering, from Reference [10]. . . . .	29
2.14 a) $3\sigma$ sensitivity contours for $\nu_\mu$ disappearance from experiments. b) Global fit of appearance and disappearance data. The Dis line exclude regions by $3\sigma$ whilst App corresponds to $3\sigma$ allowed regions. From Reference [11]. . . . .	30
2.15 $3\sigma$ sensitivity contours for $\nu_e$ appearance from experiments. a) is with all MiniBooNE data and b) is without low energy MiniBooNE data. From Reference [11] . . . . .	30
2.16 a) Allowed regions from reactor experiments. The shaded areas correspond to the allowed regions prior to reactor experiment results from Daya Bay, DANSS and NEOS. The green area corresponds to setting flux spectra to the predicated values, whilst orange allows the flux to be a free parameter. The blue area corresponds to all reactor data with the flux being free and the magenta contours indicate the global fixed analysis. The star represents the best fit. b) Is the reactor global fits overlapped with the gallium allowed region and the exclusion (to the left) from solar neutrino data and atmospheric data from Super-K, DeepCore and IceCube at 95% and 99% CL. Figures are from Reference [12]. . . . .	31
3.1 The LArTPC design. The energy deposits travel toward the readout which is made up of wire planes for 2D reconstruction. Figure from Reference [13].	35

3.2	A Landau distribution. The long asymmetric tail predicts higher energy electrons, known as delta rays, to exist within the TPC. These have been observed. The variable $x$ depends on the energy deposited within a small segment and properties of the detector. From Reference [14]. . . . .	41
3.3	Simulation of a 2 GeV electron (a) and muon (b) in a LArTPC. The image is the raw digitised output from a wire plane with wire number on the x-axis and time in digitised ticks on the y-axis. . . . .	42
3.4	The neutrino (a) and anti-neutrino (b) cross-section measurements as a function of neutrino energy. The prediction for the Quasi-Elastic (QE), Resonance (RES) and Deep Inelastic Scattering (DIS) are given by the Nuance [15] generator. Figure from Reference [16]	43
3.5	Feynman diagrams of charge current (a) and neutral current (b) quasi-elastic events upon a proton ( $p$ ) or a neutron ( $n$ ). $W = W^+/W^-$ depending on the interaction. The result is a charged lepton, $l$ , which is identified within the detector. . . . .	44
3.6	Feynman diagrams of charge current (a) and neutral current (b) resonant events upon a neutron ( $n$ ). Other permutations upon neutrons and protons are possible. . . . .	45
4.1	The positions of the detectors along the BNB beamline. The near detector SBND is positioned at 110 m from the beam, MicroBooNE is at 470 m and ICARUS is at 600 m. From Reference [17]. . . . .	47
4.2	A schematic of the BNB beam indicating the focusing horn and the decay region. From Reference [18]. . . . .	49
4.3	The simulated flux of the BNB beam at MiniBooNE split into the components of the beam. From Reference [18]. . . . .	49
4.4	The predicted flux from the BNB (a) (from Reference [19] and the NuMI beam (b) (from Reference [20]) to MicroBooNE as a function of neutrino energy. Both beams are in a mode which selects neutrinos over anti-neutrinos.	51
4.5	A schematic of SBND. The CPA lies between the two TPCs providing an electric field. The anode (APA) readout consists of 3 wire planes. From Reference [17]. . . . .	52
4.6	SBND wire readout setup. From Reference [17]. . . . .	52
4.7	A PDS module in SBND. The module is made up of 5 PMTs and 8 ARAPU-CAs which are positioned interchangeably. Mounting brackets on all four sides of the detectors hold the detectors to the frame. Credit: J. Boissevain.	53

4.8	a) Position of the CRT system in SBND. Note that the bottom of the detector is not fully covered due to struts holding the cryostat in position. From Reference [17]. b) Images of the scintillating strips before and after the Mylar has been glued on. From Reference [21]. . . . .	55
4.9	The efficiency of the prototype CRT panel as the distance from the readout increases. This is a result of a laboratory tests at the University of Bern. From Reference [21]. . . . .	55
4.10	Full path of the electronics chain in SBND. From Reference [17]. . . . .	56
4.11	The MicroBooNE sensitivity using Monte Carlo to an excess of electron neutrino events given the sterile oscillation parameters in the x and y axes. The MiniBooNE sensitivity 90% confidence limit is also shown and the LSND 90% confidence limit. Area to the right of the contour has a higher confidence value than the contour value. See Chapter 8 for further detail. From Reference [22]. . . . .	59
4.12	Schematic of the MicroBooNE cryostat and TPC. From Reference [13]. . .	59
4.13	The ICARUS sensitivity to an excess of $\nu_e$ events given the sterile oscillation parameters in the x and y axes. Area to the right of the ICARUS contour has a higher confidence value than the contour value. See Chapter 8 for further detail. From Reference [23]. . . . .	61
4.14	Diagram showing the geometry of the ICARUS TPCs. From Reference [17].	62
4.15	Oscillation probability as a function of length travel for a 0.7 GeV neutrino. Global best-fit parameters for sterile $\nu_e$ appearance excess and $\nu_\mu$ disappearance are used to calculate the probability. SBND (110 m), MicroBooNE (470 m ) and ICARUS (600 m) are indicated on the plot by their distance from the beam target. Note that the hadrons decay between the beam target and up to 50 m in the decay pipe, therefore the neutrino length of flight need not be the full distance to the detectors from the beam target. . . . .	63
4.16	SBN muon neutrino spectra after the $\nu_\mu$ CC inclusive truth-based selection. Note that no cosmic background was considered in the analysis. From Reference [17]. . . . .	65
4.17	The SBN sensitivity to muon neutrino disappearance for the proposal truth-based analysis as well as the sensitivity contour for a joint MiniBooNE and SciBooNE analysis. Area to the right of the contour has a higher confidence value than the contour value. From Reference [17]. . . . .	65
4.18	SBN electron neutrino spectra after the $\nu_e$ CC inclusive truth-based selection. From Reference [17]. . . . .	66

4.19	The sensitivity of the SBN programme to electron neutrino appearance signal for the proposal truth-based analysis. The Figure shows the allowed regions for LSND and some global best fits at the time. Area to the right of the contour has a higher confidence value than the contour value. See Chapter 8 for further detail. From Reference [17]. . . . .	67
5.1	The LArIAT tertiary beamline with the sub-detector and TPC elements. From Reference [24]. . . . .	70
5.2	The LArIAT cosmic muon tagging system. From an internal presentation, credit: J Asaadi. . . . .	71
5.3	The azimuthal (a) and zenith (b) angle distributions arising from events triggered from the cosmic muon tagging system. . . . .	72
5.4	Online monitor cubisms [25] for the noise RMS for the first 8 channels. The current RMS value is shown on the right hand side of the cube in units of ADC. The rest of the cube shows how the noise RMS has changed with time. . . . .	73
5.5	A triggered event whilst the liquid argon was filling during the second run of the VST. The top two images show the charge deposition in wire-time space. The bottom is the waveform on a particular wire. . . . .	74
5.6	Example of the reconstruction process. Hits (black and coloured dots) are identified, and presented in the wire-tick (wire-time) space. Hits are then clustered together based on their topology (coloured dots). An event such as this would be removed during the signal to noise ratio analysis. . . . .	76
5.7	Event display of raw data from the VST showing a through-going cosmic muon. Hits on the waveforms are visible and the 2D cluster in each of the views can be easily performed. There is an offset of 2560 ticks in the readout. . . . .	77
5.8	The output of the raw hit finder for the collection (a) and induction (b) planes with initial linear fits performed. Credit: Tom Brooks . . . . .	78
5.9	Tracks on the collection (a) and induction (b) planes after selection cuts. The cuts were chosen by evaluating the results on a sub-sample of the data. Credit: Tom Brooks . . . . .	79
5.10	Cuts performed on the charge vs time to select hits for the minimization. Credit: Tom Brooks . . . . .	80
5.11	All lifetime measurements (a) and the hour average with the changes to the drift field (b). Credit: Tom Brooks . . . . .	80

5.12 RMS of each wire for sub-run 322 for the induction (a) and collection (b) planes. The error bars are the standard error on the mean. Dead wires, which can be noticed by a sudden drop in RMS in the distribution, were removed from the analysis. Only the longest wires, where the RMS is the largest, were used in the run RMS analysis. Errors in y correspond to the error on the mean value. . . . .	82
5.13 The noise RMS in ADC for several sub-runs in the second run of the VST. The error bars are the standard error on the mean. The drop in the RMS in the induction plane is the result of the remaining FEMB readout wires existing at the end of the induction plane where they are smaller in length. Errors in y correspond to the error on the mean value. . . . .	82
5.14 The normalised peak heights distributions of tracks that pass the cuts for sub-run 322. The fit to the induction (a) and collection (b) distributions is a Landau-Gaussian fit. . . . .	84
5.15 The signal to noise ratio for several sub-runs in the second run of the VST. The error bars are the propagation of the statistical error on the Landau-Gaussian fit and the error on the noise RMS. Note that sub-runs 310-315 were held at the nominal cathode voltage of 23.5 kV, sub-runs 309,317-327 and 332 were held at 20 kV and sub-runs 308 and 329-331 were held at 15 kV. The errors in y are calculated by propagating the statistical errors on the noise RMS and signal value. . . . .	84
5.16 The extrapolated noise RMS as a function of wire length in the collection plane (a) and induction plane (b). . . . .	87
5.17 The signal to noise ratio as a function of drift distance in the SBND detector. The value at a drift of 0 cm is the extrapolated value discussed above. The errors are calculated by propagating the error of the original SNR to the extrapolated value and thus equal to the original error scaled by lifetime and diffusion effects. . . . .	88
5.18 The regression fit in the collection-tick plane (a) and the induction-tick plane (b). . . . .	91
5.19 The noise distribution for the collection plane (a) and induction plane (b). This corresponds to all ADC counts 2000 ticks away from the linear signal fit. . . . .	92
5.20 The noise RMS for each wire the collection plane (a) and induction plane (b). Note: the two higher capacitance wires on the induction plane are not used in the analysis. The errors in y are the standard error on the mean values. . . . .	92
5.21 The FFT for a wire on the collection plane. . . . .	93

5.22	The zenith (a) and azimuthal (b) angular distribution for all events in the run	94
5.23	The exponential plot $Q = Q_0 e^{t/\tau}$ . The errors are the errors on the most probably value that arises during the fitting procedure. The errors are calculated extracted using the MINOS technique [26].	94
5.24	The normalised charge deposition of track like particles in the TPC on the collection plane (a) and induction plane (b). Bins corresponding to less than 42 ADC for the collection plane and 8 ADC for the induction plane are not considered in the fitting to remove noise effects. Errors are Poissonian.	95
5.25	The normalised charge deposition of track-like particles in the TPC multiplied by the Gaussian coefficient on the collection plane. Bins corresponding to less than 26 ADC are not considered in the fitting to remove noise effects. Errors are Poissonian.	96
5.26	A example of an anode to cathode crossing muons in the SBND detector. Credit: Tom Brooks.	98
5.27	a) The fitted Landau-Gaussian convolution to the charge deposited on the wire plane for a specific time slice. The exponential fit of the average time of each time slice against the most probable charge. The error bars are extracted using the MINOS technique [26]. The technique finds the error on the lifetime by taking the difference of the lifetime values for which the $\chi^2$ is +1 from the minimised value. Credit: Tom Brooks	99
5.28	An example of a muon crossing parallel to the wire plane. Credit: Tom Brooks	100
5.29	a) A plot of the calculated lifetime against the simulated lifetime for each module. The error bars are calculated using the MINOS technique [26]. The technique finds the error on the lifetime by taking the difference of the lifetime values for which the $\chi^2$ is +1 from the minimised value. Credit: Tom Brooks. b) A plot of the calculated percentage of charge lost across the full drift distance of the TPC.	100
6.1	The raw (black) and deconvolved (blue) waveforms. A Gaussian (red) has been fitted to peaks using <i>GausHitFinder_module</i> . The x-axis corresponds to the time in ticks ( $0.5 \mu s$ ), and the y-axis corresponds to amplitude in ADC.	106
6.2	A Monte Carlo waveform with corresponding Gaussian fit attempts. The fit with the lowest $\chi^2/\text{Ndof}$ is chosen to model the wavepacket.	108

6.3	The 1st efficiency definition in Equation 6.1.2 for various ADC thresholds set in the <i>GausHit_Module</i> as a function of the azimuthal (XZ) and zenith (YZ) angles. The MC ADC threshold is set to 5 ADC and the <i>dt</i> merging threshold is set to 2. Errors are the standard deviation of the efficiency for the corresponding angle. . . . .	109
6.4	The 2nd efficiency definition in Equation 6.1.2 for various ADC thresholds set in the <i>GausHit_Module</i> as a function of the azimuthal (XZ) and zenith (YZ) angles. The MC ADC threshold is set to 5 ADC and the <i>dt</i> merging threshold is set to 2. Errors are the standard deviation of the efficiency for the corresponding angle. . . . .	109
6.5	The 3rd efficiency definition in Equation 6.1.2 for various ADC thresholds set in the <i>GausHit_Module</i> as a function of the azimuthal (XZ) and zenith (YZ) angles. The MC ADC threshold is set to 5 ADC and the <i>dt</i> merging threshold is set to 2. Errors are the standard deviation of the efficiency for the corresponding angle. . . . .	110
6.6	Example of a waveform pulse train where the energy is deposited along several time ticks. In this case a multi-dimensional Gaussian fit becomes unsuitable . . . . .	110
6.7	The fraction of energy reconstructed in the hits in the muon sample and electron sample. . . . .	111
6.8	The reconstruction chain for the Pandora pattern recognition. From Reference [27]. . . . .	112
6.9	The process of the <i>BlurredClustering_module</i> . The 2D hits are blurred using a Gaussian kernel then a nearest neighbour algorithm is run to merge 2D hits together. . . . .	114
6.10	Event display of a reconstructed electron shower (mostly red) and charged pion and its daughters (green and yellow) originating from the same vertex. The shower is slightly segmented, and therefore small shower clusters skew the distributions of completeness metric. . . . .	115
6.11	The shower hit completeness. The errors on the histograms are Poissonian. The histograms are normalised to the number of reconstructed events for each method individually. The errors in <i>y</i> on the energy mean are the standard error on the mean value for events with true energy within the <i>x</i> error width. The red line corresponds to the ideal value. . . . .	117

---

6.12 The shower hit purity. The errors on the histograms are Poissonian. The histograms are normalised to the number of reconstructed events for each method individually. The errors in $y$ on the energy mean are the standard error on the mean value for events with true energy within the $x$ error width. The red line corresponds to the ideal value. . . . .	118
6.13 The space-point geometry projection matching score for the collection plane. The induction plane results are similar. The errors on the histograms are Poissonian. The histograms are normalised to the number of reconstructed events for each method individually. The red line corresponds to the ideal value. . . . .	119
6.14 The shower segmentation in the events. If shower reconstruction was perfect, the value should always be one. The induction plane results are similar. The errors on the histograms are Poissonian. The histograms are normalised to the number of reconstructed events for each method individually. The errors in $y$ on the energy mean are the standard error on the mean value for events with true energy within the $x$ error width. The red line corresponds to the ideal value. . . . .	120
6.15 Particle identification using the $dE/dx$ vs residual range. From Reference [28].	123
6.16 Particle Identification using the PIDA method. From Reference [29]. . . . .	123
6.17 Pathway for TRACS reconstructions . . . . .	128
6.18 The shower reconstructed metrics for the tuned TRACS tools used in the Analyses in Chapters 7 and 8. The errors on the histograms are Poissonian and the errors in $y$ on the energy mean magnitude are the standard error on the mean value for events with $x$ error width. The histograms are the number of entries. . . . .	134
6.19 The full TRACS reconstruction pathway. The green boxes correspond to calculated shower variables. The blue and red tools have been separately tuned using the Bayesian Optimisation. . . . .	134
6.20 The magnitude of the difference between the reconstructed and true start position of the shower 0-5 cm (a,b). The mean magnitude is also plotted as a function of the true shower energy (c,d). The histograms are normalised in order to consider only the number of events within the range. The errors on the histograms are Poissonian. The error in $y$ on the graphs is the standard error on the beam for the entries with the $x$ error. . . . .	135

- 6.21 The TRACS *Smart Trajectory* tool algorithm. The algorithm iteratively checks the angles (between the black lines) between the trajectory points (in red) and the base point (in green) until both angles are below the set threshold. The resultant direction is shown as the blue line. The dynamic sliding of the base point can be turned off and the shower start position or the initial track stub start position (pink) can be used instead. If the directional information is used, the last trajectory point is not required. . . . . 137
- 6.22 A comparison of the direction calculations of all the existing shower modules. The errors on the histograms are Poissonian. The histograms are normalised in order to only consider the number of events within the range. The errors in  $y$  on the energy mean are the standard error on the mean value for events with true energy within the  $x$  error width. . . . . 138
- 6.23 TRACS track hit finder algorithms. For the *Shower Residual* (a) Blue lines indicate successful PCA fits and red lines indicate failures. Red dots with a blue outline indicate hits removed to force the initial seed to fit. Blue hits indicate points which failed the residual cuts, orange hits are removed due to pruning and cyan stars indicate new seed positions. Red points are identified as the initial track stub hits. For the *2D regression* (b) the blue line is the resultant fit of the red and blue hits. The blue points are removed by either the perpendicular tolerance (cyan lines) or the max number of hits cut (after the pink line). The red line is the next fit in the iteration. For the *3D cylinder* tool (c) a cylinder (red rectangle) is created from the start along the shower direction (blue line). All hits within the cylinder are used. . . . . 139
- 6.24 TRACS reconstructed event display for the various track hit finder and track maker tools. The 3D spacepoints of the reconstructed shower are in black. The red 3D points correspond to the spacepoints which have been identified as being part of the initial track stub. The light blue triangles are the 3D trajectory points made by the track making tools. The dark blue triangle is the start position of the 3D initial track stub. The green square is the reconstructed shower start position and the pink line corresponds to the reconstructed shower direction. Other spacepoints in blue correspond to other reconstructed particles. . . . . 140
- 6.25 ROC Curves for the three track hit finder algorithms for identifying vertex electrons and rejecting vertex photons using a  $dE/dx$  cut. The shower start position must be reconstructed within 1 cm of true start position. . . . . 141

6.26 The $\cos(\theta)$ between the true and reconstructed directions for the various track makers in TRACS using the <i>Smart Trajectory</i> tool. The errors on the histograms are Poissonian. The histograms are normalised by the number of events within the range. . . . .	142
6.27 Diagram of photon decaying into $e^+e^-$ pair. One of the pair stops (red hits) within the initial track stub and therefore a subset of the hits in the initial track have a $dE/dx$ corresponding to an electron landau (outside the blue box). . . . .	143
6.28 The electron and photon priors used in the <i>Bayesian Truncating</i> tool. . . . .	144
6.29 The $dE/dx$ reconstruction for the various TRACS algorithms and existing shower characterisation algorithms. The errors on the histograms are Poissonian. The histograms are normalised by the number of events reconstructed. The ROC curve corresponds to the efficiency of keeping vertex electrons (a) and removing vertex photons (b). . . . .	145
6.30 Linear charge energy calibration from an isotropic sample of simulated muon. . . . .	146
6.31 The shower reconstructed energy for the module currently running in SBND. The errors on the histograms are Poissonian and the errors in $y$ on the energy are the standard error on the mean value for events with true energy within in the $x$ error width. . . . .	147
6.32 The shower reconstructed opening angle for the TRACS tools and the <i>PandoraShower_module</i> . The errors on the histograms are Poissonian and the errors in $y$ are the standard error on the mean value for events with energy within $x$ error width. The histograms are normalised using the number of reconstructed events. . . . .	148
6.33 The shower reconstructed shower length for the TRACS tools and the <i>PandoraShower_module</i> . The errors on the histograms are Poissonian and the errors in $y$ are the standard error on the mean value for events with energy within $x$ error width. The histograms are normalised using the number of reconstructed events. . . . .	149
6.34 The Bayesian Optimisation process. Where the points are added depends on the maximum of the acquisition function. Once the points are added the posterior probability changes with the new information. Figure from Reference [30]. . . . .	151

6.35 The Bayesian Optimisation process over 200 iterations. The <i>3DCylinder</i> tool, see Section 6.2.2, is being tuned to maximise the $dE/dx$ electron photon separation. The red markers represent the Bayesian Optimisations result at that iteration. The black line corresponds to the best result at the iteration. The green line represents the experts best guess based on the track hit length and width discussed in Section 6.2.2. . . . .	153
7.1 The interaction rates of CC0Pi events in GENIE v2_8 and v2_12. The errors on the plot are Poissonian. . . . .	159
7.2 The interaction rates of NC resonant interactions on protons producing a final state neutral pion for GENIE v2_8 and v2_12. The errors on the plot are Poissonian. . . . .	160
7.3 Efficiency of removing events where there are two showers greater than 100 MeV for the $\nu_\mu$ background events in the active volume (red) and the beam dirt background events (orange). This efficiency is in terms of true incoming neutrino energy (a) and the true outgoing charged lepton/neutrino energy (b). The initial distributions are shown for the proposal (solid line) and the modern (dashed line) samples. The efficiency is presented for both samples and the errors are calculated using the normal approximation to $1\sigma$ . . . . .	161
7.4 Diagram of a $\nu$ interaction. The vertex is visible due to hadronic activity (red dots). The black dots refer to a photon shower. Due to the hadronic activity, the conversion gap (red line) is visible. The initial track stub (blue line) is used to calculate the $dE/dx$ of the shower. . . . .	162
7.5 The background rejection from removing events with a visible vertex and conversion gap greater than 3 cm for the $\nu_\mu$ CC (black) and the $\nu_\mu$ NC (red) background events in the active volume. This is displayed as a function of the true incoming neutrino energy (a) and the true outgoing charged lepton/neutrino energy (b). The distributions correspond to the events which pass the multiple shower cut in the proposal (solid line) sample and the modern (dashed line) sample. The errors are calculated using the normal approximation to $1\sigma$ . . . . .	162

7.6	The resultant background rejection due to removing events with a muon with a track length greater than 1 m for the $\nu_\mu$ CC background events in the active volume (black). This is shown in terms of the true incoming neutrino energy (a) and the true outgoing charged lepton energy (b). The distributions of the events after the conversion gap cut and the two shower cut are shown for the proposal (solid line) and the modern (dashed line) samples. The errors are calculated using the normal approximation to $1\sigma$ . . . . .	163
7.7	The background rejection due to removing events where the primary shower has less than 200 MeV of energy for the $\nu_e$ intrinsic CC background events (a,b) (pink), $\nu_\mu$ CC background events (c,d) (black) and $\nu_\mu$ NC background events (e,f) (red) in the active volume. The oscillated $\nu_e$ efficiency (blue) is drawn on each graph for reference. The background rejection is presented in terms of the true neutrino energy (left-hand side) and the true outgoing neutrino/charged lepton energy (right-hand side). The distributions of events after the conversion gap cut, the muon length cut and the two shower cut are shown for the proposal (solid line) and the modern (dashed line) samples. The background rejection is presented for both distributions and the errors are calculated using the normal approximation to $1\sigma$ . . . . .	164
7.8	a) The distribution of photons of energy $>100$ MeV from dirt events within the active volume in terms of the drift direction (x) and the beam direction (z). b) The distance to the closest wall of the active volume bounding box for photons of energy $>100$ MeV that enter the active volume from dirt events.	166
7.9	The total background rejection in the truth-based analysis for removing dirt events with respect to the neutrino energy (a) and outgoing charged lepton/neutrino energy (b). The distributions of events after the two shower cut are shown for the proposal (solid line) and the modern (dashed line) samples. The oscillated $\nu_e$ efficiency (blue) is drawn for reference. The background rejection is presented for both distributions and the errors are calculated using the normal approximation to $1\sigma$ . . . . .	167
7.10	a) The distribution of photons, with energy greater than 100 MeV, from cosmic events within the active volume in terms of the vertical direction (y) and the beam direction (z). b) The distance to the closest wall of the active volume bounding box for photons that enter the active volume from cosmic events. . . . .	168

---

7.11 The background rejection when removing cosmic events with an energy greater than 200 MeV. The errors are calculated using the normal approximation to $1\sigma$ . . . . .	168
7.12 A diagram of the cylinder cut from Reference [17]. . . . .	169
7.13 The total background rejection of the analysis for the intrinsic $\nu_e$ CC (a,b) (pink),The efficiencies are presented in terms of the true neutrino energy (left-hand side) and the true outgoing neutrino/charged lepton energy (right-hand side). The oscillated $\nu_e$ efficiency (blue) is drawn on each graph for reference. The background rejection is presented for both the proposal and modern distributions and the errors are calculated using the normal approximation to $1\sigma$ . . . . .	170
7.14 The resultant spectra from the truth-based analysis in SBND using the "modern sample". The ratio between the "modern" sample and "proposal" sample for the total event distribution and the different interaction types of the analysis is shown. Errors are derived from the Poissonian errors of the histogram. A weighting of $\sin(\theta_{\mu e}) = 0.013$ and $\Delta m_{41}^2 = 0.43 \text{ eV}^2$ has been used on the oscillation sample [31]. . . . .	171
7.15 The resultant spectra from the truth-based analysis in MicroBooNE using the "modern" sample. The ratio between the "modern" sample and "proposal" sample for the total event distribution and the different interaction types of the analysis is shown. Errors are derived from the Poissonian errors of the histogram. A weighting of $\sin(\theta_{\mu e}) = 0.013$ and $\Delta m_{41}^2 = 0.43 \text{ eV}^2$ has been used on the oscillation sample [31]. . . . .	171
7.16 The resultant spectra from the truth-based analysis in ICARUS using the "modern" sample. The ratio between the "modern" sample and "proposal" sample for the total event distribution and the different interaction types of the analysis is shown. Errors are derived from the Poissonian errors of the histogram. A weighting of $\sin(\theta_{\mu e}) = 0.013$ and $\Delta m_{41}^2 = 0.43 \text{ eV}^2$ has been used on the oscillation sample [31]. . . . .	172
7.17 The resultant spectra from the truth-based analysis in SBND using the "proposal" sample. The ratio between the total event distribution and the different interaction types of the analysis against the spectra from 2012 is shown. Errors are derived from the Poissonian errors of the histogram. A weighting of $\sin(\theta_{\mu e}) = 0.013$ and $\Delta m_{41}^2 = 0.43 \text{ eV}^2$ has been used on the oscillation sample [31]. . . . .	172

---

7.18 The resultant spectra from the truth-based analysis in MicroBoone using the "proposal sample". The ratio between the total event distribution and the different interaction types of the analysis against the spectra from 2012 is shown. Errors are derived from the Poissonian errors of the histogram. A weighting of $\sin(\theta_{\mu e}) = 0.013$ and $\Delta m_{41}^2 = 0.43 \text{ eV}^2$ has been used on the oscillation sample [31]. . . . .	173
7.19 The resultant spectra from the truth-based analysis in ICARUS using the "proposal" sample. The ratio between the total event distribution and the different interaction types of the analysis against the spectra from 2012 is shown. Errors are derived from the Poissonian errors of the histogram. A weighting of $\sin(\theta_{\mu e}) = 0.013$ and $\Delta m_{41}^2 = 0.43 \text{ eV}^2$ has been used on the oscillation sample [31]. . . . .	173
7.20 SBN electron neutrino spectra after the $\nu_e$ CC inclusive truth-based selection.	174
7.21 The selection efficiency of CC $\nu_e$ interactions and the background rejection of $\nu_\mu$ interactions in the active volume for removing events where a PFP reconstructed neutrino does not exist in the event drift window. The underlying neutrino distribution is also plotted. The errors on the figures are the $1\sigma$ Clopper-Pearson intervals [32]. . . . .	176
7.22 The selection efficiency of CC $\nu_e$ interactions and the background rejection of $\nu_\mu$ interactions in the active volume for removing events where a reconstructed shower does not exist in the event drift window. The underlying neutrino distribution is also plotted. The errors on the figures are the $1\sigma$ Clopper-Pearson intervals [32]. The results correspond to the sample created after performing the PFP cut. . . . .	177
7.23 The efficiency (blue), background rejection (red), efficiency $\times$ background rejection (black), purity (red) and efficiency $\times$ purity (black) for selecting events requiring a) one shower above the energy threshold and b) no more than one shower above threshold. The error on the curves are the cumulative statistical error. . . . .	178
7.24 The efficiency for reconstructing the neutrino vertex within 5 cm of the true vertex as a function of the true hadronic energy deposited in the TPC. The analysis is performed on the $\nu_\mu$ events after the PFP neutrino and no shower cuts are applied. The underlying true vertex distribution is also plotted, along with the correctly reconstructed distribution. The errors on the figures are the $1\sigma$ Clopper-Pearson intervals [32]. . . . .	179

7.25 The conversion gap distance for signal (blue hatched) and background (red hatched) events. The efficiency (blue), background rejection (red), efficiency $\times$ background rejection (black), purity (red) and efficiency $\times$ purity curves are also plotted. The error on the curves are the cumulative statistical error.	179
7.26 $dE/dx$ for signal (blue hatched) and background (red hatched) for all events (a) and for events where the reconstructed shower energy is greater than 225 MeV (b). The efficiency (blue), background rejection (red), efficiency $\times$ background rejection (black), purity (red) and efficiency $\times$ purity (black) curves are also plotted. The error on the curves are the cumulative statistical error. . . . .	180
7.27 Length of the longest track for signal (blue hatched) and background (red hatched) events. The efficiency (blue), background rejection (red), efficiency $\times$ background rejection (black), purity (red) and efficiency $\times$ purity (black) curves are also plotted. The error of the curves is the cumulative statistical error. . . . .	181
7.28 The energy of the largest shower for signal (blue hatched) and background (red hatched) events. The efficiency (blue), background rejection (red), efficiency $\times$ background rejection (black), purity (red) and efficiency $\times$ purity (black) curves are also plotted. The errors on the curves are the cumulative statistical error. . . . .	182
7.29 The efficiency of selecting CC $\nu_e$ events (blue), the background rejection for removing $\nu_\mu$ events (red) and the purity (black) of the proposal-like selection as a function the true neutrino energy. The underlying distributions before selection for the signal (red) and background (blue) are also plotted, along with the efficiency (dashed orange) and background rejection (dashed light blue) for the POT and oscillation normalised events. The efficiencies are calculated from the remaining events after the pre-selection cuts and the full selection. For the unnormalised events the errors are the $1\sigma$ Clopper-Pearson intervals [32], whilst for the normalised distributions the normal approximation is used. . . . .	183
7.30 The steps in the residual analysis. The distance from the start position of the secondary showers (blue, pink and red dots) is evaluated with respect to the primary shower (black dots). The residual of the secondary shower is calculated using the primary shower direction (pink line). Showers outside the black line (made from the 2 cm line (blue) and the cone outline (red)) are removed. . . . .	184

7.31	The distribution of residual values of secondary showers. . . . .	185
7.32	The PIDA score for the largest shower for signal (blue hatched) and background (red hatched) events. The efficiency (blue), background rejection (red), efficiency $\times$ background rejection (black), purity (red) and efficiency $\times$ purity (black) curves are also plotted. The errors on the curves are the cumulative statistical error. . . . .	186
7.33	The 2D signal (a) and background (b) distributions of the length and the PIDA score of the longest track. . . . .	186
7.34	The selection efficiency of CC $\nu_e$ events (blue), the background rejection for removing $\nu_\mu$ events (red) and the purity (black) of the updated selection as a function of the true neutrino energy. The underlying distributions before selection for the signal (red) and background (blue) are also plotted, along with the efficiency (dashed orange) and background rejection (dashed light blue) for the POT and oscillation normalised events. The efficiencies are calculated from the remaining events after the pre-selection cuts and the full selection. For the unnormalised events, the errors are the $1\sigma$ Clopper-Pearson intervals [32] whilst for the normalised distributions the normal approximation is used. . . . .	187
7.35	An example of the first decision tree from the BDTG algorithm from the root TMVA package. The hyperparameters of the BDT are shown in Table 7.5 and the input parameters are shown in Table 7.6. The signal and background, Figure 7.38, are scaled to 1:1 before the BDT analyses the data. . . . .	191
7.36	The BDT response for the standard BDT for the validation (points) and the training (bars) data for the signal (blue) and background (red). The errors are Poissonian. . . . .	192
7.37	The BDT response for the scaled BDT for the validation (points) and the training (bars) data for the signal (blue) and background (red). The errors are Poissonian. . . . .	193
7.38	The input parameter distribution for the BDTs. The inputs are the remaining events after the PFP neutrino and at least one shower pre-selection cuts. The blue histograms correspond to the $\nu_e$ CC signal and the red to the $\nu_\mu$ background. The distributions go in descending order of the Table 7.6 from the top left Figure to the bottom right Figure. . . . .	195
7.39	Correlation matrix between the input parameters of the BDTs, listed in Table 7.6, for the $\nu_e$ CC signal data. The correlation is represented as a percentage.	196

---

7.40 Correlation matrix between the input parameters of the BDTs, listed in Table 7.6, for the $\nu_\mu$ background data. The correlation is represented as a percentage.	196
7.41 The efficiency (blue), background rejection (red), purity (red), efficiency $\times$ background rejection (black) and efficiency $\times$ purity (black) for selecting events above the BDT score cut off from the standard trained BDT. The errors on the curves are the cumulative statistical error. . . . .	197
7.42 The efficiency (blue), background rejection (red), purity (red), efficiency $\times$ background rejection (black) and efficiency $\times$ purity (black) for selecting events above the BDT score cut off from the scaled trained BDT. The BDT inputs have been normalised using the POT and oscillation weighting. The errors on the curves are the cumulative statistical error. . . . .	198
7.43 The efficiency of selecting CC $\nu_e$ events (blue), the background rejection for removing $\nu_\mu$ events (red) and the purity (black) of the standard BDT selection with a BDT cut off score at 0, a fiducial volume cut and a minimum shower energy cut of 100 MeV, as a function of true neutrino energy. The underlying distributions before selection for the signal (red) and background (blue) are also plotted, along with the efficiency (dashed orange) and background rejection (dashed light blue) for the POT and oscillation normalised events. For the unnormalised events, the errors are the $1\sigma$ Clopper-Pearson [32] intervals whilst for the normalised distributions the normal approximation is used. . . . .	199
7.44 The different selections in efficiency background rejection space for active volume events only. The cut-based selections are optimised by maximising either the efficiency $\times$ background rejection for each cut or the efficiency $\times$ purity. The BDT curves are calculated by varying the BDT score threshold within the selection. The subfigure is the whole curve. The 80% reconstruction efficiency weighting is applied to the truth-based analyses. . . . .	200
7.45 The total selection efficiency as a function of the true neutrino energy. The underlying signal distribution (blue dashed) and background distribution (red dashed) are shown for convenience. The 80% reconstruction efficiency weighting is not applied. The errors on the reconstruction analyses are the $1\sigma$ Clopper-Pearson intervals [32] whereas, for the truth-based analyses the normal approximation is used. . . . .	201

---

7.46 The total background rejection as a function of the true neutrino energy. The underlying signal distribution (blue dashed) and background distribution (red dashed) are shown for convenience. The 80% reconstruction efficiency weighting is not applied. The errors are calculated using the normal approximation. . . . .	202
8.1 The spectra from the three detectors for the truth-based selection. The selection is performed on the modern-like sample described in Chapter 7, Section 7.1.1. The errors are statistical. . . . .	208
8.2 The neutrino energy resolution for Charged Current (CC) $\nu_e$ events after the pre-selection cuts in Chapter 7, Section 7.4.3. The analysis is performed on the Short-Baseline Near Detector (SBND) Booster Neutrino Beam (BNB) events. . . . .	209
8.3 The spectra from the three detectors for the proposal-like selection with the parameters discussed in Chapter 7, Section 7.4.3 set to maximise the efficiency $\times$ background rejection. The spectra are created using the reconstruction information of the Time Projection Chamber (TPC) signal and backgrounds and the truth-based dirt and cosmic background. The errors are statistical. . . . .	209
8.4 The fractional covariance matrices for the flux (a) and interaction (b) for the truth-based modern selection discussed in Chapter 7, Section 7.1.1. The matrices are created from spectra, Figure 8.1. . . . .	212
8.5 The fractional covariance matrices for the flux (a) and interaction (b) for the reconstruction selection discussed in Chapter 7, Section 7.4.3. The matrices are created from spectra, Figure 8.3, which uses the proposal selection, maximising the efficiency $\times$ background rejection. The new binning scheme has been used. . . . .	213
8.6 The $\nu_e$ appearance sensitivity for the truth-based analyses. The modern (red) and proposal-like (blue) samples are shown as well as the modern sample with the new binning scheme (purple). The sensitivities without the systematic penalty are shown via the dashed lines. The 2012 proposal sensitivity (black dashed) is shown for comparison. . . . .	214
8.7 The 90% CL $\nu_e$ appearance sensitivity for the reconstruction analyses. The sensitivity without the systematic penalty is shown via the dashed lines and the systematic errors with the solid line. The 2012 proposal sensitivity (black dashed) and the modern truth-based analysis with the new binning scheme (purple line) is shown for comparison. . . . .	215

8.8	The area covered by the 90% $\nu_e$ appearance sensitivity for the reconstruction analyses as a function of the BDT score used. The cut-based methods are shown as a horizontal line as they do not depend on the BDT score. The $x$ errors for each BDT area are the distance to the next point and the $y$ errors are calculated by propagating the size of the steps used to form the sensitivity region around the contour. The modern truth-based analysis (purple line) with the new binning scheme is shown for comparison. . . . .	216
8.9	The $\nu_e$ appearance sensitivity for the truth-based analyses with the proposal binning (red) and the BDT reconstruction selection with a score of 0.2 (black). The 90%, $3\sigma$ and $5\sigma$ are presented for each analysis. The LSND allowed regions and the MiniBooNE sensitivity is also plotted to show the SBN coverage of the low energy excess. . . . .	217
9.1	The Sheffield liquid argon rig. A) The PMT, B) the protector mesh, C) mesh cathode, D) TPC field cage, E) pixel plane, F) LArASIC chip, G) adapter board, H) argon capacitor level gauge, I) PT-100 RTD resistive temperature sensor, J) threaded struts, K) purifier output, L) purifier input, and M) argon input. . . . .	220
9.2	The Sheffield liquid argon rig. A) The PMT, B) the protector mesh, C) mesh cathode, D) TPC field cage, E) pixel plane, F) LArASIC chip, and G) adapter board. . . . .	221
9.3	The face of the Bern pixel readout. From Reference [33]. . . . .	223
9.4	The response to a test pulse being supplied to the ASIC chip via the Arduino nano. . . . .	223
9.5	The response to a test pulse being supplied to the ASIC chip via a charged wire. . . . .	224
9.6	A flow diagram of the electronics chain in the Sheffield rig. . . . .	224
9.7	The response of the Hamamatsu R11065 PMT to a Cs-137 source when coupled to sodium iodide crystal. A tick is 1 ns in the DAQ. Credit: Ed Tyley for redoing the analysis as the initial data was lost during PC hardware upgrades. . . . .	226
9.8	The time distribution of the peak heights from one of the larger cosmic paddles when triggering with the other. . . . .	226
9.9	Online monitoring diagram showing the average noise RMS on the CAEN DAQ channels every hour. . . . .	228

9.10	The sub-detector event display for an event which triggered the PMT and two large CRTs. The top waveform (in ticks = 1 ns) corresponds to the internal PMT. The next two waveforms are the side CRTs and the bottom two are the small vertical CRTs. Underneath each waveform is the distribution of the waveform in ADC. A Gaussian is fitted to the distribution in the first (left-hand side) and second (right-hand side) iteration of the analysis. The integral of the waveform is taken within the two vertical red lines, whereas the baseline is calculated using the information outside the lines. The green vertical line represents the peak height point and the black horizontal line represents the baseline. . . . .	229
9.11	The mean noise RMS on the pixel channels (a) and the ROI channels (b). The errors are the standard deviation on the RMS calculated for each event.	231
9.12	FFT from software triggered data. The sub-figure is the entire frequency range available. . . . .	231
9.13	a) A test pulse on a pixel (left) and an ROI (right) from the Arduunio. Since the image, alterations to the grounding scheme were undertaken and the oscillatory noise has been removed. b) The same waveforms but with an FFT noise filter applied. . . . .	232
A.1	The 1st efficiency definition in Equation 6.1.2 for various MC $dt$ thresholds as a function of the azimuthal (XZ) and zenith (YZ) angles. The MC ADC threshold is set to 5 ADC. Errors are the standard deviation of the efficiency for the corresponding angle. . . . .	260
A.2	The 2nd efficiency definition in Equation 6.1.2 for various MC $dt$ thresholds as a function of the azimuthal (XZ) and zenith (YZ) angles. The MC ADC threshold is set to 5 ADC. Errors are the standard deviation of the efficiency for the corresponding angle. . . . .	260
A.3	The 3rd efficiency definition in Equation 6.1.2 for various MC $dt$ thresholds as a function of the azimuthal (XZ) and zenith (YZ) angles. The MC ADC threshold is set to 5 ADC. Errors are the standard deviation of the efficiency for the corresponding angle. . . . .	261
A.4	The 1st efficiency definition in Equation 6.1.2 for various MC ADC thresholds as a function of the azimuthal (XZ) and zenith (YZ) angles. The $dt$ merging threshold is set to 2 ticks. Errors are the standard deviation of the efficiency for the corresponding angle. . . . .	261

A.5	The 2nd efficiency definition in Equation 6.1.2 for various MC ADC thresholds as a function of the azimuthal (XZ) and zenith (YZ) angles. The <i>dt</i> merging threshold is set to 2 ticks. Errors are the standard deviation of the efficiency for the corresponding angle. . . . .	262
A.6	The 3rd efficiency definition in Equation 6.1.2 for various MC ADC thresholds as a function of the azimuthal (XZ) and zenith (YZ) angles. The <i>dt</i> merging threshold is set to 2 ticks. Errors are the standard deviation of the efficiency for the corresponding angle. . . . .	262
A.7	The 1st efficiency definition in Equation 6.1.2 for various ADC thresholds set in the <i>GausHit_Module</i> as a function of the azimuthal (XZ) and zenith (YZ) angles. The MC ADC threshold is set to 5 ADC and the <i>dt</i> merging threshold is set to 2. Errors are the standard deviation of the efficiency for the corresponding angle. . . . .	263
A.8	The 2nd efficiency definition in Equation 6.1.2 for various ADC thresholds set in the <i>GausHit_Module</i> as a function of the azimuthal (XZ) and zenith (YZ) angles. The MC ADC threshold is set to 5 ADC and the <i>dt</i> merging threshold is set to 2. Errors are the standard deviation of the efficiency for the corresponding angle. . . . .	263
A.9	The 3rd efficiency definition in Equation 6.1.2 for various ADC thresholds set in the <i>GausHit_Module</i> as a function of the azimuthal (XZ) and zenith (YZ) angles. The MC ADC threshold is set to 5 ADC and the <i>dt</i> merging threshold is set to 2. Errors are the standard deviation of the efficiency for the corresponding angle. . . . .	264
B.1	Structure of the <code>ShowerElementHolder</code> which the user directly interacts with.	266
B.2	Structure of the <code>ShowerElementBase</code> which is held in the <code>ShowerElementHolder</code> and holds a specific data product the user has created. . . . .	267
B.3	The structure of the <code>ShowerProducedPtrHolder</code> which handles the objects to be stored in the <code>art::Event</code> . . . . .	268
C.1	The interaction rates of CC0Pi events in GENIE v2_8 and v2_12. The errors on the plot are Poissonian. . . . .	270
C.2	Figures showing the differences in rates of CC resonant interaction types from GENIE version v2_12 and v2_8. The errors arise from the Poissonian errors on the bin. . . . .	270

---

C.3	Figures showing the differences in rates of NC interactions between GENIE versions v2_12 and v2_8. Note that a neutral pion is not always created in DIS and coherent scattering events. The errors arise from the Poissonian errors on the bin. . . . .	271
C.4	Figures showing the differences in the hadronic energy of NC interactions between GENIE versions v2_12 and v2_8. The errors arise from the Poissonian errors on the bin. . . . .	272
C.5	Figures showing the differences in the conversion distance of the most energetic photon arising from vertex interactions of NC interactions between GENIE versions v2_12 and v2_8. The errors arise from the Poissonian errors on the bin. . . . .	272
C.6	Figures showing the differences in the energy of a final state neutral pion in NC interactions between GENIE versions v2_12 and v2_8. The errors arise from the Poissonian errors on the bin. . . . .	273
C.7	Figures showing the differences in the energy of the most energetic photon arising from the vertex in NC interactions between GENIE versions. The errors the Poissonian. . . . .	273
C.8	Figures showing the differences in the energy of the second most energetic photon arising from the vertex of NC interactions between GENIE versions v2_12 and v2_8. The errors arise from the Poissonian errors on the bin. . . . .	274
D.1	The proton reconstruction efficiency for the different Pandora pattern recognition reconstruction pathways, discussed in Chapter 6, Section 6.1.3, as a function of true energy deposited by the proton. The true deposited energy distribution is plotted (shaded black). The errors on the figures are the $1\sigma$ Clopper-Pearson intervals [32]. . . . .	276
D.2	The mean shower hit completeness for the different Pandora pattern recognition reconstruction pathways, discussed in Chapter 6, Section 6.1.3, as a function of true shower energy. The errors in $y$ are the standard error on the mean for values within the true energy range $x - x_{err} \rightarrow x + x_{err}$ where $x_{err}$ corresponds to the error bar bar on the $x$ axis. The pink line corresponds to the ideal value. . . . .	276

E.1	The $\nu_e$ efficiency and $\nu_\mu$ background rejection for the CC interactions (a) and NC interactions (c) when removing events with no reconstructed PFP neutrino. b) The CC interactions efficiency and background rejection as a function of the outgoing charged lepton energy. The underlying neutrino distribution is also plotted. The errors on the figures are the $1\sigma$ Clopper-Pearson intervals [32]. . . . .	277
E.2	The total oscillated $\nu_e$ efficiency and $\nu_\mu$ background rejection when removing events with no PFP reconstructed neutrino. The efficiency is split between the different interaction modes, a) Elastic , b) Resonant and c) DIS, with the underlying initial neutrino energy distribution also plotted. The errors on the Figures are the $1\sigma$ Clopper-Pearson intervals [32]. . . . .	278
E.3	a) The total oscillated $\nu_e$ efficiency and $\nu_\mu$ background rejection when removing events with no reconstructed shower after the PFP neutrino cut. The efficiency is split between the different interaction modes a) QE, b) Resonant and c) DIS, with the underlying initial neutrino energy distribution also plotted. The errors on the figures are the $1\sigma$ Clopper-Pearson intervals [32]. . . . .	279
E.4	The $\nu_e$ efficiency and $\nu_\mu$ background rejection for the CC interactions and NC interactions for removing events with no reconstructed shower as a function of the true incoming neutrino energy (a) and as a function of the outgoing charged lepton energy (b). The underlying neutrino distribution is also plotted. The errors on the figures are the $1\sigma$ Clopper-Pearson intervals [32]. . . . .	279
F.1	The energy reconstruction correctness for tracks (a), showers (b) and neutrinos (c). The neutrino reconstructed energy is calculated after corrections have been applied to the track and shower for only events above the reconstructed energy of 225 MeV. The $\nu_\mu$ sample is used for the tracks and the $\nu_e$ sample for the showers. All samples used underwent the no PFP neutrino and no reconstructed shower cuts. Errors on the graphs are statistical. . . .	282

---

F.2 Neutrino energy correctness of CC $\nu_e$ signal (blue) events and the $\nu_\mu$ (red) background events as a function of the neutrino interaction variables: a) the true neutrino energy, b) the outgoing charged lepton/neutrino energy, c) the reconstructed neutrino energy, d) the Bjorken parameter $x$ , e) the inelasticity $y$ , f) the momentum transfer $Q^2$ , g) the angle between the outgoing charged lepton/neutrino and the incoming neutrino, h) the incoming transverse momentum, and i) the hadronic invariant mass. The errors on the y-axis are the errors on the mean from the distribution of values between $x - \delta x \rightarrow x + \delta x$ . All samples used underwent the no PFP neutrino and no reconstructed shower cuts and only the track and shower energies have been corrected for. A 225 MeV reconstructed energy cut has been applied to the Figures to remove poor reconstruction. . . . .	284
G.1 The truth-based $dE/dx$ distributions from electrons (a) and photons (b) from the vertex simulated sample as well as electrons (c) and photons (d) from the BNB sample. The photon samples are stacked and split into the different end processes of the photon, Compton scattering (blue) and pair production (red). Credit: Ed Tyley. . . . .	286
H.1 The efficiency of selecting CC $\nu_e$ events (blue), the background rejection for removing $\nu_\mu$ events (red) and the purity (black) of the proposal-like selection as a function of the neutrino interaction variables: a) the outgoing charged lepton/neutrino energy, b) the reconstructed neutrino energy, c) the Bjorken parameter $x$ , d) the inelasticity $y$ , e) the momentum transfer $q^2$ , f) the angle between the outgoing charged lepton/neutrino and the incoming neutrino, g) the incoming transverse momentum, and h) the hadronic invariant mass. The underlying distributions before selection for the signal (red) and background (blue) are also plotted, along with the efficiency (dashed orange) and background rejection (dashed light blue) for the POT and oscillation normalised events. The efficiencies are calculated from the remaining events after the pre-selection cuts and the full selection. For the unnormalised events the errors are the $1\sigma$ Clopper-Pearson intervals [32], whilst for the normalised distributions the normal approximation is used. . . . .	289

H.2 The selection efficiency of CC $\nu_e$ events (blue), the background rejection for removing $\nu_\mu$ events (red) and the purity (black) of the updated selection as a function of the neutrino interaction variables: a) the outgoing charged lepton/neutrino energy, b) the reconstructed neutrino energy, c) the Bjorken parameter $x$ , d) the inelasticity $y$ , e) the momentum transfer $q^2$ , f) the angle between the outgoing charged lepton/neutrino and the incoming neutrino, g) the incoming transverse momentum, and h) the hadronic invariant mass. The underlying distributions before selection for the signal (red) and background (blue) are also plotted, along with the efficiency (dashed orange) and background rejection (dashed light blue) for the POT and oscillation normalised events. The efficiencies are calculated from the remaining events after the pre-selection cuts and the full selection. For the unnormalised events, the errors are the $1\sigma$ Clopper-Pearson intervals [32] whilst for the normalised distributions the normal approximation is used. . . . .	291
H.3 The efficiency of selecting CC $\nu_e$ events (blue), the background rejection for removing $\nu_\mu$ events (red) and the purity (black) of the standard BDT selection with a BDT cut off score at 0, a fiducial volume cut and a minimum shower energy cut of 100 MeV, as a function of the neutrino interaction variables: a) the outgoing charged lepton/neutrino energy, b) the reconstructed neutrino energy, c) the Bjorken parameter $x$ , d) the inelasticity $y$ , e) the momentum transfer $q^2$ , f) the angle between the outgoing charged lepton/neutrino and the incoming neutrino, g) the incoming transverse momentum, and h) the hadronic invariant mass. The underlying distributions before selection for the signal (red) and background (blue) are also plotted, along with the efficiency (dashed orange) and background rejection (dashed light blue) for the POT and oscillation normalised events. For the unnormalised events, the errors are the $1\sigma$ Clopper-Pearson intervals [32] whilst for the normalised distributions the normal approximation is used. . . . .	293
I.1 The 2D distributions of the efficiency (a), background rejection (b), purity (c), efficiency $\times$ background rejection (d) and efficiency $\times$ purity (e) for the normalised signal and background events when removing the events with a secondary shower, defined by residual analysis. The secondary shower is only considered for the analysis if the shower has a greater energy than the energy cut and the residual of the shower obeys the equality, Equation I.1, for the residual fraction in the plot. . . . .	297

---

I.2	The 2D signal (a) and background (b) distributions of the length and the PIDA score of the longest track. Also shown are the 2D efficiency (b), background rejection (c), purity (d), efficiency $\times$ background rejection, (e) and efficiency $\times$ purity (f) for the normalised signal and background events when removing events where the length is greater than the length cut and the PIDA score is less than the PIDA cut off. . . . .	298
I.3	The largest shower length for electrons (blue) and photons (red) for all events where the reconstructed shower energy is greater than 200 MeV for the vertex sample. This is shown as a total distribution (a) and as a function of the true energy (b). The errors on the distribution are the standard deviation of the events between energies $x - \delta x \rightarrow x + \delta x$ . . . . .	299
I.4	Largest shower length for signal (blue hatched) and background (red hatched) events. The efficiency (blue), background rejection (red), efficiency $\times$ background rejection (black), purity (red) and efficiency $\times$ purity (black) curves are also plotted. The errors on the curves are the cumulative statistical error. . . . .	300
I.5	The largest shower opening angle for electrons (blue) and photons (red) for all events where the reconstructed shower energy is greater than 200 MeV for the vertex sample. This is shown as a total distribution (a) and as a function of the true energy (b). The errors on the distribution are the standard deviation of the events between energies $x - \delta x \rightarrow x + \delta x$ . . . . .	301
I.6	The largest shower opening angle for signal (blue hatched) and background (red hatched) events. The efficiency (blue), background rejection (red), efficiency $\times$ background rejection (black), purity (red) and efficiency $\times$ purity (black) curves are also plotted. The errors on the curves are the cumulative statistical error. . . . .	301
I.7	The average relative energy density of the shower segments as a function of the distance from the start position of the largest shower in the vertex sample events. The points represent the average value for one of the ten segments analysed. The errors correspond to the error on the mean for the values within the segment. The blue points are from the electron sample and the blue line is the best fit for the electrons. The red points are from the photon sample and the red line is the best fit. A $\chi^2$ minimisation is performed to do the fit. . . . .	303

I.8	The relative fit energy density gradient fit parameter (a) and the power fit parameter (b) for the largest shower for the signal (blue hatched) and background (red hatched) events. The efficiency (blue), background rejection (red), efficiency $\times$ background rejection (black), purity (red) and efficiency $\times$ purity (black) curves are also plotted. The errors of the curves are the cumulative statistical error. . . . .	303
I.9	The track stub length of the largest shower for electrons (blue) and photons (red) for the vertex sample events where the reconstructed shower energy is greater than 200 MeV. The total distribution is shown in (a) and as a function of the true energy in (b). The errors on the distribution are the standard deviation of the events between energies $x - \delta x \rightarrow x + \delta x$ . . . . .	304
I.10	The track stub width of the largest shower for electrons (blue) and photons (red) for events where the reconstructed shower energy is greater than 200 MeV. a) Shows the total distribution and b) shows the events as a function of the true energy for the vertex sample. The errors on the distribution are the standard deviation of the events between energies $x - \delta x \rightarrow x + \delta x$ . . . . .	305
I.11	The track stub length (a) and width (b) for largest shower in the events for BNB oscillated CC $\nu_e$ signal (blue hatched) and BNB $\nu_\mu$ background (red hatched). The efficiency (blue), background rejection (red), efficiency $\times$ background rejection (black), purity (red) and efficiency $\times$ purity (black) curves are also plotted. The errors on the curves are the cumulative statistical error. . . . .	305
I.12	The number of neutrinos in the event for BNB oscillated CC $\nu_e$ signal (blue hatched) and BNB $\nu_\mu$ background (red hatched) events. The efficiency (blue), background rejection (red), efficiency $\times$ background rejection (black), purity (red) and efficiency $\times$ purity (black) curves are also plotted. The errors on the curves are the cumulative statistical error. . . . .	306
J.1	The spectra from the three detectors for the truth based selection. The selection is performed on the proposal-like sample described in Chapter 7, Section 7.1.2. . . . .	307
J.2	The spectra from the three detectors for the truth-based selection. The selection is performed on the modern-like sample described in Chapter 7, Section 7.1.1. The errors are statistical. . . . .	307
J.3	The spectra from the three detectors for the truth-based selection with the new binning. The selection is performed on the modern-like sample described in Chapter 7, Section 7.1.1. The errors are statistical. . . . .	308

---

J.4	The spectra from the three detectors for the proposal-like selection with the parameters discussed in Chapter 7, Section 7.4.3 set to maximise the efficiency $\times$ background rejection. The spectra are created using the reconstruction information of the Time Projection Chamber (TPC) signal and backgrounds and the truth-based dirt and cosmic background. The errors are statistical. . . . .	308
J.5	The spectra from the three detectors for the proposal-like selection with the parameters discussed in Chapter 7, Section 7.4.3 set to maximise the efficiency $\times$ purity. The spectra are created using the reconstruction information of the TPC signal and backgrounds and the truth-based dirt and cosmic background. The errors are statistical. . . . .	308
J.6	The spectra from the three detectors for the proposal-like selection with the parameters discussed in Chapter 7, Section 7.4.3 set to the proposal cut values. The spectra are created using the reconstruction information of the TPC signal and backgrounds and the truth-based dirt and cosmic background. The errors are statistical. . . . .	309
J.7	The spectra from the three detectors for the new selection with the parameters discussed in Chapter 7, Section 7.4.4 set to maximise the efficiency $\times$ background rejection. The spectra are created using the reconstruction information of the TPC signal and backgrounds and the truth-based dirt and cosmic background. The errors are statistical. . . . .	309
J.8	The spectra from the three detectors for the new selection with the parameters discussed in Chapter 7, Section 7.4.4 set to maximise the efficiency $\times$ purity. The spectra are created using the reconstruction information of the TPC signal and backgrounds and the truth-based dirt and cosmic background. The errors are statistical. . . . .	309
J.9	The spectra from the three detectors for the Boosted Decision Tree (BDT) selection with the parameters discussed in Chapter 7, Section 7.4.5 for a BDT score of 0.2. The BDT where the signal and background have been normalised to 1:1 is used. The spectra are created using the reconstruction information of the TPC signal and backgrounds and the truth-based dirt and cosmic background. The errors are statistical. . . . .	310

J.10	The spectra from the three detectors for the BDT selection with the parameters discussed in Chapter 7, Section 7.4.5 for a BDT score of -0.99. This is the BDT where the signal and background have been normalised with the Protons On Target (POT) and oscillation parameters described in Chapter 7. The spectra are created using the reconstruction information of the TPC signal and backgrounds and the truth-based dirt and cosmic background. The errors are statistical. . . . .	310
K.1	The fractional covariance matrices for the flux (a) and interaction (b) for the truth-based proposal selection, described in Chapter 7, Section 7.1.2. The matrices are created from spectra, Figure J.1. . . . .	312
K.2	The fractional covariance matrices for the flux (a) and interaction (b) for the truth-based modern selection discussed in Chapter 7, Section 7.1.1. The matrices are created from spectra, Figure J.2. . . . .	312
K.3	The fractional covariance matrices for the flux (a) and interaction (b) for the truth-based modern selection discussed in Chapter 7, Section 7.1.1. The matrices are created from spectra, Figure J.3. The new binning scheme has been used. . . . .	313
K.4	The fractional covariance matrices for the flux (a) and interaction (b) for the reconstruction selection discussed in Chapter 7, Section 7.4.3. The matrices are created from spectra, Figure J.4, which uses the proposal selection, maximising the efficiency $\times$ background rejection. The new binning scheme has been used. . . . .	313
K.5	The fractional covariance matrices for the flux (a) and interaction (b) for the reconstruction selection discussed in Chapter 7, Section 7.4.3. The matrices are created from spectra, Figure J.5, which uses the proposal selection, maximising the efficiency $\times$ purity. The new binning scheme has been used.	314
K.6	The fractional covariance matrices for the flux (a) and interaction (b) for the reconstruction selection discussed in Chapter 7, Section 7.4.3. The matrices are created from spectra, Figure J.6, which uses the proposal selection using the proposal cut values. The new binning scheme has been used. . . . .	314
K.7	The fractional covariance matrices for the flux (a) and interaction (b) for the reconstruction selection discussed in Chapter 7, Section 7.4.4. The matrices are created from spectra, Figure J.7, which uses the new selection, maximising the efficiency $\times$ background rejection. The new binning scheme has been used. . . . .	315

K.8 The fractional covariance matrices for the flux (a) and interaction (b) for the reconstruction selection discussed in Chapter 7, Section 7.4.4. The matrices are created from spectra, Figure J.8, which uses the new selection, maximising the efficiency $\times$ purity. The new binning scheme has been used.	315
K.9 The fractional covariance matrices for the flux (a) and interaction (b) for the reconstruction selection discussed in Chapter 7, Section 7.4.5. The matrices are created from spectra, Figure J.9, which uses the BDT selection. The BDT where the signal and background have been normalised to 1:1 is used with a BDT score threshold of 0.2. The new binning scheme has been used. . . . .	316
K.10 The fractional covariance matrices for the flux (a) and interaction (b) for the reconstruction selection discussed in Chapter 7, Section 7.4.5. The matrices are created from spectra, Figure J.10 which uses the BDT selection. The BDT where the signal and background have been normalised with the POT and oscillation parameters described in Chapter 7 is used. A BDT score threshold of -0.99 was used. The new binning scheme has been used. . . . .	316



# List of tables

2.1	Best fit oscillation parameters of the PMNS matrix, assuming the normal and inverted hierarchy. Errors are the $1\sigma$ confidence intervals. From Reference [34] . . . . .	12
4.1	Hadrons created from the BNB by impacting protons on the target and their branching ratios to neutrinos. From Reference [18]. . . . .	48
4.2	The composition of the BNB. From Reference [18]. . . . .	49
4.3	Variations in the total flux for each neutrino species. From Reference [18]. .	50
5.1	Cut values used to select straight through-going cosmic muon tracks. Credit: Tom Brooks . . . . .	79
5.2	Selection cuts used in the lifetime selections. . . . .	98
6.1	The quality cuts applied to the simulated data before analysis the success of the reconstruction . . . . .	115
6.2	The list of tools available in the TRACS software. The property that the tool calculates is given ( SP="Start Position", SD="Shower Direction", ITH="Initial Track Hits", IT="Initial Track", SE="Shower Energy", SL="Shower Length" and SOA="Shower Opening Angle"). A description of the tool is also given along with a reference to a figure of merit for the tool. In addition, the origin of the tool methods within shower reconstruction is given. . . . .	132
7.1	The number of protons on target for each sample generated. . . . .	158
7.2	Fiducial volume cut values showing the distance from the active volume faces to the fiducial volume. The cut is applied to all three detectors in the SBN programme. A 1.5 cm cut is applied either side of cathode so that interactions in the cathode are not considered. . . . .	165

7.3	Selection cuts in the truth-based analysis and the percentage of events which pass the cuts for the main beam background modes and the oscillated $\nu_e$ signal in SBND. The proposal sample (Prop) and the modern sample (Mod) efficiencies are presented. The cuts are placed sequentially with the top cut of the Table being applied first. The efficiency of the next cut is then based on the remaining events which passed the previous cut. Events where there is no shower above 100 MeV do not pass the initial cut. Errors are calculated using the normal approximation to $1\sigma$ . . . . .	165
7.4	The cosmic background rejection for each of the cosmic cuts in SBND. The errors on the figures are the $1\sigma$ Clopper-Pearson intervals [32]. . . . .	169
7.5	The BDT hyperparameters for the tuned BDTs. . . . .	193
7.6	Selection parameters used in the analyses. The fraction of times the parameter is used in the standard BDT is shown (similar for the scaled BDT). The efficiency, background rejection and purity for the best cut with respect to maximising the efficiency $\times$ background rejection and efficiency $\times$ purity for each variable independently is shown. The Figures associated to the cuts are presented. . . . .	194
8.1	The interaction systemic errors used in the analysis and the $1\sigma$ fractional uncertainties used when creating the "Universes". Below the line, the event reweighting GENIE package is not used. . . . .	211
8.2	The flux systemic errors used in the analysis and the $1\sigma$ fractional uncertainties used when creating the "Universes". . . . .	212

# Nomenclature

The next list describes several symbols that will be later used within the body of the document

## Acronyms

ADC Analog Digital Converter

APA Anode Plane Assembly

API Application Programming Interface

ARAPUCA Argon R&D Advanced Program at UniCAMP (light trap detector)

ArgonNeuT Argon Neutrino Test Stand

Art analysis reconstruction framework

ASIC Application-Specific Integrated Circuit

BAO Baryon Acoustic Oscillation

BDT Boosted Decision Tree

BDT Support Vector Machine

BNB Booster Neutrino Beam

BNL Brookhaven National Laboratory Experiment

Borexino Italian diminutive of BOREX (BORon solar neutrino EXperiment)

CC Charge Current interaction

CMB Cosmic Microwave Background

CMOS Complementary Metal-Oxide-Semiconductor

CNN Convolutional Neural Network

CORSIKA COsmic Ray SImulations for KAscade

COTS Commercial-Off-The-Shelf

CP Charge Parity

CPA Cathode Plane Assembly

CPT Charge Parity Time

CPU Central Processing Unit

CRT Cosmic Ray Tagging system

CV Central Value

DAQ Data AcQuisition systems

DIS Deep Inelastic Scattering interaction

DUNE Deep Underground Neutrino Experiment

EPICS Experiment Physical and Industrial Control System

FEB Front end Electronics Board

FEMB Front End Mother Board

FFT Fast Fourier Transform

FPGA Field-Programmable Gate Array

FSI Final State Interaction

FTS File Transfer System

GALLEX Gallium Experiment

GENIE Generates Events for Neutrino Interaction Experiments

Hyper-K Hyper-Kamiokande

ICARUS Imaging Cosmic And Rare Underground Signals experiment

J-PARC Japan Proton Accelerator Research Complex

JSNS2	J-PARC Spallation Neutron Source (JSNS2) experiment
JUNO	Jiangmen Underground Neutrino Observatory
K2K	KEK to Kamioka experiment
KamLAND	Kamioka Liquid Scintillator Antineutrino Detector
KARMEN	KArlsruhe Rutherford Medium Energy Neutrino experiment
KATRIN	KArlcruhe TRItium Neutrino experiment
KS	Kolmogorov–Smirnov score. A test to compare the similarities between data
LANSCE	Los Alamos Neutron Science Center
LArIAT	Liquid Argon In A Testbeam
LArSoft	Liquid Argon Software
LArTPC	Liquid Argon Time Projection Chamber
LBL	Long Base-Line
LNGS	Laboratori Nazionali del Gran Sasso
LSND	Liquid Scintillator Neutrino Detector
MC	Monte Carlo
MEC	Meson Exchange Current interaction
MicroBooNE	Micro Booster Neutrino Experiment
MiniBooNE	Mini Booster Neutrino Experiment
MINOS	Main Injector Neutrino Oscillation Search (a + is added for the MINOS upgraded detector)
MIP	Minimum Ionizing Particle
MSW	Mikheyev–Smirnov–Wolfenstein effect
MVA	Multivariate Analysis
NC	Neutral Current interaction

ndof	Number of Degrees of Freedom
NIST	National Institute of Standards and Technology
Nova	NuMI Off-axis $v_e$ Appearance experiment
NOMAD	Neutrino Oscillation MAgnetic Detector
NU	National Instruments
NuMI	Neutrinos at the Main Injector beam
OPERA	Oscillation Project with Emulsion-tRacking Apparatus
ORCA	Oscillation Research with Cosmics in the Abyss experiment
PCA	Principle Component Analysis
PDS	Photon Detection System
PFParticle	Particle Flow Particle
PID	Particle Identification
PIDA	Particle Identification value A
PINGU	Precision IceCube Next Generation Upgrade
PMA	Projection Matching Algorithm
PMNS	Pontecorvo-Maki-Nakagawa-Sakata
PMT	Photo-Multipler Tube
POT	Protons On Target
QE	Quasi-Elastic interaction
RBF	Radial Basis Function
RENO	Reactor Experiment for Neutrino Oscillation
ROI	Region of Interest
SAGE	Soviet American Gallium Experiment
SAM	Sequential Access via Metadata

SBN	Short Baseline Neutrino programme
SBND	The Short Baseline Near Detector
SerDes	Serializer/Deserializer
SiPM	Silicon PhotonMultipler
SM	Standard Model
SNO	Sudbury Neutrino Observatory
SNR	Signal to Noise Ratio
SPT K11	South Pole Telescope,Keisler et al, 2011
SQL	Structured Query Language
Super-K	Super-Kamiokande
T2K	Tokai to Kamioka experiment
TMVA	Toolkit for Multivariate Data Analysis
TPB	Tetra-Phenyl-Butadiene
TPC	Time Projection Chamber
TRACS	Tool-based Reconstruction Algorithm for Characterising Showers
UCB	Upper Confidence Bounds
V-A	Vector-Axial Theory
VALOR	VALencia-Oxford-Rutherford neutrino oscillation fitting framework
VST	Vertical Slice Test
WIB	Warm-Interface Board
WMAP7	The Wilkinson Microwave Anisotropy Probe 7-year result

### **Reconstruction Nomenclature**

*cluster* A 2D reconstructed collection of hits in wire tick space.

*hit* Reconstructed charge deposition identified on the waveform.

*Pandora Pattern Recognition* Software which takes *hits* and clusters them into 2D *clusters* and 3D clusters *PFParticle*.

*shower* A collection of particles which arises from a cascade in the argon. These are commonly electromagnetic showers  $e^{+-}/\gamma$  In reconstruction is a *PFParticle* with the topological characteristics of a track.

*track* Particle which ionises the argon in straight line. In reconstruction is a *PFParticle* with the topological characteristics of a track.

## Units

ADC The voltage step size in a Analog Digital Converter

ENC Electron Noise Count, i.e.  $e^-$

tick The unit of time measurement corresponding to frequency of the DAQ system

# Chapter 1

## Introduction

Over the past century, neutrino physics has been developed and incorporated into the Standard Model (SM) of particle physics. The SM defines neutrinos as massless and left-handed; however, several experiments have found evidence for neutrino oscillation. This evidence has confirmed neutrino oscillation theory which also requires neutrinos to have mass. More recently, there have been measurements of neutrino rates which contradict this neutrino oscillation theory. To account for these discrepancies, one or more extra neutrinos which do not interact with matter have been proposed: sterile neutrinos. Additionally, if a right-handed sterile neutrino does exist, this could explain why neutrino masses are lower than expected. Chapter 2 will introduce the theory of neutrinos within the standard model as well as neutrino oscillations and the experimental evidence for them. The Chapter will then discuss the current status of the neutrino masses, sterile neutrino theory and the observations of anomalies which could be accounted for by sterile neutrinos.

Concerning experimental evidence for the existence of sterile neutrinos, anomalous results in two experiments, the Liquid Scintillator Neutrino Detector (LSND) and the Mini Booster Neutrino Experiment (MiniBooNE) are of particular interest. The Short Baseline Neutrino (SBN) programme has been designed to confirm or refute these anomalies. The SBN programme employs the use of three Liquid Argon Time Projection Chambers (LArTPCs): The Short Baseline Near Detector (SBND), Micro Booster Neutrino Experiment (MicroBooNE), and the Imaging Cosmic And Rare Underground Signals (ICARUS) detector. The LArTPC design and physics response will be discussed generally in Chapter 3 and a particular focus on the SBN programme design will be discussed in Chapter 4. Initial testing of the electronic chain in SBND is discussed in Chapter 5. A signal to noise measurements of the electronics, which is key to understanding the capabilities of the programme, are discussed in detail within this Chapter.

The SBN detectors are placed in the Booster Neutrino Beam (BNB) beamline at 110 m (SBND), 470 m (MicroBooNE) and 600 m (ICARUS) such that high precision measurements of the muon neutrino flux in the near detector can constrain the flux in the far detectors. Due to the proximity to the BNB target, the expected number of neutrino events for the programme is in the millions. Therefore, robust automated software is required to reconstruct, simulate and select neutrinos. The software exists in the LArSoft framework, which is shared across other LArTPC experiments. The particular reconstruction used in SBND is discussed in Chapter 6. The reconstruction is similar to the other detectors in the programme. The SBN programme will be particularly effective due to the powerful calorimetric and topological reconstruction of LArTPCs.

One of the primary goals of the SBN programme is to perform an electron neutrino appearances sensitivity analysis in the muon BNB beam. In order to measure the number of Charge Current (CC) electron neutrinos in the detectors, it is essential to identify the electron leptons arising from the neutrino interactions. Electrons form electromagnetic showers in LArTPCs and can be identified from the shower-like topology and their calorimetric information. Therefore, a new framework to characterise the properties of the electromagnetic showers was developed that allows for various methods to be easily combined. This, in turn, maximises the efficiency of shower reconstruction. The new framework and methods are discussed in Chapter 6.

The main background for electron neutrino CC events are events where a photon is misidentified as an electron shower. Hence, a CC electron neutrino selection was developed on SBND simulated data with a particular focus on removing the photon background. Several selections were performed and compared in Chapter 7. The selections were then performed on simulated data in the SBN detectors and the resultant event distributions were used to perform an electron neutrino appearance sensitivity analysis in the VALEncia-Oxford-Rutherford (VALOR) fitting framework. Details of the sensitivity analysis can be found in Chapter 8.

An additional aim of the SBND detector was to provide a single-phase wire TPC prototype for the Deep Underground Neutrino Experiment (DUNE) far detector. However, due to a high pile-up rate at the DUNE near detector, the wire plane readout employed in SBND is not feasible. It is not feasible because, at a high pile-up rate, ambiguities will be produced in the readout, which cannot be deconvolved. Hence, new readout methods are being developed. Of particular interest is a type of 3D pixel readout method. At the University of Sheffield, a liquid argon rig has been developed in order to test new electronics and readouts for future liquid argon experiments. The design, electronics and software will be discussed in Chapter 9.

## Chapter 2

# The Case for the Sterile Neutrino: An Overview of Neutrino Physics

The Standard Model (SM) describes neutrinos as weakly interacting, neutral particles with no mass. However, it was postulated that neutrinos oscillate between different flavours during propagation [35] [36]. The theory of neutrino oscillations requires that at least two neutrinos have an absolute mass. This contradicts the SM. Several experimental investigations of neutrino oscillation have been carried [37, 3, 4, 4, 38] out in order to pin-point the parameters which describe oscillation theory which are described in below. A subset of these experiments [39, 10, 40–42] have observed anomalous results which may indicate one or more non-interacting neutrinos [12], known as sterile neutrinos [35].

To better understand why there are currently ongoing searches for eV sterile neutrinos, this Chapter will briefly explain, with historical context, the changes in neutrino theory from postulation up to the SM in Section 2.1. Going beyond the SM, theory and observations of neutrino oscillations are then discussed in Section 2.2.1. In order for the neutrino flavour eigenstates to oscillate neutrinos require mass. Therefore, searches for neutrino mass and a theoretical mechanism for the eV neutrino masses, known as the seesaw mechanism, will then be discussed in Sections 2.2.2 and 2.2.3 respectively. The seesaw mechanism is of particular interest as it gives a theoretical case for the sterile neutrino. One requirement of this seesaw mechanism is that neutrinos are Majorana particles [43]. Traditionally the added right-handed neutrino in the seesaw mechanism has a mass of  $\sim 10^6$  as to explain the smallness of the neutrino mass with regards to their lepton partners; however, seesaw theories do exist to allow for eV scale sterile neutrinos [44].

This Chapter also discusses experimental and cosmological limits on neutrino masses, in Section 2.2.2 and the number of neutrinos in Section 2.2.4. The sections discuss how the mass can be used as a probe the Majorana nature of neutrinos whilst the number of

neutrinos is important for understanding the degrees of freedom in oscillation theory. A theoretical review of sterile neutrinos is presented in Section 2.2.5 to explain how modern beam experiments identify sterile signals. Finally, the experiments that observed the neutrino oscillation anomalies are outlined, and the current experimental searches for sterile neutrinos are discussed in Section 2.2.6.

## 2.1 Historical Overview of Neutrino Physics Theory

The existence of the neutrino was first proposed by Pauli in the 1930s [45] to account for discrepancies in the energy spectrum of the electrons in beta decay experiments which started with Meitner and Hahn in 1911. Pauli suggested that a light, neutral particle was also emitted during beta decay which would account for the missing energy and allow the spectrum of the  $e^-$  to be continuous rather than discrete. In 1934, Fermi developed the theory of the weak interaction [46]. Fermi's theory described a point-like transition interaction called the Fermi Transition, see Figure 2.1. The theory was parity conserving and introduced a constant  $G_F$  to describe the strength of the interaction. The Fermi interaction is described by the following interaction Hamiltonian [47]

$$\mathcal{H}_I = G_F \sum_{i=1}^4 (\psi_p^* \gamma_i \psi_n) (\psi_e^* \gamma_i \psi_v) + h.c., \quad (2.1)$$

where  $\psi_{e/n/p/v}$  are the Dirac wave functions.

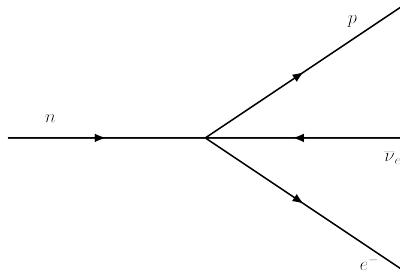


Fig. 2.1 Feynman diagram for the Fermi transition for beta decay.

In 1936, the Hamiltonian in Equation 2.1 was generalised by Gamow and Teller to account for terms including the products of scalars (S), vector (V), tensor (T), pseudovector (A) and pseudoscalar (P) terms,

$$\mathcal{H}_I = \sum_{i=S,V,T,A,P} G_i \psi_p^* O_i \psi_n (\psi_e^* O_i \psi_v) + h.c., \quad (2.2)$$

where  $O_i$  are the unitary Pauli and gamma matrices.  $G_i$  are the corresponding coupling constants [47].

Cowan and Reines first detected neutrinos in 1953 [48] [49]. The experiment used to achieve this detected anti-neutrinos arising from nuclear reactors. Electorn anti-neutrinos were detected as they underwent inverse beta decay on hydrogen nuclei, producing a positron and a neutron. The positron would then annihilate producing two  $\sim 0.51$  MeV, detectable, gamma-ray photons in opposite directions. The experiments used 200 L of water doped with cadmium chloride to capture the neutron. This capture occurred several microseconds after the neutrino interaction. Upon capture, gamma-rays with a combined energy totalling 9 MeV were released. The unique signal caused by the instant gamma rays from the position decay and the delayed gamma rays from the neutron capture was used to identify interactions and could, therefore, be used to detect the neutrinos. To detect the gamma rays, three tanks of the doped water were sandwiched between two liquid scintillator tanks, which were coupled to PhotoMultiplier Tubes (PMTs). The result of this  $\bar{\nu}_e$  experiment was that they were able to detect and confirm the existence of neutrinos.

It was also in 1956 that Yang and Lee [50] reviewed the experimental evidence for parity conservation in the weak interaction. After finding no evidence they suggested a general interaction Hamiltonian of

$$\mathcal{H}_I = \sum_i \psi_p^* \mathcal{O}_i \psi_n \psi_e^* \mathcal{O}^i (G_i - G'_i \gamma_5) \psi_v^* + h.c, \quad (2.3)$$

where the difference between  $G_i$  and  $G'_i$  denotes the strength of the parity violation. In 1957, Wu et al. [1] considered changes in the directional rate of  $e^-$  from beta decay of oriented  $^{60}\text{Co}$  to show that the weak interaction does not conserve parity. It was known that the spin vector would stay constant under a parity flip, while the momentum of the electron would switch by 180 degrees. Therefore parity is only conserved if the count rate of the electrons is constant when the  $^{60}\text{Co}$  spin is flipped. Figure 2.2 shows the change in rate when flipping the  $^{60}\text{Co}$ , which proves parity is not conserved.

As a result of the Wu et al. parity observation, Salam [51], Landau [52], and Lee and Yang [53] proposed that neutrinos are described by a two-component wavefunction  $\psi = \psi_L + \psi_R$  in the standard Hamiltonian, Equation 2.2 [54]. The theory also stipulated that the neutrino field only appears in interactions as  $(1 - \gamma_5)\psi = \psi_L = v_L$ . It also stipulated that the neutrino mass is zero.

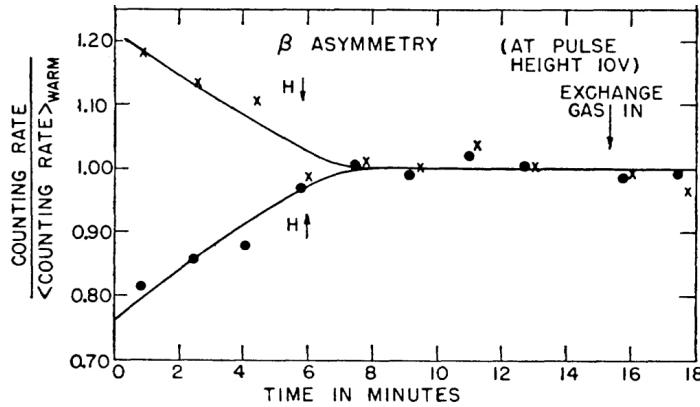


FIG. 2. Gamma anisotropy and beta asymmetry for polarizing field pointing up and pointing down.

Fig. 2.2 The asymmetry in the electron count rate for the change in spin orientation of Co(60) in the Wu et al. experiment. From Reference [1]

The fact that neutrinos are left-handed was observed by Goldhaber et al. [55]. They calculated the helicity of the neutrino by considering the polarisation of light emitted from resonance fluorescence in  $^{152}\text{Sm}$ . The experiment considered an initial  $^{152}\text{Eu}$  which would undergo electron capture of an orbiting electron, releasing a neutrino. The resultant excited  $^{152}\text{Sm}$  nucleus would then decay releasing a photon. The photon would only have sufficient energy to interact with another Sm nuclei if the photon direction was roughly opposite to the neutrino direction. This caused the photon to be Doppler shifted to higher energies due to the nuclear recoil. In this case, the photon would have the same helicity as the neutrino. The photon helicity was measured by applying a magnetic field to the de-excitation photon. If the photon was oppositely aligned to the magnet it could undergo Compton scattering whilst photons with an aligned spin state could not. The photons which were not effected would go on to produce fluorescent light on contact with a  $\text{Sm}_2\text{O}_3$  source. Therefore, the difference in the number of photons detected when the magnetic field was aligned and when it was not aligned indicated the polarisation of the light and the helicity of the neutrino.

After other experiments also indicated that the weak interaction occurs with only left-handed particles, in 1957 Feynman and Gell-Mann [56] and Sudarshan and Marshak [57] generalised the theory proposed by Yang and Lee to account for all particles in the weak interactions. Forcing interactions to occur via the left-handed field,  $\psi_L$ , reduced the interaction Hamiltonian to vector and axial-vector terms only and hence the theory is known as the V-A theory. A charged weak current [47],

$$j^\alpha = 2(\hat{\rho}_L \gamma^\alpha n_L + \bar{\nu}_{\mu L} \gamma^\alpha e_L + \bar{\nu}_{e L} \gamma^\alpha \mu_L), \quad (2.4)$$

was also added to V–A theory to account for the  $\mu - e$  weak interaction resulting in the following interaction Hamiltonian [47]

$$\mathcal{H}_I = \frac{G_F}{\sqrt{2}} j^\alpha j_\alpha^+. \quad (2.5)$$

V–A theory was able to explain the various observed Beta decay transition as well as the ratio of pion decays to muons with respect to electrons [58]. However, at the time the theory had two problems. The first was that only one flavour of neutrino was assumed to exist. In this situation some reactions which were allowed by the theory had not been observed, such as the decay of a muon to an electron and a gamma ray [59]. Secondly, is that cross sections increases as function of incoming particles momentum [59]. This means the theory eventually breaks down at higher energies. The solution to first problem is to introduce an additional neutrino, the muon neutrino. The solution to the second is to introduce a mediator Boson. However introducing a boson implies the branching ratio of  $(\mu \rightarrow e + \gamma)/(\mu \rightarrow e + \nu + \bar{n}\bar{\nu}) \sim 10^{-4}$  [2] where in fact, as eluded to above, the  $\mu \rightarrow e + \gamma$  reaction had not been detected and limits on the branching ratio were less than  $10^{-8}$  [2, 60, 47].

In 1962, an experiment at Brookhaven performed by Lederman, Schwartz and Steinberger et al. [2] discovered the muon neutrino. In order to produce the neutrino beam required by the experiment, protons were fired at a beryllium target producing pions. The pions then decayed resulting in the neutrino beam. The outgoing charged particles created from pion decays were stopped by a 13.5 m thick iron shield wall. This wall only allowed the neutrinos (mainly muon neutrinos) to pass through to a spark chamber in sufficient quantities for detection [2]. A spark chamber is made up of several parallel plates of which alternating plates are connected to a fast high voltage pulse. Between the plates is a gas, e.g. neon, which is ionised as charged particles pass through [59]. If a high voltage pulse is applied when the charged particle passes through, the ionised electrons are accelerated. This causes an avalanche effect resulting in a spark. The sparks are then captured in a photograph. From the photograph, the direction and an estimate of the particle momentum can be obtained by considering how many plates the particle passes through and the location of the sparks. The Brookhaven experiment used 4 foot squared plates with 9 plates in a single module (on a single layer).

Neutrinos that interacted within the detector were identified by the topology of the decay products. If there were only one flavour of neutrino, then the neutrinos from the beam would equally produce muons and electrons from the interactions in the detector. However, if muon neutrinos exist and the beam neutrinos were specifically muon neutrinos, then only muon

events would be detected. The experiment observed 34 muon events and only 6 electron events [2]. The electron events can be explained via background cosmic events and there are not sufficient events to confirm the null hypothesis, therefore this experiment successfully proved the existence of muon neutrinos [2]. This also in turn showed that the decay of  $\mu \rightarrow e + \gamma$  was forbidden, easing the pressure on a mediator boson theory. The result of the experiment also prompted Pontecorvo to consider oscillations between the two flavours [35].

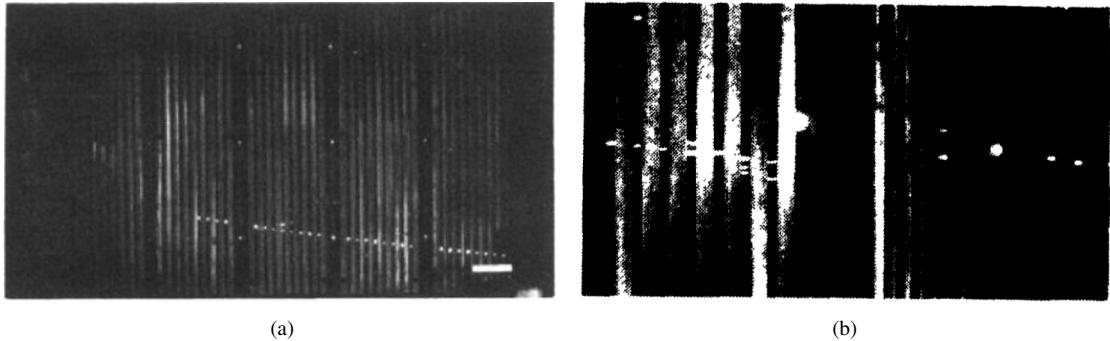


Fig. 2.3 A Muon track (a) and an electron shower (b) in the Brookhaven spark chamber detector. From Reference [2]

The current theory that describes the weak force in the SM is the electroweak theory, pioneered by Weinberg [61], Salam [62] and Glashow [63] in the 1960s. The theory is based on the  $SU(2) \times U(1)$  symmetry breaking where the  $U(1)$  symmetry is known as hypercharge. The theory consists of four mediator bosons,  $B_\mu$ ,  $W_\mu^\alpha$  ( $\alpha = 1 \rightarrow 3$ ). The Lagrangian that describes the boson interactions is [64]

$$\mathcal{L}_{int} = -\frac{1}{4}(W_{\mu\nu}^\alpha)^2 - \frac{1}{4}B_{\mu\nu}^2 + (D_\mu H)^\dagger(D_\mu H) + m^2 H^\dagger H - \lambda(HH^\dagger)^2, \quad (2.6)$$

where  $H$  is the Higgs field,  $D_\mu H = \partial_\mu H - igW_\mu^\alpha \tau^\alpha H - \frac{1}{2}ig'B_\mu H$ ,  $W_{\mu\nu}^\alpha$  is the field strength and  $\tau$  are matrix generators. One can extract the mass terms of the gauge bosons in the theory by setting  $\tau = \frac{\sigma^a}{2}$ , then expanding the Higgs fields around the vacuum expectation value, and finally setting to the unitary gauge. From this, the linear combinations that describe the photon and the weak bosons can be deduced [64]:

$$\begin{aligned} A_\mu &= \sin(\theta_w)W_\mu^3 + \cos(\theta_w)B_\mu \\ Z_\mu &= \cos(\theta_w)W_\mu^3 - \sin(\theta_w)B_\mu \\ W_\mu^+ &= \frac{1}{\sqrt{2}}(W_\mu^1 - iW_\mu^2) \\ W_\mu^- &= \frac{1}{\sqrt{2}}(W_\mu^1 + iW_\mu^2). \end{aligned} \quad (2.7)$$

Combinations of the  $W^3$  and  $B$  fields and the rotation angle  $\theta_w$  can be used to extract the photon ( $A$ ) and the Z boson ( $Z$ ).

The left-handed fermion fields form lepton doublets,  $L^i = \begin{pmatrix} v_{iL} \\ e_{Li} \end{pmatrix}$ , and quark doublets,  $Q^i = \begin{pmatrix} u_{iL} \\ d_{Li} \end{pmatrix}$ , that couple to the weak interaction as a SU(2) doublet. The right-handed particles form singlets,  $e_R^i$ ,  $u_r^i$  and  $d_R^i$ .

The tau neutrino was not discovered until the year 2000 by the DONUT (Direct observation of the nu tau, E872) experiment [65] and was the last lepton predicted by the SM to be detected. DONUT was situated on the Fermilab Tevatron particle accelerator and detected tau neutrino interactions arising from protons interacting in a 1 m tungsten "beam dump" (or target) [65]. The majority of the tau neutrinos were specifically created during the decay of  $D_S$  mesons,  $D_S \rightarrow \tau + \nu_\tau$ , which were created during the proton interactions. The  $\nu_\tau au$  then interacted in an nuclear emulsion detector, 36 m downstream of the beam dump [65]. A variety of magnets were used to remove charged particles from the beam and additional detectors were used to help identify the flavour and momentum of charged particles arising from neutrino interaction in the emulsion detector. Once the plates of interest were identified, the plates underwent an automated optical scanning procedure. Typically, at the energy of the experiment the resultant tau from  $\nu_\tau$  interactions decays within 2 mm to a single charged daughter [65]. Therefore plates of interest containing a vertex and a track with a kink were considered tau neutrino candidates and analysed. Four tau neutrino interactions with an estimated background of  $0.34 \pm 0.05$  events were identified [65].

In summary, neutrinos were postulated by Pauli and proven to exist by Cowan and Reines. The work of Wu et al. and Goldhaber et al. showed neutrinos to only interact and exist in the left-handed neutrino field, as proposed by Yang and Lee. The existence of muon neutrinos was proven by the Brookhaven experiment. This led to the theoretical work of Salam, Landau, Lee and Yang, Feynman and Gell-Man, Sudarshan and Marshak, Weinberg and Glashow, which incorporated the left-handed helicity of the neutrino into theory. This work resulted in the SM, which defines the neutrino to be left-handed, massless and only interact via the weak force. The tau neutrino was discovered by the DONUT experiment, which was the final piece of the SM.

However, strong evidence, see Section 2.2.1, now proves neutrino oscillations which require neutrinos to have mass. Therefore, one must go beyond the original formulation of the SM to explain these phenomena. This is discussed in the next Section.

## 2.2 Beyond the Standard Model in Neutrino Physics

Experimental observations have provided evidence that contradicts the SM. In particular, evidence for neutrino oscillations where one type of neutrino changes to another during propagation has been observed which would require neutrinos to have mass. The theory of neutrino oscillations is discussed in the next Section, and the evidence for neutrino oscillations follows in Section 2.2.1. However, the neutrino mass has not been directly measured because it is too small: smaller than its other lepton partner. Experimental searches for neutrino mass are discussed in Section 2.2.2 and have found a current upper limit of 1.1 eV (90% confidence level) (from the KArlsruhe TRItium Neutrino experiment (KATRIN) [66]) for the absolute mass scale of neutrinos. One explanation of why the neutrino mass is smaller than its other lepton partner is the seesaw mechanism, which is discussed in Sections 2.2.3. This mechanism requires the particles to be Majorana and requires the existence of a right-handed neutrino field, i.e. a sterile neutrino. As well as this, the seesaw mechanism links back to neutrino oscillations and provides a theoretical justification for them. The addition of another neutrino field in the seesaw mechanism can be probed by considering the number of neutrinos. The experiments that consider the number of neutrinos are discussed in Section 2.2.4. The theory behind sterile neutrinos is then discussed in Section 2.2.5. Finally, the neutrino oscillation anomalies which have led to the search for sterile neutrinos, as well as the current experimental searches for sterile neutrinos, are discussed in Section 2.2.6.

### 2.2.1 Neutrino Oscillations

Neutrino oscillations were first postulated by Pontecorvo in 1957 [67]. Originally Pontecorvo proposed [67] the oscillation of neutrinos to anti-neutrinos and visa-versa. Oscillation theory continued to developed with Maki, Nakagawa and Sakata [36] who considering oscillations between  $\nu_e$  and  $\nu_\mu$  flavour eigenstates. Later work by Pontecorvo also considered flavour oscillations [35] and also introduced oscillations between  $N$  neutrinos [68]. Pontecorvo, Bilenky [69] and others [70][71] then utilised an analogy of quark mixing to form the current neutrino vacuum oscillation formalism. The theory now describes the mixing of the mass eigenstates ( $\nu_i$  i=1,3) and the flavour eigenstates ( $\nu_\alpha$  ( $\alpha = e/\mu/\tau$ )) by a mixing matrix called the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix [36]. This Section will describe the theory behind neutrino oscillations [72, 73] and the experimental evidence for the phenomena.

## Neutrino Oscillation Theory

The evolution of a neutrino,  $\nu$ , is described by the time dependent Schrödinger equation

$$\frac{d}{dt}\nu = (H_0 + H_I)\nu, \quad (2.8)$$

where  $H_0$  is the Hamiltonian operator which describes the neutrino propagation in space and  $H_I$  ( $= 0$  in a vacuum) is the interaction Hamiltonian. The mass eigenstates of the neutrinos, which describe how the neutrinos propagate, and the flavour eigenstates, which describe the way the particles interact, are different. The mixing of the mass eigenstates  $|\nu_i\rangle$  and flavour eigenstates  $|\nu_\alpha\rangle$  is described by the PMNS matrix, i.e.

$$|\nu_\alpha\rangle = U_{PMNS} |\nu_i\rangle. \quad (2.9)$$

The PMNS matrix,  $U_{PMNS}$ , is a  $3 \times 3$  unitary matrix, which, in a vacuum, can be parameterised as three rotation matrices [36]

$$\begin{bmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{bmatrix} \begin{bmatrix} c_{13} & 0 & s_{13}e^{-i\delta_{CP}} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta_{CP}} & 0 & c_{13} \end{bmatrix} \begin{bmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2.10)$$

$U_{li}$  describes the amplitude of the mass eigenstate,  $i$ , in flavour eigenstate,  $l$ :  $c_{ij} = \cos \theta_{ij}$ ,  $s_{ij} = \sin \theta_{ij}$  and  $\theta_{ij}$  are the mixing angles between mass eigenstates  $i$  and  $j$ :  $\delta_{CP}$  is the parity violating phase. Note: a fourth additional matrix with two phases is added if neutrinos are Majorana. This however does not affect the oscillation probability. When considering the passage through matter, further corrections are also required to the PMNS matrix to account for the Mikheyev–Smirnov–Wolfenstein (MSW) effect [74]. The MSW effect describes the change in the neutrinos' mass eigenstates due to coherent forward scattering of electrons within matter. This is equivalent to introducing an interaction term to the Hamiltonian ( $H_I \neq 0$ ).

The probability of neutrino  $\alpha$  oscillating to neutrino  $\beta$  is then given by

$$P_{\alpha\beta} = \left| \langle \nu_\beta | \nu_\alpha(t) \rangle \right|^2 = \left| \sum_{i=1}^n \sum_{j=1}^n U_{\alpha i}^* U_{\beta j} \langle \nu_j | \nu_i(t) \rangle \right|^2. \quad (2.11)$$

Usually the wavefunction of a neutrino is described as a plane waveform  $|\nu_i(x, t)\rangle = e^{-ip_\mu x^\mu} |\nu_i(0)\rangle$ , where  $p_\mu = -E, p_x, p_y, p_z$  and  $x_\mu = t, -x_x, -x_y, -x_z$  and therefore  $p_\mu x^\mu = Et - \mathbf{p} \cdot \mathbf{x}$ . It should be noted that the neutrino can also be considered as a wave-packet,

which provides an alternative approach to calculating the oscillation probability [75]. In most cases, due to experimental design, oscillation theory can be approximated to the two neutrino form to facilitate comprehension and is used below. However, in modern experiments the three flavour formalism is used in analyses due to the sensitivity of the experiments. Assuming neutrinos are described by the plane wave solution and are relativistic, then  $p_i \approx E_i + (m_i^2/2E_i)\hat{p}$  and  $t = x = L\hat{p}$ . Therefore, the two flavour probability can be written as

$$P_{\alpha \rightarrow \beta} = \sin^2(2\theta) \sin^2\left(\frac{\Delta m^2 L}{4E}\right), \quad (2.12)$$

where  $E_\alpha = E_\beta = E$  and  $\Delta m^2 = m_\alpha^2 - m_\beta^2$ . The three flavour neutrino probability is given by:

$$\begin{aligned} P_{\alpha \rightarrow \beta} = & \delta_{\alpha\beta} - 4 \sum_{i>j} \operatorname{Re}(U_{\alpha i}^* U_{\beta i} U_{\alpha j} U_{\beta j}^*) \sin\left(\Delta m_{ij}^2 \frac{L}{4E}\right) \\ & + 2 \sum_{i>j} \operatorname{Im}(U_{\alpha i}^* U_{\beta i} U_{\alpha j} U_{\beta j}^*) \sin\left(\Delta m_{ij}^2 \frac{L}{4E}\right) \end{aligned} \quad (2.13)$$

Table 2.1, from Reference [34], shows the current measured values of the oscillation parameters. The Table shows the case for the ordered hierarchy, where the  $\nu_e$  is the lightest of the three neutrinos, and the inverted hierarchy, where the  $\nu_\tau$  is the lightest. The next Section will discuss the experimental observations corresponding to the elements in Table 2.1.

	Normal Ordering	Inverted Ordering
$\sin^2 \theta_{12}$	$0.310^{+0.013}_{-0.012}$	$0.310^{+0.013}_{-0.012}$
$\theta_{12}/^\circ$	$33.82^{+0.78}_{-0.76}$	$33.82^{+0.78}_{-0.76}$
$\sin^2 \theta_{23}$	$0.582^{+0.015}_{-0.019}$	$0.582^{+0.015}_{-0.018}$
$\theta_{23}/^\circ$	$49.7^{+0.9}_{-1.1}$	$49.7^{+0.9}_{-1.0}$
$\sin^2 \theta_{13}$	$0.02240^{+0.00065}_{-0.00066}$	$0.02236^{+0.00065}_{-0.00066}$
$\theta_{13}/^\circ$	$8.61^{+0.12}_{-0.13}$	$8.65^{+0.12}_{-0.13}$
$\delta_{CP}/^\circ$	$217^{+40}_{-28}$	$280^{+25}_{-28}$
$\Delta m_{21}^2 / 10^{-5} \text{ eV}^2$	$7.39^{+0.21}_{-0.20}$	$7.39^{+0.21}_{-0.20}$
$\Delta m_{31}^2 / 10^{-3} \text{ eV}^2$	$2.525^{+0.033}_{-0.031}$	$-2.512^{+0.034}_{-0.031}$

Table 2.1 Best fit oscillation parameters of the PMNS matrix, assuming the normal and inverted hierarchy. Errors are the  $1\sigma$  confidence intervals. From Reference [34]

### Experimental Observations of Neutrino Oscillations

The Homestake experiment [37], performed by Ray Davis from 1965 to 1994, was the first neutrino experiment which provided some evidence for neutrino oscillations. The experiment detected solar neutrinos via capture upon  $^{37}\text{Cl}$  atoms in  $380\text{ m}^3$  of  $\text{C}_2\text{Cl}_4$  in the Homestake Gold Mine in South Dakota (1478 m below sea level). The capture of electron neutrinos upon a  $^{37}\text{Cl}$  atom, which has a threshold of 0.814 MeV, results in the creation of an  $^{37}\text{Ar}$  atom [76]. The argon atoms were then counted every few months using a proportional counter to calculate the neutrino rate.

$0.478 \pm 0.03$   $^{37}\text{Ar}$  were detected per day which is equivalent to a  $^{37}\text{Ar}$  production rate of argon via solar neutrinos of  $2.56 \times 10^{-36}\text{ sec}^{-1}(^{37}\text{Cl}_{\text{atom}})^{-1}$  [76]. This result is significantly less than the emission rate of solar neutrinos calculated by Bahcall of  $9.3 \times 10^{-36}\text{ sec}^{-1}(^{36}\text{Cl}_{\text{atom}})^{-1}$  [77, 76].

Further experiments followed which also showed the deficit of solar neutrinos to be approximately 1/3. Kamiokande [3] and then Super-Kamiokande (Super-K) [4, 78, 79], which are both water-based Cherenkov detectors of the same style, discovered a deficit in the solar neutrino flux in the elastic scattering channel, see Figure 2.4 for the Feynman diagram. Super-K, situated in Japan 1000 m underground, contains 50 ktons of pure water surrounded by 50 cm Photo-Multiplier Tubes (PMTs). When neutrinos interact with either a nucleus within the water molecules or the electrons, final state charged particles, with sufficient kinematic energy, produce Cherenkov radiation which is detected as rings by the PMTs. The Cherenkov radiation occurs from the charged particles, travelling faster than speed of light in the medium. The charged particles electromagnetically interact with the medium which results in a wave front analogous to a sonic boom and creates a cone of light. Therefore for charge current events where there is an outgoing charged lepton partner, the radius and location of the ring indicate the position of the interaction and the direction of travel of the lepton partner. Particle identification is performed by considering the fuzziness of the rings. Electron events are more blurred than muons at high energies due to electromagnetic showers, and at low energies due to the multiple scattering.

Kamiokande showed a deficit of neutrinos coming from the sun, see Figure 2.5, with the rate being 46% of the expected value [78]. This is shown in Figure 2.5 as a deficit against the solar prediction when considering events with an angle corresponding to the direction of the sun. A modern solar analysis in Super-K, selecting events below  $\sim 60$  MeV, found the solar flux to be  $2.34 \times 10^6 \text{ cm}^2/\text{s}$  which was  $\sim 40\%$  of the flux predicted by the solar models.

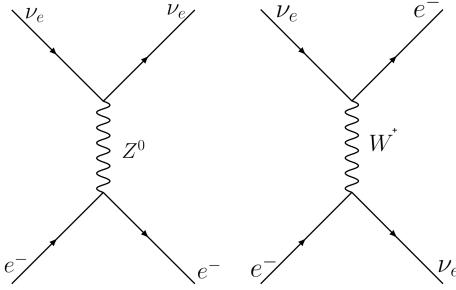


Fig. 2.4 Feynman diagrams of the elastic scattering interactions between  $\nu_e$  and  $e^-$ .

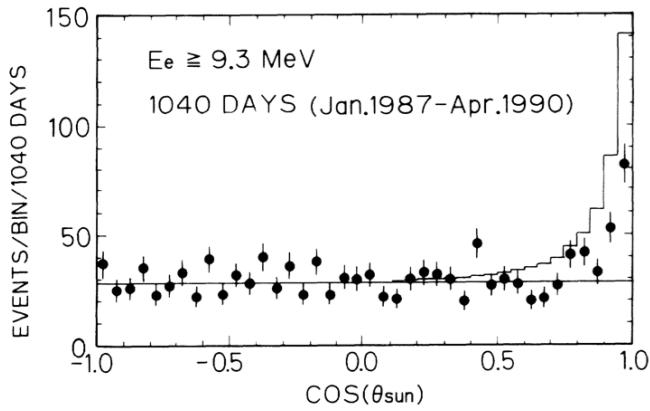


Fig. 2.5 The solar deficit observed in Kamiokande as a function of the angle from the sun. The solid line indicates the Monte Carlo prediction. From Reference [3].

The proof that the solar deficit was due to neutrino oscillations came from measurements of atmospheric neutrinos [4]. Decays of  $\pi^+$  and  $\mu^+$  cosmic rays in the atmosphere are expected to give a  $\nu_\mu/\nu_e$  ratio of 2. An analysis was performed on atmospheric neutrinos to identify and reconstruct the direction of atmospheric neutrinos in Super-K [4, 80–82]. Events with energy below 100 MeV were cut from the analysis to remove solar neutrinos and the outgoing lepton partner direction was used as the neutrino direction. The ratio between the prediction from Monte Carlo and the experimental data was found to be  $\sim 0.61$  [80]. A zenith angle asymmetry was observed in the muon neutrino data and no asymmetry in the electron neutrino data, see Figure 2.6, indicating that the ratio changed as a function of the zenith angle. Figure 2.6 shows the deficit in the predicted muon rate within different energy ranges. This deficit showed that neutrinos created on the opposite side of the world oscillate from muon neutrinos to tau neutrinos [83]. The data from atmospheric neutrinos experiments provide measurements of  $\Delta m_{32}^2 = m_3^2 - m_2^2$  and  $\theta_{23}$  which describe the survival probability in

the two neutrino approximation,

$$P(\nu_\mu \rightarrow \nu_\mu) = 1 - \sin^2(\theta_{23}) \sin^2 \left( \frac{1.267 \Delta m^2 (\text{eV}^2) L (\text{km})}{E (\text{GeV})} \right). \quad (2.14)$$

The IceCube [84, 85] experiment is also sensitive to atmospheric neutrinos and provides measurements of  $\Delta m^2_{32}$  and  $\theta_{23}$ . The experiment employs an array of 4800 Digital Optical Modules (DOMs) installed on 80 strings between 1450 m and 2450 m below the surface of south pole [84]. The detector is enclosed in a 1 km<sup>2</sup> area and uses PMTs to detect Cherenkov light emitted by charged particles in the ice. Thus a neutrino interaction is detected using the final state charged particles in the interaction. The data from the DOMS is converted into light patterns which are used to reconstruct the direction and energy of the interactions. Similar to the water Cerenkov detectors the type of neutrino can be identified from final state charged lepton partner in a charge current interaction. Muon neutrino events can be identified from track-like topology of the muon whereas electron neutrino events are more spherical due to the electron cascade. The results of IceCube [85] contribute to the global fit parameters of  $\Delta m^2_{32}$  and  $\theta_{23}$  in Table 2.1.

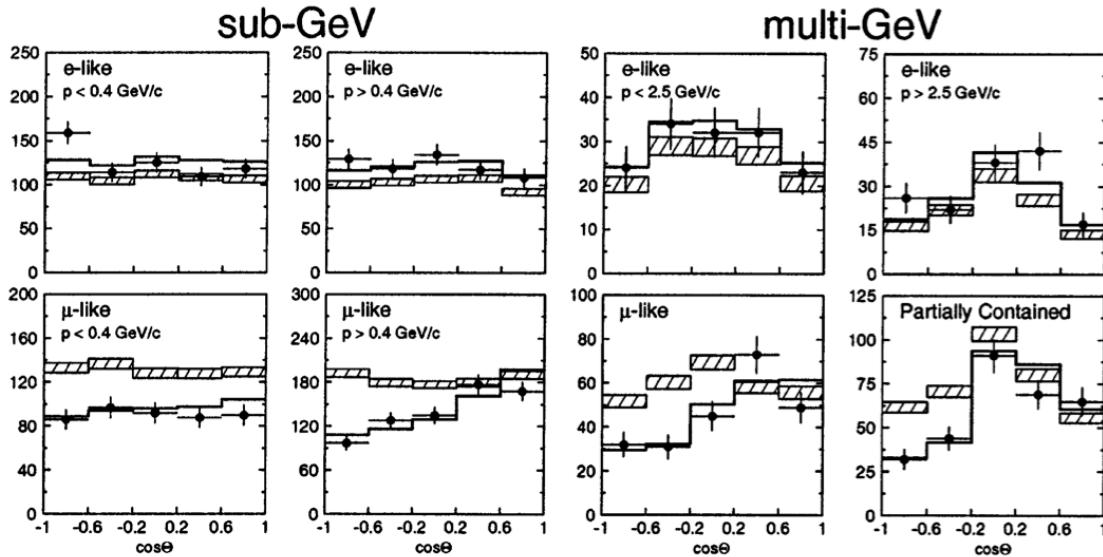


Fig. 2.6 Electron and muon rates in Super-K as a function of the zenith angle. The hatched squares correspond to Monte Carlo simulations assuming no neutrino oscillations while the black line corresponds to the best fit of the data. From Reference [4].

Further proof of neutrino oscillations was presented by the Sudbury Neutrino Observatory (SNO) [38] collaboration after the results from Super-K [5, 86]. The SNO detector was a 7 ktonne, water-based Cherenkov detector placed 2 km underground near Sudbury Canada.

The detector consists of 1 ktonne of heavy water, D<sub>2</sub>O, and 6 ktonne of pure water surrounded by PMTs [38]. Using heavy water, the SNO detector was able to identify solar neutrinos through three interaction channels [5]

$$\begin{aligned} \nu_e + d &\rightarrow p + p + e^- \text{ (Charge Current (CC))}, \\ \nu_x + d &\rightarrow p + n + \nu_x \text{ (Neutral Current (NC))}, \\ \nu_x + e^- &\rightarrow \nu_x + e^- \text{ (Elastic Scattering (ES))}. \end{aligned}$$

Only the  $\nu_e$ s interacted via the CC channel as the energies of neutrinos from the sun are lower than the mass of the muon and tau leptons. For the NC events, the freed neutrons capture upon deuteron nuclei, releasing gamma photons. The gamma photons, in turn, cause Cherenkov radiation which is detected.

Fitting the data, the flux of each interaction was found and used to calculate the flux of solar neutrinos [5]. Figure 2.7 shows how the flux of the three neutrino interaction types combine in the SNO experiment to identify the flux of the neutrino flavours from the sun. From the first data report, the flux was found to be [5]

$$\begin{aligned} \Phi_{CC} &= 1.76^{+0.06}_{-0.05} \text{(stat.)}^{+0.09}_{-0.09} \text{(syst.)} \times 10^6 \text{ cm}^{-2} \text{s}^{-1}, \\ \Phi_{NC} &= 2.39^{+0.24}_{-0.23} \text{(stat.)}^{+0.12}_{-0.12} \text{(syst.)} \times 10^6 \text{ cm}^{-2} \text{s}^{-1}, \\ \Phi_{ES} &= 5.09^{+0.44}_{-0.43} \text{(stat.)}^{+0.46}_{-0.43} \text{(syst.)} \times 10^6 \text{ cm}^{-2} \text{s}^{-1}. \end{aligned}$$

This implies a total  $\nu_\mu$  and  $\nu_\tau$  flux of  $3.41 \times 10^6 \text{ cm}^{-2} \text{s}^{-1}$  from solar neutrinos, where the Standard Solar Model (SSM) expect no flux [5].

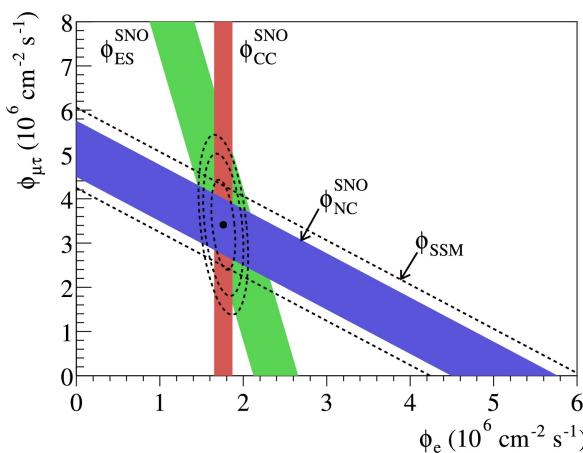


Fig. 2.7 The flux,  $\phi_{\mu\tau}$ , of the  $\nu_\mu$  and  $\nu_\tau$  neutrinos from B<sup>8</sup> interactions in the sun against the flux,  $\phi_e$ , of  $\nu_e$  for the three types of interactions in SNO. The contours indicate the best fit and the dashed lines indicated the predicated flux from the SSM. From Reference [5].

The results from the solar experiments (Homestake) [76], the Gallium Experiment (GALLEX/GNO) [87, 88], Soviet American Gallium Experiment (SAGE) [89], Super-K [79], SNO [5] and Borexino [90] (Italian diminutive of BOREX (BORon solar neutrino EXperiment)) can be used with results from KamLAND [91] (Kamioka Liquid Scintillator Antineutrino Detector), a long baseline ( $\sim 180$  k) reactor experiment, to provide measurements of  $\Delta m_{21}^2$  and  $\theta_{21}$  [34]. A global combined fit of these measurements result in the values shown in Table 2.1. Note that the MSW effect allows for the sign of  $\Delta m_{21}^2$  to be determined [92].

Once neutrino oscillations were proven, many experiments were then designed to calculate the mixing parameters of the PMNS matrix, Equation 2.10, such as Long BaseLine neutrino (LBL) experiments. LBL experiments usually consist of two detectors: a near detector, which is normally situated  $\sim 1$  km away from the production of the neutrino beams, and a far detector, normally situated several hundred kilometres away from the production of the neutrino beam. The positioning of far detector is usually fixed due to location requirements i.e.  $\sim 1$  km underground in a mine. In order to get a maximum oscillation signal, the mean energy of the neutrino beam has to be tuned. It is the ratio of  $L/E$  (distance travelled divided by the neutrino beam energy) which is important when defining an effective oscillation experiment.

In LBL experiments, the neutrino rates in near detectors provide well constrained flux systematics allowing for precise neutrino oscillation measurements. The beams are either muon neutrinos or muon anti-neutrinos. This is because neutrino beams are made from impacting protons onto targets resulting in high energy mesons, such as pions. The pion decay to an electron and electron neutrino is suppressed due V-A nature of the weak force (the right-handed helicity of the lighter electron (hence more relativistic) is suppressed). Therefore, the experiments are sensitive to either muon disappearance or electron appearance. For example the Main Injector Neutrino Oscillation Search (+) (MINOS(+)) [93] and NuMI Off-axis  $\nu_e$  Appearance (NovA)[94] experiments are situated on the  $\nu_\mu$  NuMI [95, 96] (Neutrinos at Main Injector) beamline at Fermilab which peaks at approximately 5 GeV. Other LBL experiments, such as T2K [97] (Tokai to Kamioka, Japan) and K2K [98] (KEK to Kamioka), also search for  $\nu_\mu$  disappearance with beam energies  $\mathcal{O}1$  GeV. At such energies,  $\Delta m_{32}^2 L/(2E) \sim 1$  and  $\Delta m_{21}^2 L/(2E) \ll 1$  as  $\Delta m_{32}^2/\Delta m_{21}^2 \sim 33$ . Therefore,  $\sin(\Delta m_{21}^2 L/(2E))$  can be approximated to 0 and the oscillation probability, Equation 2.13, can be approximated as [97]

$$P(\nu_\mu \rightarrow \nu_\mu) \approx 1 - 4 \cos^2(\theta_{13}) \sin^2(\theta_{23}) [1 - \cos^2(\theta_{13}) \sin^2(\theta_{23})] \sin^2 \left( \frac{1.267 \Delta m_{32}^2 (\text{eV}^2) L (\text{km})}{E (\text{GeV})} \right). \quad (2.15)$$

Therefore, these LBL experiments searching for disappearance signals in muon neutrino beams have sensitivity to  $\Delta m_{32}^2$  and  $\theta_{23}$ . This approximation is similar to the two neutrino approximation that approximately describes in the atmospheric case, Equation 2.14, as  $\theta_{13}$  is small. The Oscillation Project with Emulsion-tRacking Apparatus (OPERA) [99] experiment has also detected  $\nu_\tau$  appearance from a muon neutrino beam providing further proof of neutrino oscillations.

T2K [100], MINOS(+) [101] and NovA [102] are also sensitive to  $\nu_e$  appearance and therefore can provide a measurement of  $\theta_{13}$ . This is seen by approximating  $\Delta m_{21}^2 \sim 0$  and simplifying Equation 2.13 to [103]

$$P(\nu_\mu \rightarrow \nu_e) = \sin^2(2\theta_{13}) \sin^2(\theta_{23}) \sin^2 \left( 1.267 \Delta m_{32}^2 (\text{eV}^2) \frac{L (\text{km})}{E (\text{GeV})} \right). \quad (2.16)$$

The reactor neutrino detectors such as Daya Bay [104, 105], Double Chooz [106] (Chooz [107]) and Reactor Experiment for Neutrino Oscillation (RENO) [108] are also sensitive to  $\theta_{13}$ . This is because reactor experiments measure the disappearance of electron anti-neutrinos over distances of  $\sim \mathcal{O}1$  km. The probability for electron neutrino disappearance is described as

$$\begin{aligned} P_{dis} = & \cos^4(\theta_{13}) \sin^2(\theta_{12}) \sin^2(\Delta_{21}) \\ & + \cos^2(\theta_{12}) \sin^2(2\theta_{13}) \sin^2(\Delta_{31}) \\ & + \sin^2(\theta_{12}) \sin^2(2\theta_{13}) \sin^2(\Delta_{32}), \end{aligned} \quad (2.17)$$

where  $\Delta_{ij} = 1.267 \Delta m_{ij}^2 (\text{eV}^2) \times 10^3 \frac{L (\text{km})}{E (\text{MeV})}$  and  $\Delta m_{ij} = m_i^2 - m_j^2$ . As  $\theta_{13} < 10^\circ$ ,  $\cos^4(\theta)_{13} \sim 1$  [104]. Also as  $m_{ij}^2 \ll 1 \times 10^{-2} (\text{eV}^2)$  then  $\Delta_{ij} < 1$  for distances of 1 km and neutrino energies of 1-10 MeV. In addition,  $\Delta m_{12}^2 \ll |\Delta m_{13}^2| \sim |\Delta m_{32}^2|$  and hence the  $\Delta_{13}$  term dominates. Therefore, for reactor experiments, the electron anti-neutrino disappearance probability can be approximately described by the two flavour oscillation probability [106]:

$$P_{dis} \approx P_{13} = 1 - \sin^2(2\theta_{13}) \sin^2 \left( 1.267 \Delta m_{13}^2 (\text{eV}^2) \frac{L (\text{km})}{E (\text{GeV})} \right). \quad (2.18)$$

Due to the energy of reactor neutrinos, the reactor experiments employ a gadolinium doped liquid scintillator detector. Inverse beta decay interactions are then identified via PMTs that detect gamma rays from both positron annihilation and neutron capture upon  ${}^1\text{H}$  or the gadolinium. Using near and far detectors, uncertainties on the models of reactor cores are constrained, resulting in a precise  $\theta_{13}$  measurement [109]. In reactor experiments,

a difference between the observed event rate and predicted event rate has been observed with a ratio of  $R = 0.940 \pm 0.011(\text{stat.}) \pm 0.004(\text{syst.})$  [110] [42, 111]. This is known as the reactor anti-neutrino anomaly and is yet to be explained.

The CP phase ( $\delta_{CP}$ ) can be detected in oscillation experiments (T2K [112], NOvA [102]) which have sensitivity to neutrino and anti-neutrino appearance. Differences in the observed neutrino and anti-neutrino oscillations are theoretically due to the CP phase, and analyses require other oscillation parameters in order to calculate the CP phase. IceCube [113] has also placed constraints on the CP phase by considering ultra high energy neutrinos created in cosmic events. Future experiments HyperK [114] and Deep Underground Neutrino Experiment (DUNE) [115] are being constructed in order to pin down the value of the CP phase.

All the experiments discussed in this Section provide evidence for neutrino oscillations and contribute to the calculations of the parameters in the PMNS matrix, which describes neutrino oscillation probability. The oscillation parameters are calculated from the global fits of these experiments, as well as others. A detailed description can be found in Reference [34].

### 2.2.2 Experimental Search for Neutrino Mass

The absolute mass of each neutrino eigenstate is currently not known, and also the ordering is not known. From the oscillation experiments described above, it is known that  $|\Delta m_{32}^2| \approx |\Delta m_{31}^2| >> \Delta m_{12}^2$ . It is also known that  $\Delta m_{12}^2$  is positive. The sign of  $|\Delta m_{32}^2|$  ( $|\Delta m_{13}^2|$ ) determines the ordering. As can be seen in Figure 2.8, if  $\Delta m_{23}^2$  is positive the order is known as "normal", if negative it is known as "inverted". Accelerator experiments, such as NOvA [102], DUNE [115] and HyperK [114], have sensitivity to the mass hierarchy through matter effects [116]. Jiangmen Underground Neutrino Observatory (JUNO) [6, 117], a reactor-based experiment, aims to determine the mass hierarchy by considering searches for "ripples" in the electron anti-neutrino survival rates [6, 7]. The ripples can be seen in Figure 2.9 and are due to the atmospheric mass squared difference ( $\Delta M_a = |m_3 - m_1|$ ). The phase of the ripple is linked to the ordering of the mass hierarchy. Super-K [82] can also take advantage of matter effects to determine the mass hierarchy using atmospheric neutrinos. Precision IceCube Next Generation Upgrade (PINGU) [118] and Oscillation Research with Cosmics in the Abyss (ORCA) [119] are two further experiments which also use atmospheric neutrinos to determine the mass hierarchy.

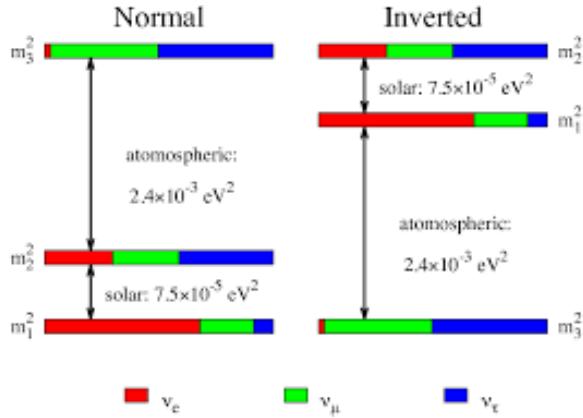


Fig. 2.8 Pictorial explanation of the normal and inverted mass hierarchy, from Reference [6]. The contributions of the flavour eigenstates to the mass eigenstates are shown in the coloured bars. For the first mass eigenstate, the main contribution is the  $\nu_e$ . For the third mass eigenstate, the  $\nu_\mu$  and  $\nu_\tau$  contribute more.

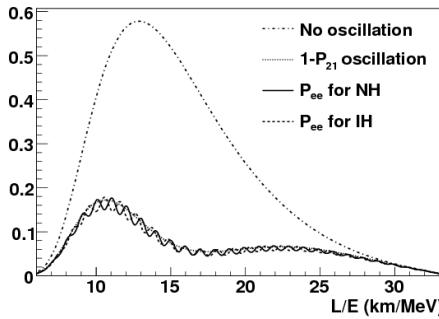
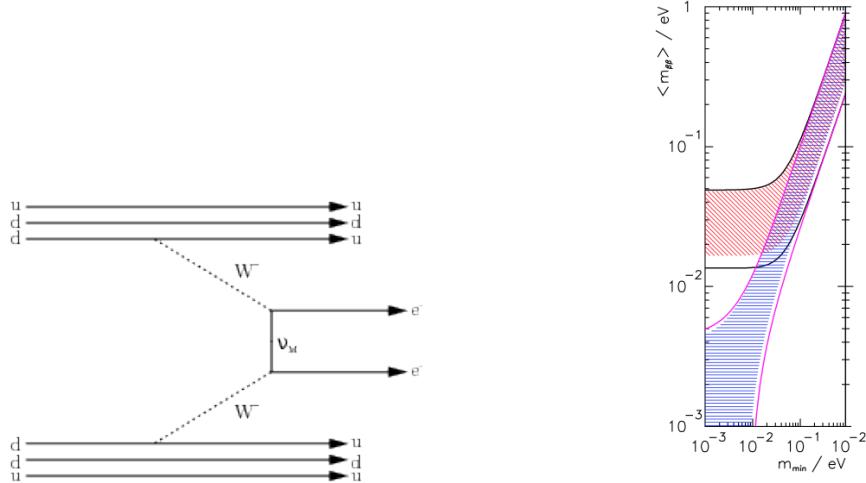


Fig. 2.9 The oscillation probability in reactor experiments with a baseline of 60 km, from Reference [7]. The figure shows the difference in the probability for the normal ordering (black line) and the inverse ordering (dashed line).

Another potential way to measure the mass hierarchy is through neutrinoless double beta decay [120]. As can be seen in Figure 2.10a, neutrinoless double beta decay is where two nucleons undergo beta decay and a neutrino mediates the interaction between the two nucleons [121]. Neutrinoless double beta decay can only occur if the neutrino is a Majorana particle. The decay rate of this transition,  $(T_{1/2}^{0\nu})^{-1}$ , is proportional to the effective Majorana mass  $\langle m_{\beta\beta} \rangle = |\sum_i U_{ei}^2 m_{vi}|^2 = |U_{e1}m_1^2 + U_{e2}m_2^2 + U_{e3}m_3|^2$  [92, 122]. This mass is the sum of the PMNS matrix elements of the  $i^{th}$  mass eigenstate and the electron flavour state,  $U_{ei}$ , multiplied by the mass of the mass eigenstate. For the normal ordering hierarchy,  $m_3 \gg m_2 > m_1$  and, therefore, there is a larger contribution from the  $U_{e3}$  term than for the inverse ordering hierarchy where  $m_3 \ll m_1 < m_2$  [122]. Therefore, if neutrinoless double beta decay is observed and if the PMNS variables are well known, the decay rate can give an



(a) Neutrinoless double beta decay of two neutrons. The interaction is mediated by a neutrino  $\nu_m$ . From Reference [120].

(b) The effective Majorana mass as a function of the smallest neutrino mass within the 90% confidence bounds of the oscillation parameters, from Reference [123]. The blue hatched area corresponds to the normal ordering and the red corresponds to the inverted.

indication of the mass hierarchy as well as mass of the smallest neutrino, as shown in Figure 2.10b. Currently there is no experimental evidence for neutrinoless double beta decay [120]. The neutrinoless double beta decay experiments [121] also have sensitivity to the sum of the neutrino masses and the smallest neutrino mass when the PMNS matrix elements are known.

neutrino masses, specifically the effective electron neutrino mass is being probed by considering the kinematics of electrons in single beta decay interactions [66]. The maximum energy the electron can receive from a beta decay interaction is limited to the energy of the emission minus the neutrino mass. By looking at the spectral distortion of the emitted electron energy, experiments have the possibility of being sensitive to the neutrino mass [124]. The current best sensitivity limits for the effective electron neutrino mass neutrino mass is  $m_i < 1.1 \text{ eV}$  (90% confidence level) from the KARlsruhe Tritium Neutrino experiment (KATRIN) [66].

Finally, cosmology can also provide further limits on the sum of the neutrino masses. The Cosmic Microwave Background (CMB) [125] and the Matter Power Spectrum [126], which describes the relative mass density fluctuations of the Universe, are sensitive to the neutrino masses. At early times of the Universe decouple from baryonic matter and free stream, damping small-scale density fluctuations. The lack of fluctuations at small scale reduces the matter power spectrum at small scale as the fluctuations of baryonic and dark matter are reduced. The amount of relativistic matter (also called radiation) can affect the epoch of matter-radiation equality. This change affects the growth rate of the Universe and therefore the size of fluctuations of the CMB in the sky. Assuming three neutrino species

and combining data from the Plank CMB data, along with the affects of Baryon Acoustic Oscillation (BAO) data and gravitational lensing (both not discussed here) results in a limit of  $\sum m_\nu < 0.12$  eV [125].

### 2.2.3 The Seesaw Mechanism

This Section describes the "type 1" seesaw mechanism [127] and follows arguments from References [128, 73]. Although the "type 1" seesaw mechanism provides theoretical motivation for an additional right-handed (sterile) field, to explain the mass differences between neutrinos and their lepton partners the mass of the right-handed field has to be of order  $\sim 10^6$  eV. This is much larger than the eV scale sterile neutrinos that have been suggested to resolve the anomalies in oscillation experiments. However, alternative seesaw models have been suggested to account for the eV scale anomalies, see Reference [44] for a review. It should also be noted that the seesaw mechanism is not the only theoretical mechanism to describe the neutrino mass. The neutrino field could be described by the Dirac mass term only with a smaller Yukawa couple compared to the charged lepton partners. Or one can introduce a new Higgs isospin triplet which couples to the neutrino to form a mass term [129].

[129]

As discussed in Section 2.1, neutrinos in the SM are left-handed with no mass. The masses of the other Fermi fields ( $\psi$ ) are given by the Yukawa couplings ( $\bar{\psi}\phi\psi$ ) to the Higgs field ( $\phi$ ). A left-handed only Yukawa coupling is prohibited in the SM. Therefore, to provide a Dirac mass to the neutrino, a new field  $\bar{v}$  is required so that the new field couples to the neutrino field  $v$  in a new Yukawa coupling. In the unitary gauge, such a coupling can be described by

$$\mathcal{L}_{v \text{ Yuk}} = -\frac{1}{y} \sqrt{2}(v + H)(v\bar{v} + \bar{v}^\dagger v^\dagger), \quad (2.19)$$

where  $y$  is the coupling constant,  $H$  is the real scalar Higgs field, and  $v$  is the vacuum expectation [128]. The mass of the neutrino, in this case, is given by  $m_\nu = yv/\sqrt{2}$ . The mass term,

$$\mathcal{L}_{\bar{v} \text{ mass}} = -\frac{1}{2}M(\bar{v}\bar{v} + \bar{v}^\dagger \bar{v}^\dagger) \quad (2.20)$$

where  $M$  is constant, is also allowed. In this case,  $\bar{v}$  can be thought of as the right-handed component of the neutrino field and is Majorana ( $\bar{v} = v_R = C\bar{v}_L^T$ ). In this case the neutrino is its own anti-particle. Equations 2.19 and 2.20 can be combined to form the the following mass term [128]

$$\mathcal{L}_{v \bar{v} \text{ mass}} = \begin{pmatrix} v & \bar{v} \end{pmatrix} \begin{pmatrix} 0 & m \\ m & M \end{pmatrix} \begin{pmatrix} v \\ \bar{v} \end{pmatrix} + h.c. \quad (2.21)$$

This equation can be used to provide a reason as to why the neutrino mass is smaller than the corresponding partner lepton mass. When assuming  $M \gg m$ , one can diagonalise the matrix which results in eigenvalues of  $M$  for  $\bar{v}$  and  $-m^2/M$  for  $v$ . If  $M$  is assumed to be  $\sim 1$  MeV and the Yukawa coupling is of the same magnitude as the charged leptons, then the equation requires the mass of the neutrino to be  $\sim 1$  ZeV, which is in the region of experimental limits but greater than cosmological constraints.

As described in Reference [128], the new neutrino field can be integrated out and, keeping the leading order term, and the effective Lagrangian can be shown as

$$\mathcal{L}_{v \text{ Yuk mass}} = -\frac{1}{2}m_v(vv + v^\dagger v^\dagger)(1 + H/v)^2. \quad (2.22)$$

This formulation results in the same neutrino mass terms as before. Extending to three generations results in the charge interaction currents gaining a new matrix. This new matrix describes the switch from flavour eigenstates to mass eigenstates, otherwise known as the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix [36], which is described in Section 2.2.1. The new neutrino fields added to the theory can be thought of as sterile as they do not interact via the weak force.

In summary, by introducing a new neutrino field for which Yukawa coupling is still of the same magnitude and where the coupling constant of the new field is much larger than the coupling constant of the neutrino field, one can explain why the mass of the neutrino is comparatively smaller than the corresponding charged lepton partner mass. By applying the above to all three neutrinos, the theory requires the PMNS matrix which describes neutrino oscillations. In addition, the new neutrino field, or fields, can be thought of as sterile as they do not interact via the weak force. Therefore, the theory described in this Section, which is known as the "type 1" seesaw mechanism, justifies the neutrino mass, neutrino oscillations and the potential existence of sterile neutrinos at MeV scales.

#### 2.2.4 Calculating the Number of Neutrinos

Cosmology also provides constraints on the effective number of neutrinos,  $N_{\text{eff}}$  [125].  $N_{\text{eff}}$  describes the degrees of freedom arising from the number of neutrino fields at the time of decoupling. Decoupling is an era in time where mean free path of the particle becomes higher than the horizon of the Universe. Hence the particle able to free stream. It should be noted that due to neutrino decoupling not being instantaneous,  $N_{\text{eff}}$  is not the exact number of neutrinos, but rather the effective number. The CMB and the Matter Power spectrum are sensitive to the number of neutrinos. This is because the total radiation energy density of the

Universe,

$$\rho_{\text{total,rad}} = \rho_\gamma + \rho_\nu = \rho_\gamma + \frac{7}{8} \left( \frac{4}{11} \right)^{4/3} N_{\text{eff}} \rho_\gamma, \quad (2.23)$$

depends on contributions from the neutrino energy density,  $\rho_\nu$ , and photon energy density,  $\rho_\gamma$ . Therefore, at higher values of  $N_{\text{eff}}$ , for a constant  $\rho_{\text{total}}$ , more of the fraction of the radiation density and gravitational force is a result of more evenly distributed free streaming neutrinos. Therefore relative fluctuations are weaker at smaller scales. This can be seen in Figure 2.11, where changes in the effective number of neutrinos visibly affected the CMB power spectrum. Also, the density of neutrinos affects the expansion of the Universe and hence the temperature of the Universe. Therefore the ratio of the photon diffusion length and the sound length is affected. The effect of changing this ratio is that an increase  $N_{\text{eff}}$  causes a loss of fluctuations as photons diffuse across a larger space [8, 130, 131]. A combined fit using CMB data, lensing and BAO data results in  $N_{\text{eff}} = 2.96^{+0.34}_{-0.33}$  [125] which means the theoretical value of  $N_{\text{eff}} = 3.046$  [125] is within the error bounds of the analysis. This result is in agreement with the SM and places cosmology constraints on additional neutrinos outside of the SM.

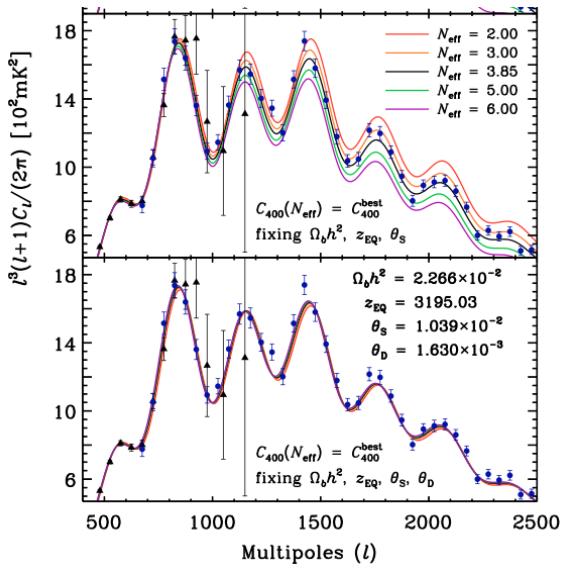


Fig. 2.11 The CMB power spectrum and the effect of increasing  $N_{\text{eff}}$ . The black line corresponds to the best fit of WMAP7 (black points) (The Wilkinson Microwave Anisotropy Probe 7-year result) and SPT K11 (blue points) (South Pole Telescope, Keisler et al, 2011) data. The top graph fixes the matter-radiation equality redshift ( $z_{\text{eq}}$ ), and the angular size of the sound horizon  $\theta_s$  as well as baryon density. The bottom graph also keeps the fraction of baryonic mass in helium,  $Y_p$ , constant in order to fix the angular size of the photon diffusion,  $\theta_D$ . As the difference in the different  $N_{\text{eff}}$  is small, when  $\theta_D$  is fixed, it shows that the impact is related to the diffusion of photons, i.e. Silk damping. From Reference [8]

Accelerator experiments have also determined the number of light neutrinos. Collisions of  $e^+e^-$  at a centre of mass energies around the mass of the Z boson (91 GeV) allows for the production of Z bosons [132, 133]. Approximately 20% of the interactions produce a  $\nu\bar{\nu}$  pair which results in the interaction being invisible [133] to the detector. The contribution to the Z-width,  $\Gamma$ , from the invisible interactions can be calculated indirectly from

$\Gamma_{inv} = \Gamma_{tot} - \Gamma_{vis}$ . The measured cross-section of the Z decays depends on both the partial width of the interaction type and the total width. Because of this, both widths can be extracted from the measured data of the Z decays [132]. The invisible width can also be measured directly when a photon is also present, e.g  $e^+e^- \rightarrow \gamma Z \rightarrow \gamma\nu\bar{\nu}$ . The number of neutrinos can then be extracted by considering the ratio of the invisible width and the lepton width

$$\frac{\Gamma_{inv}}{\Gamma_{ll}} = N_\nu \left( \frac{\Gamma_{vv}}{\Gamma_{ll}} \right)_{SM}, \quad (2.24)$$

where  $(\Gamma_{inv}/\Gamma_{ll})_{SM} = 1.9912 \pm 0.0012$  [133] is the SM prediction of the ratio. The results for the number of neutrino species was found to be  $2.9840 \pm 0.0082$  [133], which is in agreement with the three neutrinos in the SM. It should be noted that a sterile neutrino does not contribute to the effective number of neutrinos in this analysis as sterile neutrinos do not interact via the weak force.

### 2.2.5 Sterile Neutrino Theory

The theoretical justification for a sterile neutrino was discussed in Section 2.2.3. Coupled with the experimental evidence for neutrino oscillations, which require neutrinos to have mass, the addition of a sterile neutrino to current theory is plausible. However, the sterile neutrino typically has a mass much greater than an eV in theoretical models. This Section describes the eV sterile oscillation models used to explain anomalies in recent neutrino results.

Although sterile neutrinos do not interact via the weak force, it is thought that sterile neutrinos could interact with the SM neutrinos via neutrino oscillations. In this case, the PMNS matrix is converted from a  $3 \times 3$  matrix to a  $6 \times 6$ ,

$$U_{3+3} = \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} & U_{e4} & U_{e5} & U_{e6} \\ \dots & & & & & U_{\mu 6} \\ \dots & & & & & U_{\tau 6} \\ \dots & & & & & U_{s_1 6} \\ \dots & & & & & U_{s_2 6} \\ \dots & \dots & \dots & \dots & \dots & U_{s_3 6} \end{pmatrix}. \quad (2.25)$$

This increase is to account for the three flavours of the standard model and the sterile neutrinos,  $s_i$ . Oscillation probabilities are then calculated using the standard procedure. Commonly, only one sterile neutrino (3+1) framework [134] is considered due to its simplicity, and  $U_5$  and  $U_6$  terms are removed. In addition, for a short-baseline analysis the short-baseline approximation is used,  $\Delta m_{solar}^2 = \Delta m_{atmospheric}^2 = 0$  [134]. Therefore, in the case of a muon beam where the detector is  $\sim 1$  km away from the start of the beam, oscillations to the sterile neutrino can be described by the following probabilities [134]:

$$\begin{aligned} P_{\nu_\mu \rightarrow \nu_\mu} &= 1 - 4(1 - |U_{\mu 4}|^2)|U_{\mu 4}|^2 \sin^2(1.26 \Delta m_{41}^2 (\text{eV}^2) L (\text{km})/E (\text{GeV})), \\ P_{\nu_\mu \rightarrow \nu_e} &= 4|U_{e4}|^2|U_{\mu 4}|^2 \sin^2(1.26 \Delta m_{41}^2 (\text{eV}^2) L (\text{km})/E (\text{GeV})). \end{aligned} \quad (2.26)$$

In this case no CP violation is detected and it is also usually assumed that the oscillations are CPT invariant which requires that  $P(\nu_i \rightarrow \nu_j) = P(\bar{\nu}_j \rightarrow \bar{\nu}_i)$ .

When the analysis is operating in a 2D framework, the matrix elements are usually replaced with the following:

$$\begin{aligned} \sin^2 2\theta_{\mu\mu} &= (1 - |U_{\mu 4}|^2)|U_{\mu 4}|^2, \\ \sin^2 2\theta_{e\mu} &= 4|U_{e4}|^2|U_{\mu 4}|^2, \end{aligned}$$

which are used as fit parameters in data analysis. A similar procedure is performed for the 3 + 2 and the 3 + 3 sterile cases resulting in more complex definitions for the oscillation probabilities as well as more oscillation parameters.

## 2.2.6 Anomalies in Neutrino Experiments

Several neutrino experiments have attempted to identify sterile signatures within data. Some experiments have seen indications of sterile physics. However, tensions in global fits and null results from other experiments have placed doubt on the results [135]. This Section describes the short-baseline neutrino beam experiments, the Liquid Scintillator Neutrino Detector (LSND) [9] and MiniBooNE [136] (Mini-Booster Neutrino Experiment), and discusses

the low energy excess observed in these experiments. In addition, other searches will be discussed, such as the reactor limits discussed in Section 2.2.1, and the "Gallium Anomaly".

LSND [9] and MiniBooNE [136] were both neutrino experiments in low energy neutrino beams, positioned a short distance from the start of the beam. Both are filled with mineral oil [136, 137] and surrounded by PMTs to detect scintillation and Cherenkov light. Using the Cherenkov rings and the timing information from the prompt scintillation allows for directional and calorimetric reconstruction.

From 1993 to 1998 LSND was situated 30 m downstream of the LANSCE (Los Alamos Neutron Science Center) accelerator beam at Los Alamos [137]. The beam provided a source of  $\nu_\mu$  and  $\bar{\nu}_\mu$  in the  $36 < E_\nu < 58.2$  MeV range. Neutrinos in the beam were created mostly from the decay at rest of  $\pi^+ \rightarrow \nu_\mu \mu^+$  where the  $\mu^+$  decayed at rest via  $\mu^+ \rightarrow \bar{\nu}_\mu \nu_e e^+$  [138]. However, there also existed decay in-flight modes such as  $\pi^+ \rightarrow e^+ \nu_e$  which, were backgrounds for the  $\nu_\mu \rightarrow \nu_e$  analysis. A contamination of  $\bar{\nu}_e$  also existed in the beam and was the primary background for the  $\bar{\nu}_e$  oscillation search. An additional background for the  $\nu_e$  and  $\bar{\nu}_e$  searches was the misidentification of  $\nu_\mu$ . This background was reduced as events were taken within an energy range of 20 MeV to 60 MeV, which is below the muon production threshold [138]. Additional backgrounds also include cosmic events, but, a veto existed around the inner detector to reduce this background [138, 137]. For  $\nu_e$  and  $\bar{\nu}_e$  appearance analysis, the electron neutrinos and anti-neutrinos interacted via beta decay and inverse beta decay. The detector was also doped with a liquid scintillator to allow neutron capture on the scintillator and hence produce a 2.2 MeV photon [39]. This allowed beta decay  $\nu_e$  interactions to be separated from the  $\bar{\nu}_e$  interactions.

An appearance search showed an excess of  $\bar{\nu}_e$  from the  $\bar{\nu}_\mu$ , indicating oscillations. Assuming a two neutrino oscillation fit, the expected number of  $\bar{\nu}_e$  from  $\mu^-$  decay at rest was  $19.5 \pm 3.9$  and the number of events from misidentification of  $\pi^-$  decay in-flight events was  $10.5 \pm 4.6$ . The LSND experiment measured a total of  $117.9 \pm 22.4$   $\bar{\nu}_e$  events resulting in an excess of  $87.9 \pm 22.4 \pm 6.0$  [9], corresponding to a  $3.8\sigma$  excess. This excess is shown in Figure 2.12 where the number of events with the beam on minus the beam off is shown. The Figure shows that the data agrees well with predicted results for the inclusion of a sterile neutrino. The events due to the presence of a sterile neutrino are shown in blue in the Figure. If the excess is interpreted as oscillations and is not statistical, the result can be accounted for by one or more sterile neutrinos [139]. An excess was also seen in a  $\nu_\mu \rightarrow \nu_e$  analysis within the energy range  $60 < E_e < 200$  MeV with an excess of  $18.1 \pm 6.6 \pm 4.6$  events [140].

MiniBooNE then investigated the LSND anomaly at the Booster Neutrino Beam [141]. The beam has a neutrino mode and an anti-neutrino mode, allowing for analysis of oscillations from both  $\bar{\nu}_\mu$  and  $\nu_\mu$ . More information about the BNB can be found in Chapter 4, Section

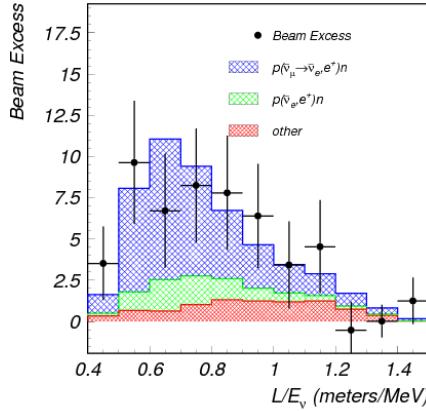


Fig. 2.12 The LSND excess results in terms of distance travelled in meters divided by the neutrino energy is determined from the measured positron energy and angle with respect to the neutrino beam, from Reference [9].

4.1. The signal events were  $\nu_e$  that predominately underwent charge current quasi-elastic interactions, see Section 3.6. Intrinsic  $\nu_e$  and resonant  $\nu_\mu$  events where a photon is released are the predominant backgrounds in the oscillation searches for MiniBooNE. Photons can be released via the decay of  $\Delta \rightarrow N\gamma$  [142] and from the decay of  $\pi^0 \rightarrow \gamma\gamma$ . If photons travel in the same direction in  $\pi^0$  decay this mimics a single photon which is mistaken for an  $e^{+-}$ . Additional backgrounds occur from external sources; however, a cosmic veto region surrounded the detector to minimise the background.

The energy of the signal events was approximated assuming the quasi-elastic kinematics which required the energy of the electron and the angle of the electron to be reconstructed. The measured number of events is currently 2437 for both neutrino and anti-neutrino mode. This corresponds to an excess of  $460.5 \pm 99.0$  events, resulting in a  $4.7\sigma$  effect, see Figure 2.13 [142]. The discrepancy between the excesses of the neutrino and anti-neutrino mode places tension on the 3+1 model; however, the data can be explained using a two sterile neutrino model [138, 143].

Excesses in neutrino flux rates have also been detected elsewhere. The anomalies in reactor experiments [42] have already been mentioned in Section 2.2.1 and are discussed in further detail in Reference [12].

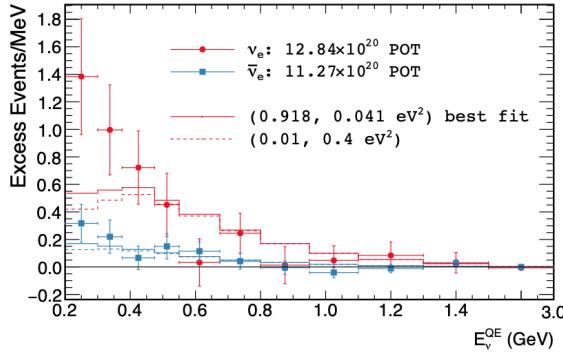


Fig. 2.13 The MiniBooNE excess from both neutrino and anti-neutrino mode data in terms of neutrino energy assuming quasi-elastic scattering, from Reference [10].

Furthermore, the gallium experiments SAGE [40] and GALLEX [41] have measured rates less than predicted from radioactive neutrino sources. The gallium (Ga) detectors identify inverse beta decay interactions upon Ga atoms by counting the resultant  ${}^{71}\text{Ge}$  atoms. The number of  ${}^{71}\text{Ge}$  atoms is counted by allowing the  ${}^{71}\text{Ge}$  to radioactively decay within a proportional counter which detects the corresponding X-rays.  ${}^{51}\text{Cr}$  and  ${}^{37}\text{Ar}$  neutrino sources were introduced to the SAGE detector and  ${}^{51}\text{Cr}$  to the GALLEX detector for calibration. A weighted average of the ratio between the expected and measured neutrino rates from the sources was found to be  $0.86 \pm 0.05$  [41] corresponding to a discrepancy greater than two sigma. In both the reactor and gallium experiments, it is thought that a likely cause of this discrepancy is overestimation of the production rates of neutrinos from the neutrino source [41, 144]. However, the lower rate could correspond to electron neutrino disappearance, which could be due to the existence of sterile neutrinos.

Several other experiments have also performed sterile searches in various forms and placed constraints on the parameter space. Experiments usually try to identify excesses or deficits in the neutrino spectra assuming the standard 3+1 model using the sterile prescription discussed in Section 2.2.5. The constraints can be seen in Figures 2.14, 2.15 and 2.16b. Depending on the experimental setup, different experiments are sensitive to either disappearance or appearance channels. For  $\nu_\mu$  disappearance MINOS+ [145], NOvA [146], Super-K (ATM) [147] and IceCube [148, 149] place constraints on the parameter space, see Figure 2.14. For  $\nu_e$  appearance KARMEN [150] (KARlsruhe Rutherford Medium Energy Neutrino experiment), NOMAD [151] (Neutrino Oscillation MAGnetic Detector), OPERA [152], BNL-E776 [153] (Brookhaven National Laboratory Experiment) and ICARUS [23] place constraints on the sterile oscillation parameters. The contours that describe the sensitivities of these experiments in terms of the 3+1 sterile parameters are shown in Figure 2.16b [12]. Global best fit results of the appearance and disappearance parameters show tension. The

tension can be seen in Figure 2.14. This combined analysis is performed in the 3+1 framework, see Section 2.2.5, and therefore does not dismiss the anomalies completely.

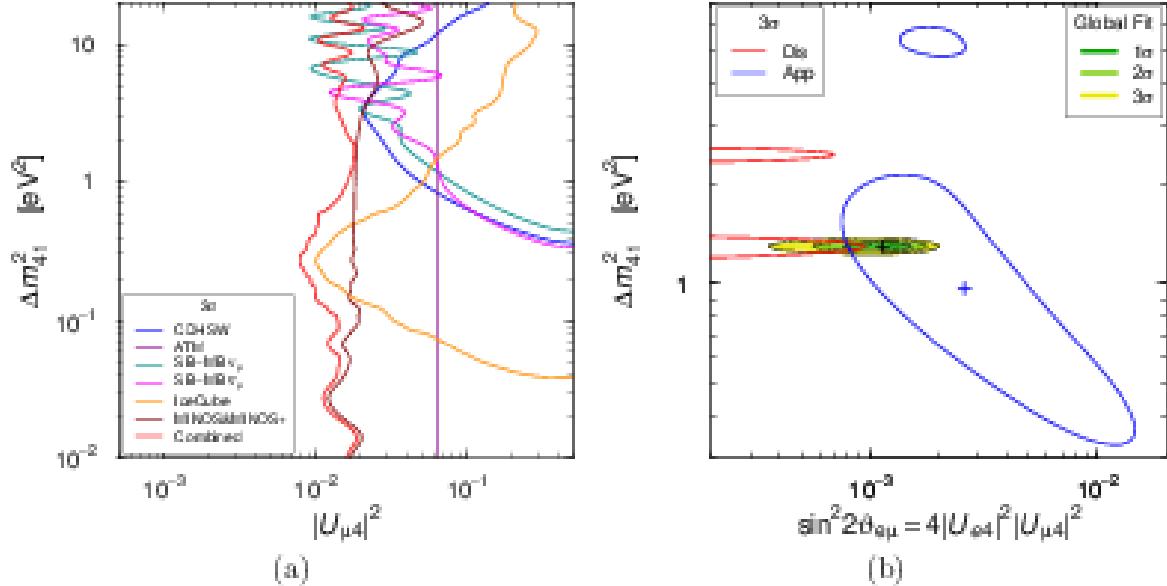


Fig. 2.14 a)  $3\sigma$  sensitivity contours for  $\nu_\mu$  disappearance from experiments. b) Global fit of appearance and disappearance data. The Dis line exclude regions by  $3\sigma$  whilst App corresponds to  $3\sigma$  allowed regions. From Reference [11].

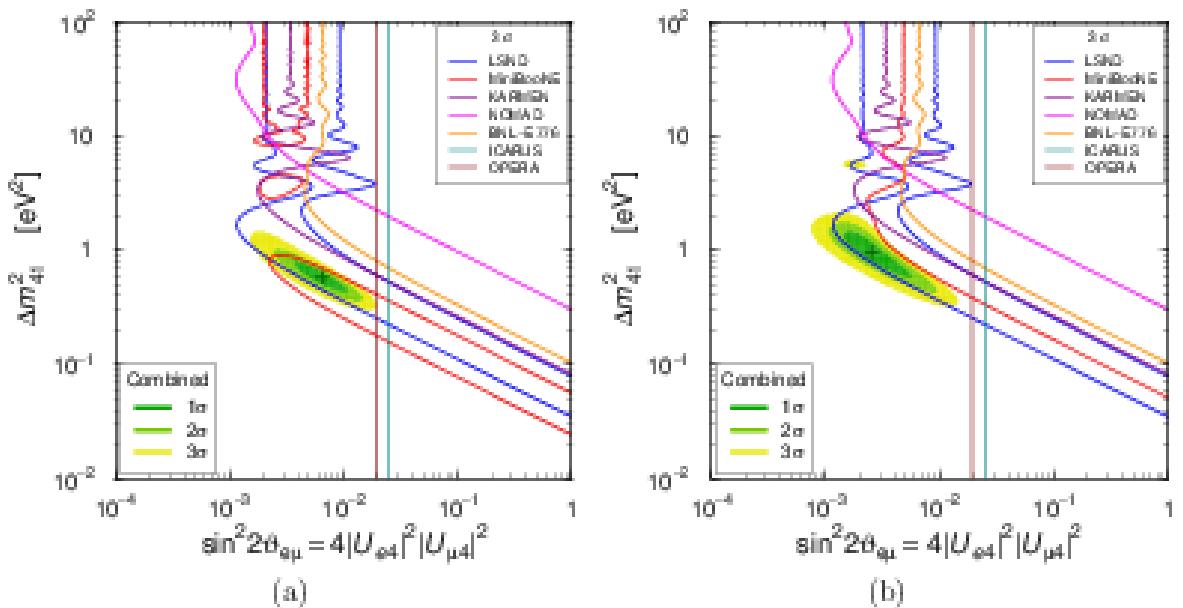


Fig. 2.15  $3\sigma$  sensitivity contours for  $\nu_e$  appearance from experiments. a) is with all MiniBooNE data and b) is without low energy MiniBooNE data. From Reference [11]

Constraints for  $\nu_e$  disappearance are set by solar neutrino data as well as bounds from Super-K, DeepCore and IceCube [11, 12, 135]. Figure 2.16a shows the allowed regions from reactor experiments with a 95% confidence regions bounds and the best fit estimates. Figure 2.16b shows the best fit from the reactor experiments [135] as well as the preferred region from the gallium experiments and the exclusion curves from Super-K, DeepCore and IceCube [11, 12].

In conclusion, several experimental anomalies have been observed resulting in possible evidence for sterile neutrinos, most notable are the anomalies observed LSND and Mini-BooNE. Therefore, to confirm or refute the anomalies discussed in this Section, further experimental results are required. Due to the mass of the  $\nu_e$  flavour state being a superposition of the mass eigenstates, searches in neutrinoless double beta decay and direct mass searches for the effects of sterile neutrinos are possible.

Additional searches are also planned in reactor experiments, J-PARC (Japan Proton Accelerator Research Complex) Sterile Neutrino Search at J-PARC Spallation Neutron Source (JSNS2) experiment [154] and the SBN programme JSNS2 will directly test the LSND anomaly by using the same neutrino source, target material, baseline and neutrino oscillation channel as LSND, with modern detection methods.

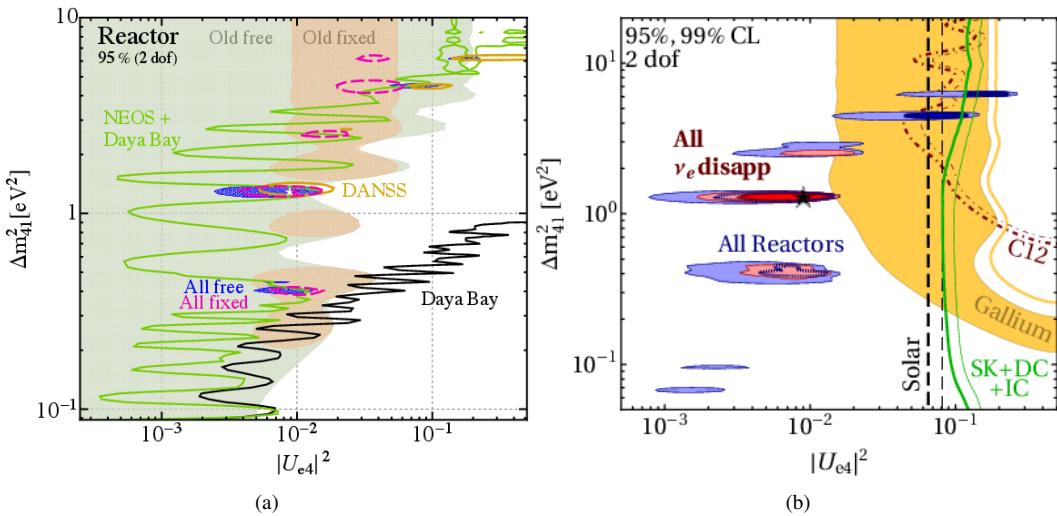


Fig. 2.16 a) Allowed regions from reactor experiments. The shaded areas correspond to the allowed regions prior to reactor experiment results from Daya Bay, DANSS and NEOS. The green area corresponds to setting flux spectra to the predicated values, whilst orange allows the flux to be a free parameter. The blue area corresponds to all reactor data with the flux being free and the magenta contours indicate the global fixed analysis. The star represents the best fit. b) Is the reactor global fits overlapped with the gallium allowed region and the exclusion (to the left) from solar neutrino data and atmospheric data from Super-K, DeepCore and IceCube at 95% and 99% CL. Figures are from Reference [12].

Another search is the MicroBooNE [13] experiment which is a Liquid Argon Time Projection Chamber (LArTPC) detector situated on the BNB next to MiniBooNE. By taking advantage of the high calorimetric and topological resolution of LArTPCs, MicroBooNE is capable of identifying and separating significant photon backgrounds from a  $\nu_e$  oscillation signal [22]. This cannot be performed in MiniBooNE as it cannot discriminate between photon and electron showers [22]. The experiment has been running since 2015 and is described in further detail in Chapter 4. The focus of this thesis is the SBN programme [17], which uses three LArTPCs (including MicroBooNE) as detectors in the BNB, see Chapter 4. The next Chapter will discuss the LArTPC technology and the interactions expected from neutrinos with energies corresponding to the BNB.

## 2.3 Concluding Remarks

This Chapter has shown how developments throughout the last century have resulted in the current SM prescription of neutrinos. The SM describes neutrinos as weakly interacting with no mass and no right-handed counterpart. However, experimental evidence has proven neutrino oscillation to exist, which require the neutrinos to have mass. Experimental limits show this mass to be small with respect to the other leptons and this difference is currently unexplained. A theoretical justification for this mass difference is the seesaw mechanism which requires the addition of a new neutrino field which does not interact, i.e. a sterile neutrino. Recently, experimental searches to determine the properties of the PMNS, which describe neutrino oscillation, have resulted in anomalous results which could be explained by one or more sterile neutrinos. However, tensions in the global fits cast doubts on this explanation. A discussion of the current outlook of sterile neutrino theory and the experimental searches for sterile neutrinos has been presented in this Chapter.

# Chapter 3

## Liquid Argon Time Projection Chamber Theory

As mentioned Chapter 2, Liquid Argon Time Projection Chambers (LArTPCs) are to be used by the Short-Baseline Neutrino (SBN) programme to search for sterile neutrinos. C. Rubbia originally suggested the LArTPC in the 1970s [155]. LArTPCs were suggested for several reasons as discussed in Reference [155]: LArTPCs provide a dense detection medium, hence there is an increase in the rate of neutrino interactions compared to a gaseous TPC. Argon is also cheap which mean large detectors (and more neutrino interactions) are possible. Also, argon is inert and has a high electron mobility. This means that ionised electrons from charge particles interacting in the argon can travel large distances without loss of calorimetric information. The inert nature also means that argon can be easily and cheaply liquefied using liquid nitrogen. Finally, as argon is inert and as the ionisation energy is relatively low LArTPC are highly precise and have low energy threshold. Along with high spatial resolution of the TPC readout several individual charged particles can be topologically and calorimetrically reconstructed and identified. By identifying the outgoing charged particles, neutrino interactions can then be identified and the neutrino flavour resolved. Because of these advantages several LArTPCs have been designed and used in experiments in various fields of research in physics [156]. However, LArTPCs do have challenges. One of the main issues in the technology is the engineering of cryostat vessels large enough to store the argon. This means LArTPC experiments are significantly smaller (hence less neutrino interactions) than the water Cerenkov neutrino experiments such at Hyper-K. Similarly there are challenges with extracting the large number of readout channels out of the cryostat. Switching from the traditional wire readout to a pixelised readout is one possible solution and is discussed in Chapter 9. Furthermore, to maintain the low thresholds of LArTPCs the

electronic noise has to be minimised and well understood. Discussions on this topic can be found in Chapter 5.

This Chapter will discuss the design of LArTPCs in Section 3.1. Then the processes that describe how charged particles are detected will be discussed. The first stage is discussed in Section 3.2, which involves the ionisation of liquid argon and the recombination of these ionised particles. Then the processes which affect the resultant ionisation electrons during transport to the readout will be discussed in Section 3.3. Finally, the form by which energy is deposited and how the ionisation electrons are used to identify particles will be discussed in Section 3.4.

Although signal-phase wire TPC technology will be the main focus, other technologies will be briefly discussed in Section 3.5.

Of particular interest to the SBN programme are neutrino interactions upon argon that occur at a specific energy range. These interactions and the mechanisms concerning them will also be discussed in Section 3.6.

## 3.1 LArTPC Design

As can be seen in Figure 3.1, a LArTPC consists of a volume of liquid argon sandwiched between a cathode and an anode. Usually, the anode consists of two or more planes of sense wires that are spaced 3-5 mm apart [17, 24, 115]. When charged particles travel through the detector, they ionise the liquid argon. The electric field supplied by the cathode, typically 500 V/cm, forces the ionised particles to drift between the two electrodes. As argon is a noble gas, the ionised electrons can freely travel to the anode where they are detected. The front-facing planes of the anode shield the ones behind. Because of this geometry, a voltage bias has to be applied to allow the front planes to be transparent [157]. The final plane collects the drift electrons and therefore has a different response to the charge deposition, see Chapter 6 for more detail.

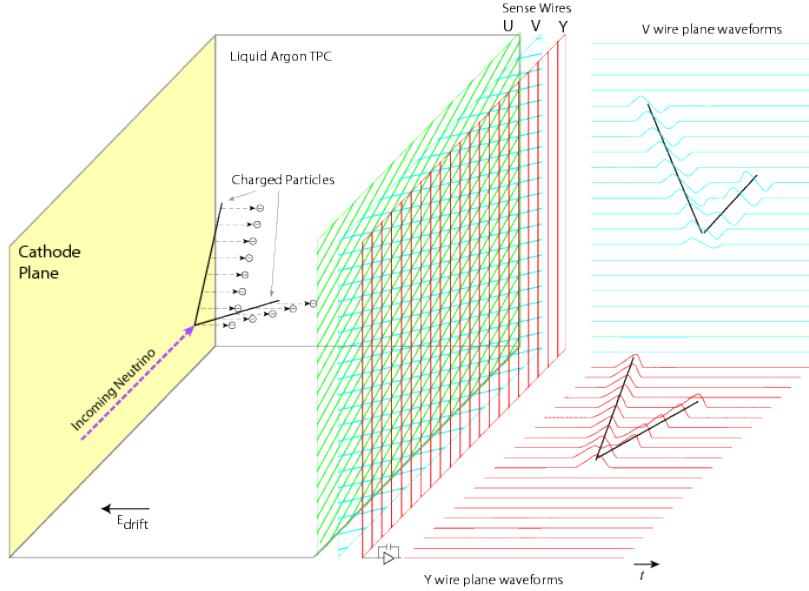


Fig. 3.1 The LArTPC design. The energy deposits travel toward the readout which is made up of wire planes for 2D reconstruction. Figure from Reference [13].

The electronic response to the energy deposits is then amplified and shaped within the cryostat before the data is processed by a digital acquisition system [157]. Placing the amplifiers within the liquid argon has the advantage of reducing the inherent noise in the amplifiers and reducing the number of channels exiting the cryostat [158].

The electronics response on the wires can then be combined and using the drift time and the wire position in a plane allows for 2D reconstruction of the energy deposits. As the planes are orientated in different directions, combining the results from planes allows for 3D reconstruction of the ionisation depositions. The 2D representation is shown in Figure 3.1.

Current large scale LArTPCs, for example MicroBooNE [13] and ICARUS [159], consist of  $\mathcal{O}10^4$  sense wires across 3 planes with a drift length of  $\mathcal{O}1$  m. The volume encompassed by the drift length and wires requires  $\mathcal{O}100$  tonnes of liquid argon to encapsulate this readout.

To provide timing information for the particles and to detect scintillation light, light detectors are usually placed in a LArTPC, typically behind the anode. PhotoMultiplier Tubes (PMTs) are currently the most common; however, for future detectors, such as DUNE, light traps which use Silicon PhotoMultipliers, called ARAPUCAs (Argon R&D Advanced Program at UniCAMP), are being developed, see Chapter 4. The scintillation light is in the Ultra Violet (UV) range [160]; therefore the light detection systems and reflector foils are coated with wavelength shifting materials so that standard PMTs can detect the light

[17, 161]. Further information on ionisation and scintillation light can be found in Section 3.2.

During the transportation to the anode, electrons from ionisation depositions diffuse and impurities diminish the signal strength. A purification system is used such that the loss due to impurities is minimal. Section 3.3 will discuss transportation and the effects which reduce the signal strength in more detail.

## 3.2 Ionisation, Recombination and Scintillation

As the ionisation and scintillation light is key to particle detection in LArTPCs, the amount of energy required to produce an electron/ion pair in liquid argon is a critical value. The value of this is approximately  $W_{ion} = 23.6$  eV [29, 161]. After a charged particle ionises argon atoms the electrons and ions can then recombine. The fraction of electrons which recombine depends on the electric field strength and the amount of localised ionisation. Therefore the recombination also depends on the energy deposited per unit distance,  $dE/dx$ , of the passing particle. The recombination has been theoretically modelled by Onsager [162], Jaffé [163] and Kramers [164] as well as Thomas and Imel [165][29, 166]. The resultant theory models the localised charge density as uniform and in a small box localised at the interaction point. The box model defines the remaining charge after recombination to be

$$Q = Q_0 \frac{1}{\eta} \ln(1 + \eta), \quad (3.1)$$

where  $\eta = \alpha Q_0/E$ ,  $Q_0$  is the initial charge,  $E$  is the electric field and  $\alpha$  is a fit parameter which describes the charge density and charge mobility. For a constant electric field  $\eta = \beta(dE/dx)$ , where  $\beta$  is a fit parameter .

An alternative approach to define the remaining charge is Birk's law [167],

$$Q = \frac{Q_0}{1 + k_B(dE/dx)/E}, \quad (3.2)$$

which arises from empirical fits to data [166]:  $k_B$  is a fit parameter. The recombination factor,  $\mathcal{R} = Q/Q_0$ , describes the conversion from energy deposition to charge after recombination via the equation

$$dE/dx = \frac{dQ/dx}{\mathcal{R} W_{ion}}. \quad (3.3)$$

One can then substitute the recombination parameter defined by Equations 3.1 and 3.2 and rearrange them to find a conversion between the  $dE/dx$  deposited and the  $dQ/dx$  after

recombination. Results from ICARUS [166] suggested modifications to the recombination parameters described in Equations 3.1 and 3.2 such that

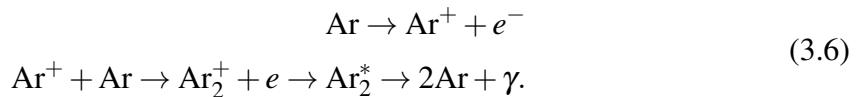
$$dE/dx = \frac{dQ/dx}{A_B/W_{ion} - k_B(dQ/dx)/E}, \quad (3.4)$$

for Birk's model and

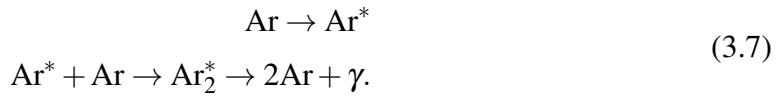
$$dE/dx = (\exp(\beta W_{ion}(dQ/dx)) - \alpha)/\beta \quad (3.5)$$

for the box model, now known as the modified box model [29]. In both cases, an extra fit parameter,  $A_B$  for Birk's model and  $\alpha$  for the box model has been added. Note that the modified box model is usually favoured as it does not suffer from high  $dE/dx$  values like Birk's model when the denominator is zero. For the modified box model the parameters  $\alpha = 0.930$  and  $\beta = 0.212/\rho E$  (MeV/cm) $^{-1}$ , where the density of the liquid used is  $\rho = 1.3954$  g cm $^{-3}$  [168]. For Birk's model the parameters  $A_B = 0.8$  cm and  $k_B = 0.0486/\rho$  (kV/cm)(MeV/cm) $^{-1}$ (cm) $^{-1}$  are used [168].  $dQ/dx$  in this case is the number of electrons per centimetre.

Approximately, the average fraction of electrons which recombine with the argon ion is 0.58 for minimum ionising particles in a 500 V/cm drift field. When those electrons recombine, scintillation light is produced through the formation of an excimer, Ar $_2^*$ , and then the deexcitation of the excimer. There are two pathways for this [160]:



Alternatively, a third pathway is the excitation of the argon atom by the through going particle



The resultant scintillation light has a wavelength of approximately 128 nm [160]. There are two excited states for the excimer: a singlet and triplet state with decay times of 5 ns and 1  $\mu$ s respectively. The relative yields are 23% (singlet) and 77% (triplet) [160]. Due to the separation length of the two particles in the excimer being shorter than the standard atoms in the argon, the energy of the photon released is lower than the required energy to excite further argon atoms. Therefore, liquid argon is transparent to its scintillation light.

Cherenkov light is also emitted as particles propagate through the argon [169]. For visible Cherenkov light, the refractive index of liquid argon is approximately constant (1.22). Therefore, for ultra-relativistic particles, the cone half-angle is approximately  $34^\circ$ . However, the number of photons produced is significantly less than the scintillation light.

### 3.3 Transportation of Ionised Electrons, Diffusion and Impurities

The electrons that do not recombine begin to move towards the anode. As they travel, the ionisation particles can diffuse diminishing the peak signal size. Furthermore, electronegative impurities (such as oxygen, water and nitrogen molecules which have not been removed from the argon) can absorb ionisation electrons further diminishing the signal strength. These effects depend upon the time it takes for the ionisation electrons to reach the anode. Therefore, the drift velocity is crucial for transportation.

The drift velocity  $v = d/t = \mu E$  (where  $d$  is the drift distance,  $t$  is the drift time,  $E$  is the electric field, and  $\mu$  is the electron mobility) has been studied and found to be  $0.1648 \text{ cm}/\mu\text{s}$  for  $E = 500 \text{ V/cm}$  [170, 171]. The drift velocity has also been modelled as a function of the electric field strength and the temperature [172]. In LArTPCs it is preferred that the drift velocity is as large as possible. This is to reduce the time required to drift electrons and thus reduce the effects of diffusion and impurities discussed below. It is also preferred that the velocity is constant throughout the detector to prevent distortion of particle trajectories. However space charge, which is discussed below, can distort the electric field and thus the drift velocity.

The electron diffusion is parameterised by the evolution of the electron cloud which is described by Fick's equation

$$\frac{\partial n}{\partial t} = D_L \frac{\partial^2 n}{\partial x^2} + D_T \left( \frac{\partial^2 n}{\partial x^2} + \frac{\partial^2 n}{\partial y^2} \right) - v \frac{\partial n}{\partial z} - \lambda v n, \quad (3.8)$$

where  $n$  is the electron cloud charge density,  $v$  is the drift velocity,  $\lambda$  is the ionisation coefficient and  $D_L$  and  $D_T$  are the longitudinal and transverse diffusion constants respectively [173]. The solution to this function for a point source in a constant electric field is a Gaussian function that depends on the coordinates  $z$  (drift direction) and  $\rho = x^2 + y^2$ ,

$$n = \frac{n_0}{4\pi D_T t \sqrt{4\pi D_L t}} \exp\left(-\frac{(z-vt)^2}{4D_L t}\right) \exp\left(-\frac{\rho^2}{4D_T t}\right), \quad (3.9)$$

where  $n_0$  is the original density at  $t = 0$ . Therefore the resolution of the detector is smeared for increasing drift distance. As is described in Reference [173], there is a relation between  $D_L$  and  $D_T$ :

$$\frac{D_L}{D_T} = 1 + \frac{E}{\mu} \frac{\partial \mu}{\partial E} \quad (3.10)$$

Measurements of the diffusion coefficients at BNL have been performed at 500 V/cm and have been found to be  $D_L = 7.2 \text{ cm}^2\text{s}^{-1}$  and  $D_T = 7.2 \text{ cm}^2\text{s}^{-1}$  [173]. This corresponds to a resolution limit of a couple of millimetres ( $\sim 1 \mu\text{s}$  in drift time) for the typical drift lengths of large LArTPCs. It should be noted that diffusion can diminish the peak signal size due to smearing. However, at this scale, where the shaping time and wire pitch are comparatively large, the diffusion effect is negligible.

The signal is further diminished due to electronegative impurities in the liquid argon, such as oxygen, nitrogen and water. The rate of electron removal from a charge cloud is modelled by [174],

$$\frac{dN_e}{dt} = -k_{tot}N_e(t), \quad (3.11)$$

where  $N_e$  is the number of electrons and  $k_{tot}$  is the probability of an electron being attached to an impurity. Integrating and describing the equation in terms of charge rather than the number of electrons and dividing by the distance travelled between two wires, the track pitch  $dx$ , results in the decay law [174],

$$\frac{dQ}{dx} = \frac{dQ_0}{dx} e^{-\frac{x}{\tau}}, \quad (3.12)$$

where  $\tau = \frac{1}{k_{tot}}$  is the electron lifetime. The electron lifetime is usually the measure given to identify the purity of the liquid argon. Experiments have seen lifetimes ranging from 0.5-10 ms [175, 176][177, 178]. The SBND simulation currently sets the lifetime at 3 ms which is based on results from the 35-tonne experiment [179]. This a conservative value compared to other similar sized detectors [178, 180]. The lifetime is usually measured using Equation 3.12 to model the most probable charge deposited by cosmic muons at various drift times in the detector. For more detail on lifetime calculation methods, see Chapter 5, Section 5.5.

Space charge is another detector effect which alters the propagation of drift electrons to the anode. Space charge is caused by the build-up of the positively charged ions from the argon ionisation, which slowly move to the cathode. These positively charged ions locally distort the electric field, which affects the recombination process and the drifting of the electrons. Space charge is particularly an issue for surface detectors where the high cosmic flux causes high amounts of ionisation [181]. Assuming the E field depends only on the drift

direction and the space charge does not evolve with time, one can take a modified continuity equation to describe the evolution of space charge [182],

$$\nabla J + \frac{\partial \rho}{\partial t} = K, \quad (3.13)$$

where  $J = \rho v$  describes the current density,  $\rho$  is the charge density,  $v$  is the velocity and  $K$  is the volume rate of creation of ion pairs. Assuming the space charge density is time independent, and ionisation travels only in the drift direction,  $z$ , this can be simplified to

$$\frac{d\rho v}{dz} = K. \quad (3.14)$$

Setting the cathode to be at  $z = D$ , the anode to be at  $z = 0$  and using the fact  $\mathbf{v} = \mu \mathbf{E}$ , where  $\mathbf{E}$  is the electric field, and  $\mu$  is the electron mobility, one can calculate the local electric field in the presence of space charge as

$$E(x) = E_0 \sqrt{\left(\frac{E_A}{E_0}\right)^2 + \alpha^2 \frac{z^2}{D^2}}. \quad (3.15)$$

$E_0$  is the nominal field,  $E_A$  is the field at the anode and  $\alpha = D/E_0 \sqrt{K/\epsilon\mu}$ , where  $\epsilon$  is the dielectric constant for the medium [182].

Space charge is measured using external calibration systems such as UV lasers or cosmic muons with known trajectories. A Fourier series is then used to map the space charge [181]. The map can then be used to spatially correct the positions of particles in the TPC as well as identify the localised electric field so that recombination is correctly determined. Note: in the analyses in Chapters 7 and 8, space charge is not simulated. However, recent efforts have been made to implement a space charge simulation and correct for it in reconstructions carried out in the SBN experiments SBND and ICARUS.

### 3.4 Particle Calorimetry and Identification

The resultant ionisation from charged particles can be used to identify the type of particles and the energy depositions within the TPC. Energy depositions in LArTPCs are described by the Bethe-Bloch Equation [14]:

$$-\langle \frac{\Delta E}{\Delta x} \rangle = \frac{4\pi N_A e^4 Z}{m_e c^2 A} \rho^2 \frac{1}{\beta^2} \left[ \frac{1}{2} \ln \frac{2m_e c^2 \beta^2 \gamma^2 T_{\max}}{I^2} - \beta^2 \right], \quad (3.16)$$

where  $N_A$  is Avogadro's number,  $A$  and  $Z$  are the mass and atomic number of the absorber respectively,  $z$  is the charge of the incoming particle,  $\rho$  is the density of the absorber,  $T_{max}$  is the maximum kinetic energy that can be transferred and  $I$  is the mean excitation energy of the absorber. The density and shell corrections have been omitted. In addition,  $1/4\pi\varepsilon_0 = 1$  and  $z = 1$ . The energy deposited within a small segment (or sense wire) by a particle, which follows the Bethe-Bloch Equation above, has been shown [14] to have the form of a Landau probability distribution (see Figure 3.2),

$$\phi(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} e^{s \ln s + xs} ds, \quad (3.17)$$

where  $c$  is an arbitrary real number.  $x$  depends on the energy lost within the segment and the properties of the detector.

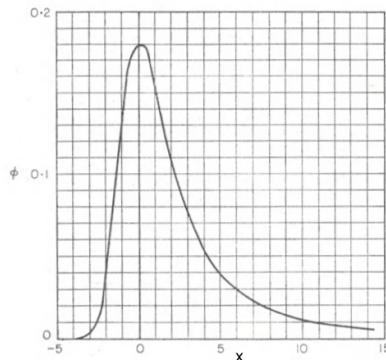


Fig. 3.2 A Landau distribution. The long asymmetric tail predicts higher energy electrons, known as delta rays, to exist within the TPC. These have been observed. The variable  $x$  depends on the energy deposited within a small segment and properties of the detector. From Reference [14].

The larger the segment, the more scattering occurs and hence, the more Gaussian-like the distribution [14]. Therefore, collections of charge deposition on sense wires can be well modelled by a Landau-Gaussian convolution, see Chapter 5, Section 5.5 for an example of this.

Once charge depositions have been collected and reconstructed on the sense wires, particles can be identified from the topology of the depositions in the detector and the calorimetry of the particles. For example, electromagnetic particles ( $e^\pm\gamma$ ) have shower-like energy depositions in the detector when the mother particle has sufficient energy, see Figure 3.3a. However, muons, protons and pions create track-like energy depositions in the detector, see Figure 3.3b.

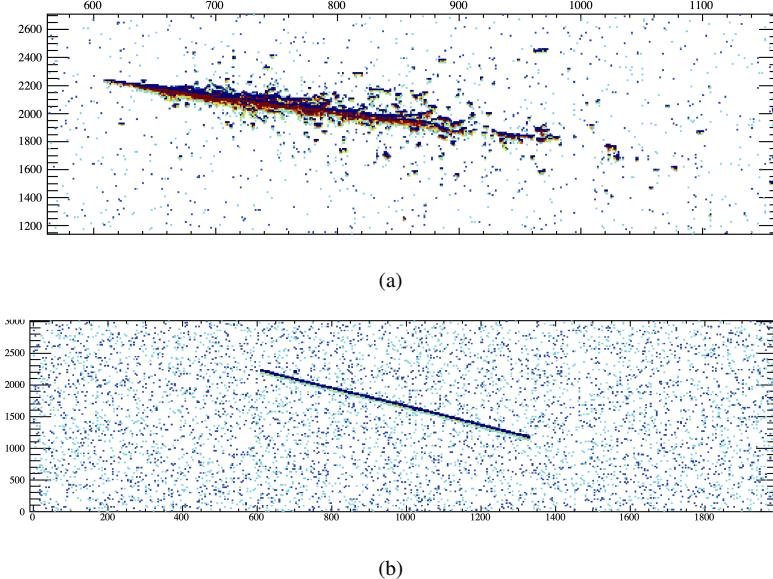


Fig. 3.3 Simulation of a 2 GeV electron (a) and muon (b) in a LArTPC. The image is the raw digitised output from a wire plane with wire number on the x-axis and time in digitised ticks on the y-axis.

### 3.5 Other LArTPC Developments

Single-phase LArTPCs with wire plane readouts are not the only technology used in LArTPCs. Alternative readout methods, such as a pixel readout [33, 183] described in Chapter 9, remove the need for delicate, fragile wires. Pixels also provide immediate 3D reconstruction removing the need for projection matching. However, this comes at the cost of requiring more channels  $N \rightarrow N^2$ . As the number of channels is already large in standard wire LArTPCs, multiplexing techniques are required which introduce ambiguities in the reconstruction.

An alternative readout to the single-phase LArTPC is the dual-phase LArTPC such as ProtoDUNE dual-phase detector[184, 185]. These employ an extraction grid which passes the drift electrons into an argon gas region. The gas is then passed through Large Electron Multipliers which amplify the signal via an avalanche effect. This has been shown to provide a gain of  $\sim 15$ . The readout anode then consists of a multi-layer PCB anode.

### 3.6 Neutrino Interactions and Cross-Sections

LArTPCs can provide precise topological and calorimetric information for the identification of particles and thus neutrino interactions. One other component to perform oscillation

physics is the knowledge of how many neutrino interactions are expected. Therefore, detailed cross-section models are required. Neutrino interactions between  $\sim 0.1 - 10$  GeV are of relevance for this thesis and will be briefly discussed in this Section.

There are four main categories for interactions in this energy range: quasi-elastic scattering, resonance productions, coherent pion production and deep inelastic scattering. Figure 3.4 shows the contribution from some of the sub-categories to the neutrino and anti-neutrino cross-sections. The interaction can either be a charge current interaction, where the product of the interaction is the corresponding charged lepton partner of, or a neutral current interaction, where the product is a neutrino of the same flavour. In the case of neutral current interactions, the energy of the neutrino cannot be fully resolved because the outgoing neutrino is not detected within LArTPCs.

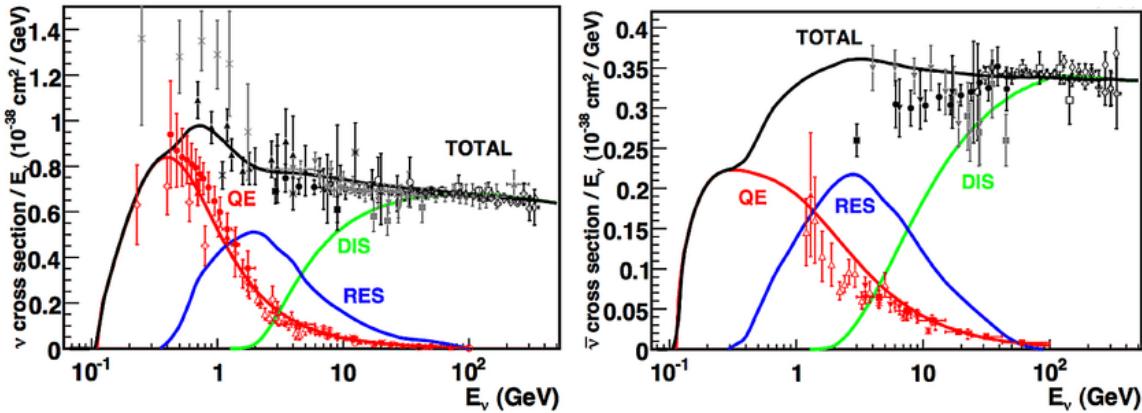


Fig. 3.4 The neutrino (a) and anti-neutrino (b) cross-section measurements as a function of neutrino energy. The prediction for the Quasi-Elastic (QE), Resonance (RES) and Deep Inelastic Scattering (DIS) are given by the Nuance [15] generator. Figure from Reference [16]

Quasi-elastic scattering [186] is the dominant mechanism below  $\sim 1$  GeV. The process involves the neutrino interacting singularly with a proton or neutron within the nuclei of the liquid argon, see Figure 3.5. Bubble chamber data between 1970-1990s provides insight into neutrino cross-section of quasi-elastic interactions using deuterium nuclei [16]. Such interactions are well explained by V-A theory, see Section 2.1; however, more modern experiments use heavier nuclei such as carbon and argon as detector media. These require additional theory: rather than nucleons being considered free particles, a form of the Fermi Gas model is applied which describes the energy states with which fermions are arranged within a potential well [16]. In addition, the impulse approximation [187] is used which assumes that the neutrino interacts with the nucleon alone (e.g. no nuclear effects between the nucleon (or neutrino) and the rest of the nucleus are considered).

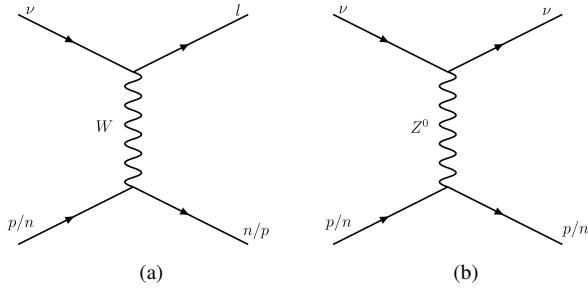


Fig. 3.5 Feynman diagrams of charge current (a) and neutral current (b) quasi-elastic events upon a proton ( $p$ ) or a neutron ( $n$ ).  $W = W^+ / W^-$  depending on the interaction. The result is a charged lepton,  $l$ , which is identified within the detector.

Meson Exchange Current (MEC) (or 2p2h) interactions have also been recently considered to explain cross-section discrepancies in data. This is where the neutrino interacts with two nucleons [188].

At higher energy levels, neutrinos are more likely to interact via deep inelastic scattering where neutrinos interact individually with the quarks inside the nucleons [186]. The deep inelastic cross-section is a function of the parton distribution functions which describe the structure of nuclei.

Resonance and coherent pion production [189] interactions produce additional particles in the process. For resonant single pion interactions, see Figure 3.6, the neutrino interaction excites the nucleon and produces a baryon resonance [16]. The excited baryon then decays releasing a pion. The Rein and Sehgal model [190] is used to model resonance interactions. In addition to pion production, baryon resonances can also create photons during the decay with a branching fraction of less than 1%. Events where a photon is created, or a neutral pion is created, which then decays into two photons, are the predominant background from the  $\nu_e$  charge current selection discussed in Chapter 7.

Multiple pion decays and other hadronic decays, such as kaon production, are also possible when the resonant baryon decays. This depends on the baryon originally created. It is also possible that the neutrino interacts with the entire nucleus producing a pion and no nuclear recoil. Such reactions are known as coherent scattering [16].

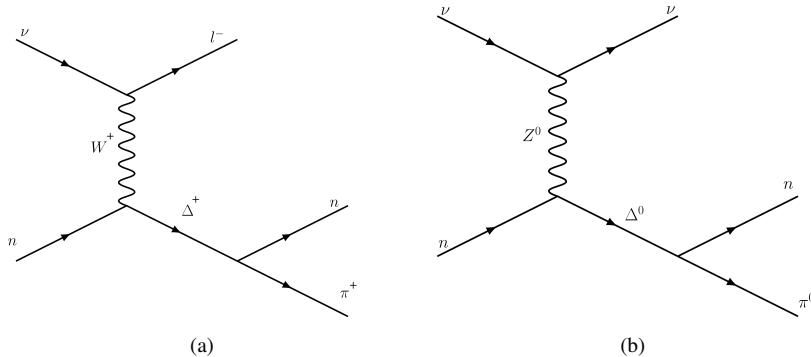


Fig. 3.6 Feynman diagrams of charge current (a) and neutral current (b) resonant events upon a neutron ( $n$ ). Other permutations upon neutrons and protons are possible.

One additional complication to the mechanisms discussed above is the exiting of the final state particles from the nucleus. During the escape from the nucleus, final state particles can interact with the nucleus causing the scattering of other particles. These other particles are also picked up by the detector and we cannot differentiate between them and final state particles experimentally. These interactions that occur after the initial neutrino interaction are known as Final State Interactions (FSI) [191]. FSIs are modelled using intranuclear cascade models [192] which describe the propagation of the initial final state particle [16]. This is done using  $\pi$ -neutron and  $\pi$ -proton cross-section data and assumes that the free  $\pi$ -proton/neutron data resembles the cross-sections within nuclei. These cascade models are used to account for the Final State Interactions in the analysis.

## 3.7 Concluding Remarks

The Short-Baseline Neutrino (SBN) programme, described in Chapter 4, has the capability to confirm or refute the low energy excess observed at LSND and MiniBooNE. The SBN programme is to employ the use of Liquid Argon Time Projection Chambers (LArTPCs) to identify neutrino interactions by providing calorimetric, topological reconstruction, see Chapter 6. This is because LArTPCs are an excellent choice of detector due the inert nature, low threshold, high density and low cost of liquid argon, as described at the start. The design and the corresponding detector effects of LArTPCs have been discussed in this Chapter as well as the interactions that occur within the detector.



# Chapter 4

## Overview of The Short-Baseline Neutrino Programme (SBN)

The SBN programme is a set of experiments situated at Fermilab in Chicago, IL. One of the primary goals of the programme is to confirm or refute the low energy excess seen in the Mini Booster Neutrino Experiment (MiniBooNE), described in Chapter 2, Section 2.2.5.

Three Liquid Argon Time Projection Chamber (LArTPC) detectors form the SBN programme: the Short-Baseline Near Detector (SBND), Micro Booster Neutrino Experiment (MicroBooNE), and the Imaging Cosmic And Rare Underground Signals (ICARUS) detector. The detectors are situated on-axis on a muon neutrino beam called the Booster Neutrino Beam (BNB) [18, 193]. This beam provides neutrinos with approximately 1 GeV of energy. Section 4.1 discusses the design and neutrino output of the beam. The three LArTPC detectors are positioned at 110 m (SBND), 470 m (MicroBooNE) and 600 m (ICARUS), see Figure 4.1, such that the flux systematics of the BNB are well constrained for analysis, see Section 4.6.1. This Chapter will discuss the design of the detectors in Sections 4.3, 4.4 and 4.5.



Fig. 4.1 The positions of the detectors along the BNB beamline. The near detector SBND is positioned at 110 m from the beam, MicroBooNE is at 470 m and ICARUS is at 600 m. From Reference [17].

The programme is expected to take three years worth of data (corresponding to  $6.6 \times 10^{20}$  Protons On Target (POT)) equating to approximately seven million neutrino interactions in the SBND. An initial analysis of the physics capabilities of the SBN programme, known as the SBN proposal [17], is discussed in Section 4.6. Due to the large data-sets expected to be taken by the programme, detailed simulation and automated reconstruction are required. Section 4.7 will discuss the simulation and Chapter 6 will discuss the reconstruction.

## 4.1 The Booster Neutrino Beam (BNB)

The BNB is a muon neutrino beam provided to the SBN programme, shown in Figure 4.2. It is created by impacting protons with a kinetic energy of 8 GeV onto a beryllium target to produce a beam of hadrons known as secondary hadrons. The resultant beam hadrons and their branching ratios can be found in Table 4.1. The protons are accelerated using the booster accelerator, which provides approximately  $5 \times 10^{12}$  protons per spill at a rate of 5 Hz. A spill lasts  $1.6 \mu\text{s}$  in length and consists of 81 bunches of protons which are approximately 2 ns wide and 19 ns apart. The separation between the bunches is significant enough to be detected by the light detection systems in the SBND detector, described in Section 4.3.

Hadron	Decay Mode	Branching Ratio (%)
$\pi^+$	$\mu^+ + \nu_\mu$	99.9877
	$e^+ + \nu_e$	0.0123
$K^+$	$\mu^+ + \nu_\mu$	63.44
	$\pi^0 + e^+ + \nu_e$	4.98
	$\pi^0 + \mu^+ + \nu_\mu$	3.32
$K^0$	$\pi^- + e^+ + \nu_e$	20.333
	$\pi^+ + e^- + \bar{\nu}_e$	20.197
	$\pi^- + \mu^+ + \nu_\mu$	13.551
	$\pi^+ + \mu^- + \bar{\nu}_\mu$	13.469
$\mu^+$	$e^+ + \nu_e + \bar{\nu}_e$	100

Table 4.1 Hadrons created from the BNB by impacting protons on the target and their branching ratios to neutrinos. From Reference [18].

As the secondary hadrons are created, they enter a pulsed electromagnet that focuses positively charged particles and defocuses negatively charged particles using a toroidal magnetic field. The beam then enters a decay region which is 50 m in length. A schematic of this is shown in Figure 4.3 [18]. In this region, the secondary hadrons decay and produce neutrinos. The output beam is made up of  $\sim 99.4\%$  muon neutrinos and anti-neutrinos along with  $\sim 0.6\%$  intrinsic electron neutrinos. Table 4.2 has a full breakdown. The predicted flux

from MiniBooNE [18] is shown in Figure 4.3. The Figure shows neutrino energies range from 0 to 5 GeV with a peak at 1 GeV. It also shows the energy distribution of the intrinsic electron neutrinos and the muon neutrinos.

Neutrino Flavour	Fraction of Beam Neutrino Mode	Fraction of Beam Antineutrino Mode
$\nu_\mu$	93.6 %	15.71%
$\nu_e$	0.52 %	0.2%
$\bar{\nu}_\mu$	5.86%	83.73%
$\bar{\nu}_e$	0.05 %	0.4%

Table 4.2 The composition of the BNB. From Reference [18].

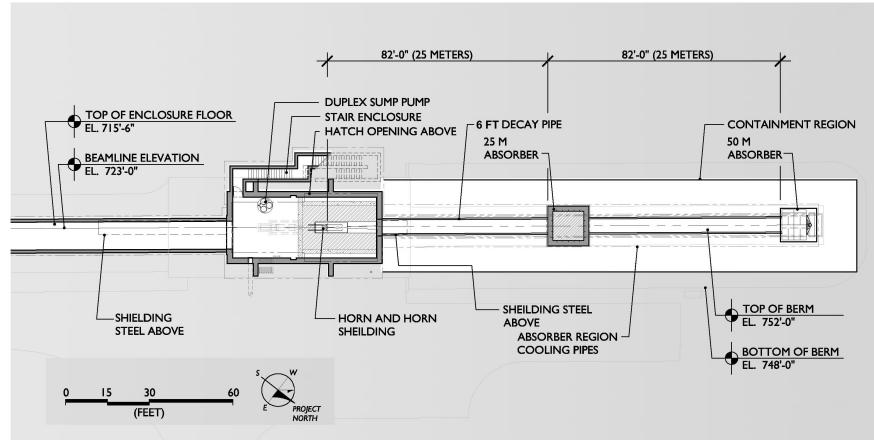


Fig. 4.2 A schematic of the BNB beam indicating the focusing horn and the decay region. From Reference [18].

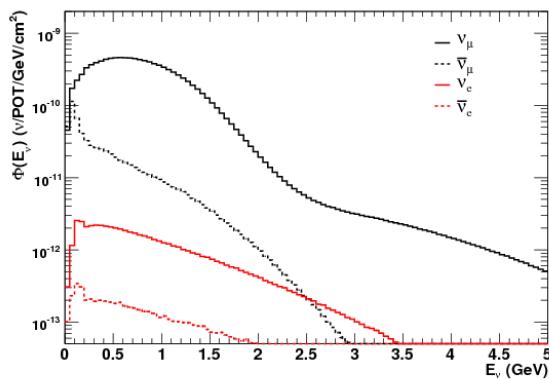


Fig. 4.3 The simulated flux of the BNB beam at MiniBooNE split into the components of the beam. From Reference [18].

There are several systematic uncertainties associated with the BNB. A detailed review of the uncertainties can be found in Reference [194] and they are further discussed in Chapter 8. The flux systematics accounted for in SBN analyses include the error on the number of protons delivered to the target, the secondary hadron and neutrino production rates, changes in the horn magnetic field, and misalignments in the beamline geometry. Table 4.3 shows the total flux uncertainty for the systematics listed above.

Source of Uncertainty	$\nu_\mu$	$\bar{\nu}_\mu$	$\nu_e$	$\bar{\nu}_e$
Proton delivery	2%	2%	2%	2%
Proton optics	1%	1%	1%	1%
$\pi^+$ production	14.7%	1.0%	9.3%	0.9%
$\pi^-$ production	0.0%	16.5%	0.0%	3.5%
$K^+$ production	0.9%	0.2%	11.5%	0.3%
$K^0$ production	0.0%	0.2%	2.1%	17.6%
Horn field	2.2%	3.3%	0.6%	0.8%
Nucleon cross-section	2.8%	5.7%	3.3%	5.6%
Pion cross-section	1.2%	1.2%	0.8%	0.7%

Table 4.3 Variations in the total flux for each neutrino species. From Reference [18].

## 4.2 Neutrinos at the Main Injector (NuMI) Beam

MicroBooNE and ICARUS are also able to detect neutrinos from the NuMI beam [195, 196]. The process to create the NuMI beam is similar to that of the BNB; however, in this case the proton target is made of graphite. The result is a beam which produces neutrinos with energies in the 1-10 GeV range, which peak at a slightly high energy than that of the BNB. However, both MicroBooNE and ICARUS are situated  $8^\circ$  and  $6^\circ$  [197] off-axis from the beamline. This means the peak energy of the beam in the detectors is shifted to lower energies, and most events detected are in the 0-3 GeV energy range [197]. The difference between the BNB flux [19] and NuMI flux in MicroBooNE [20] is shown in Figure 4.4.

One advantage of the NuMI detector is the increase in intrinsic electron neutrinos due to the enhancement in the dominant three body decay of secondary kaons [197]. Because of this, approximately 5% of the NuMI interactions that will occur in the ICARUS detector are electron neutrinos. This results in roughly 10,000 electron neutrino events per year in ICARUS [197]. Therefore, both detectors are capable of performing high statistical  $\nu_e$  cross-section measurements [20].

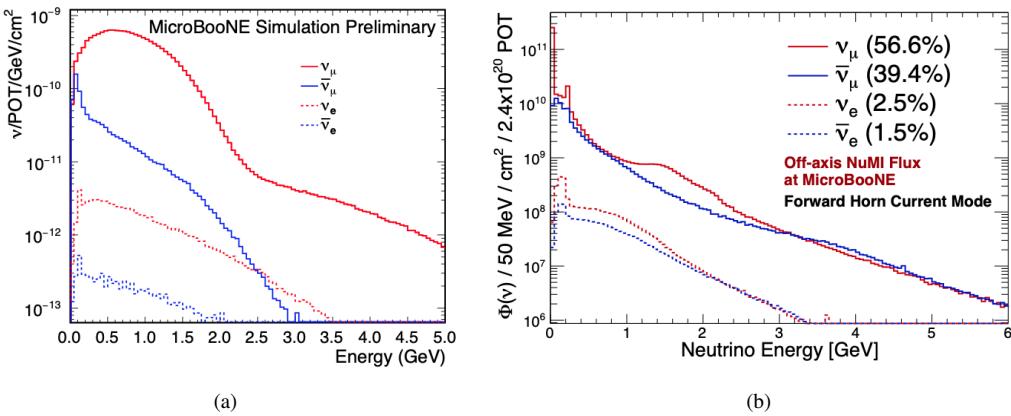


Fig. 4.4 The predicted flux from the BNB (a) (from Reference [19] and the NuMI beam (b) (from Reference [20]) to MicroBooNE as a function of neutrino energy. Both beams are in a mode which selects neutrinos over anti-neutrinos.

### 4.3 The Short-Baseline Near Detector (SBND)

### 4.3.1 Detector Design

The SBND is the near detector of the SBN programme. The SBND detector consists of two LArTPCs of dimensions 2.0 m in width, 4.0 m in height and 5.0 m in the beam direction; see Figure 4.5 for a diagram of the detector [17, 198]. The resultant active volume of this design will hold 112 tons of liquid argon. The TPCs are separated by the Cathode Plane Assembly (CPA) which provides a 500 V/cm electric field to each TPC. The CPA is a steel structure which holds a set of wire mesh frames such that the cathode provides a uniform electric field across the detector. Behind the wire mesh are Tetra-Phenyl-Butadiene (TPB) coated reflector tiles which optically isolate the TPCs [17]. The reflector tiles re-emit the ultraviolet, liquid argon, scintillation light in the visible spectrum for detection within the Photon Detection System (PDS).

The anode, also known as the Anode Plane Assembly (APA) is shown on the left and right sides of the right-hand plot of Figure 4.5. The APA for each TPC is made of three wire planes, U, V and Z, and contains a total of 5632 copper-beryllium readout wires, allowing for 3D reconstruction of events. The wires are  $150\text{ }\mu\text{m}$  in diameter and are separated by 3 mm in pitch [17]. The U and V planes are the first and second induction planes respectively and are oriented  $\pm 60^\circ$  to the vertical when facing the plane face. The collection plane, Z (or Y), is orientated such that the wires are in the vertical direction, shown in Figure 4.6. A bias of -200 V, 0 V and +500 V is applied to the U, V and Z planes, respectively, to make the induction wires transparent to drift electrons [17]. Each APA is made from two steel frames,

thus there is a gap in the middle of the wire planes. In the gap, an electron diverter diverts drift electrons on to the neighbouring planes to ensure no energy depositions are lost.

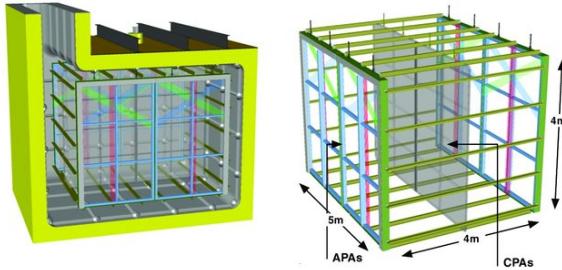


Fig. 4.5 A schematic of SBND. The CPA lies between the two TPCs providing an electric field. The anode (APA) readout consists of 3 wire planes. From Reference [17].

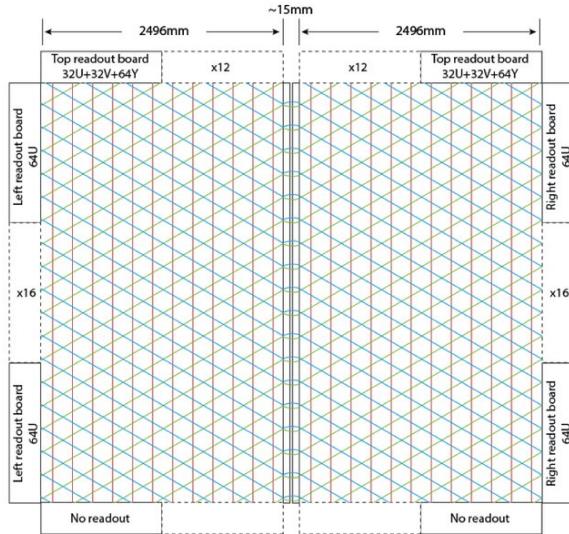


Fig. 4.6 SBND wire readout setup. From Reference [17].

The field cage for the TPCs is made of copper and is based on the 35t Deep Underground Neutrino Experiment (DUNE) prototype [179]. The detector is secured in a membrane cryostat. This consists of a steel outer supporting structure which holds further supporting structures, insulation, a vapour barrier (to prevent moisture entering the insulation), and the primary and secondary membranes. The insulation is a solid reinforced polyurethane foam. The primary membrane is made from stainless steel and holds the liquid argon whilst the second membrane is a backup in case the primary membrane leaks.

The Photon Detection System (PDS) is a set of 24 modules. One module is shown in Figure 4.7 where it can be seen that the module contains several detectors for the argon scintillation light. The PDS consists of two sets of detectors: the Photo Multiplier Tube (PMT) based system and the ARAPUCA [199] (Argon R&D Advanced Program at UniCAMP) based system. The PMT system consists of 120 Hamamatsu 8 inch R5912-mod PMTs [200, 201].

The ARAPUCA is a light-trapping device which consists of a cavity and SensL MicroFC-60035-SMT [202] silicon photomultiplier(s) (SiPM(s)). The dimensions of the cavity are 3.6 cm × 2.5 cm × 0.6 cm (0.6 cm depth). A window made of a dichronic filter, which is coated externally with a film of p-Terphenyl, allows the UV scintillation to pass into the cavity [203]. The inner side of the filter and the interior of the cavity are covered in Tetraphenyl butadiene (TPB) to shift the light to the visible region of the light spectrum for detection. Shifting the light also traps the light within the cavity. A modification to the ARAPUCA design, the X-ARAPUCA, also exists and replaces the TPB on the inner side of the filter with an acrylic bar such that SiPMs can detect light via total internal reflection as well as the existing ARAPUCA methods [204, 205]. 8 of the original ARAPUCAs and 167 of the X-ARAPUCAs are to be installed in SBND. SBND, therefore, is a good test in a neutrino beam for the (X-)ARAPUCA which are proposed for the DUNE far detector [205]. However, the primary goal of the PDS is to provide precise nanosecond timing information for events in the detector. Also, calorimetric and topological information can be extracted from the scintillation light using the PDS.

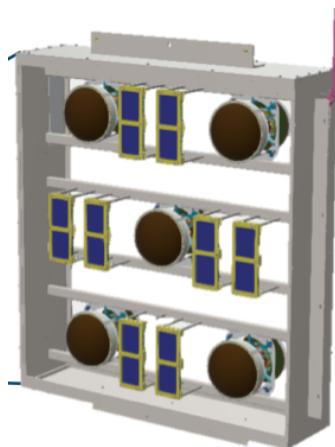


Fig. 4.7 A PDS module in SBND. The module is made up of 5 PMTs and 8 ARAPUCAs which are positioned interchangeably. Mounting brackets on all four sides of the detectors hold the detectors to the frame. Credit: J. Boissevain.

Around the cryostat, there is the Cosmic Ray Tagging (CRT) system [21] which is schematically shown in Figure 4.8a. The CRT system provides  $\sim 94\%$  coverage of the total cosmic flux in order to identify cosmic events which enter the TPC. The 6% that is uncovered is where the feet of the cryostat touch the floor. This is not shown in Figure 4.8a. SBND lies on the surface, and the cosmic muon rate through the detector is approximately five cosmic interactions in a 3 ms drift window. The drift window allows for electrons to travel the distance of the TPC for detection. This cosmic rate equates to approximately 1 in 300 beam spills, is expected to contain a cosmic ray in coincidence with the beam spill. These are called in-time cosmic interactions. This value does not fully describe the cosmic contamination in neutrino interactions. Out of time cosmic interactions can occur as the charge deposition drifts to the readout and contaminate the event. The primary aim of the CRT system is to identify all cosmic events in the detector.

The CRT system is made up of 143 scintillator modules which surround the detector. The modules, which are pictured in Figure 4.8b, consist of 16 scintillator strips each covered with a highly-reflective white coating. Adjacent to the scintillator strips, in grooves, wavelength shifting fibres transport scintillation light to a SiPM. A protective and reflective Mylar tape is placed over the groove to prevent leakage. This is shown in the right-hand image in Figure 4.8b [21]. The far end of the fibres is coated with aluminium by evaporation in order to reflect the light and improve the light detection efficiency. The reflected light intensity compared to the non-reflected light can also be used to improve spatial resolution up to 1.8 cm [21]. The strips are then encased in aluminium panels, to which they are fixed using double-sided adhesive tape. The aluminium case is sealed using optically absorbent adhesive glue and screws. The SiPMs are connected to a Front end Electronics Board (FEB) for the readout. To perform 2D reconstruction of particles, two layers of panels are oriented perpendicular to one another.

The probability of getting more than one cosmic event in a strip is of the order  $10^{-5}$ , therefore two interactions at the same time is unlikely. In addition, to improve the identification efficiency, coincidence matching occurs between the two sides of the scintillator strips to avoid misidentification of dark current pulses. The timing resolution is 1 ns [21]. Figure 4.9 shows the tagging efficiency of events as a function of the distance from the readout. As can be seen in the Figure, the efficiency increases the closer the event occurs to the readout, reaching  $> 99\%$  below 100 cm. The efficiency is always greater than 95%

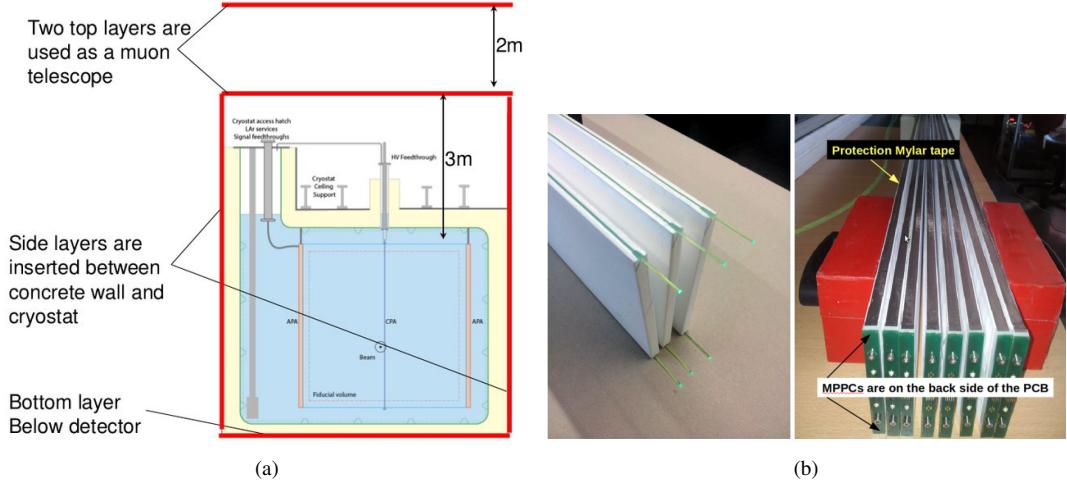


Fig. 4.8 a) Position of the CRT system in SBND. Note that the bottom of the detector is not fully covered due to struts holding the cryostat in position. From Reference [17]. b) Images of the scintillating strips before and after the Mylar has been glued on. From Reference [21].

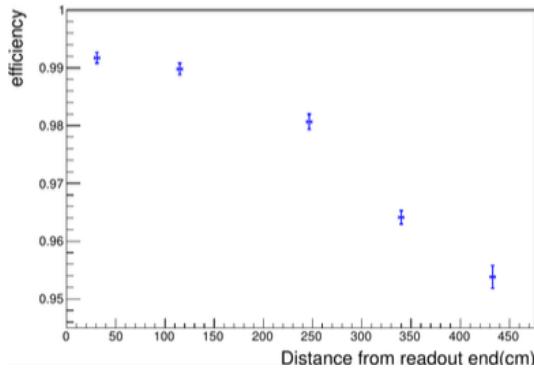


Fig. 4.9 The efficiency of the prototype CRT panel as the distance from the readout increases. This is a result of a laboratory tests at the University of Bern. From Reference [21].

### 4.3.2 Electronics Design

The SBND TPC electronics chain is depicted in a block diagram for a single channel in Figure 4.10. SBND is a prototype for the DUNE experiment and one of the goals of SBND is to introduce new electronics which reside within the liquid argon in order to reduce the number of readout channels exiting the cryostat. The electronic setup will be discussed in this Section.

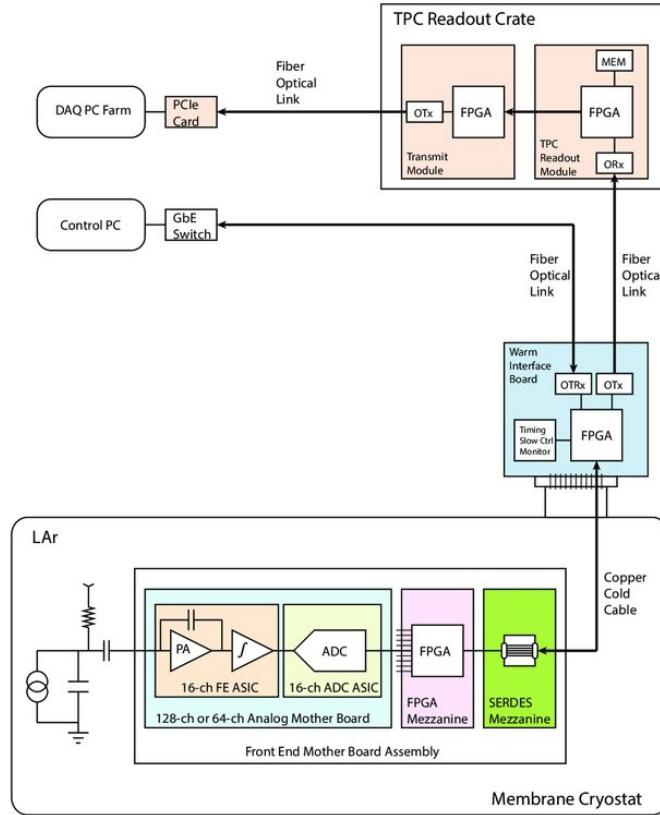


Fig. 4.10 Full path of the electronics chain in SBND. From Reference [17].

Once a charge induces a current on the readout wires, the signal is passed to a Complementary Metal-Oxide-Semiconductor (CMOS) front end Application-Specific Integrated Circuit (ASIC) known as the LArASIC-v7 [206–208]. The LArASIC preamplifies and shapes the signal using an anti-aliasing filter based on a fifth-order, low-pass network designed to obtain an impulse response close to a Gaussian distribution in the time domain [158, 209]. Hence the charge on the collection plane for an impulse signal can be approximated by finding the peak height of the Gaussian response. Or, equivalently, one can integrate the charge of the response and multiply it by  $\frac{1}{2\sqrt{2\pi}\tau}$ .  $\tau$  is the shaping time in  $\mu\text{s}$  and the factor of 2 arises from converting the integral of the response from the timing tick of the Analogue Digital Converters (ADC) to  $\mu\text{s}$ . A more sophisticated way to extract the charge is to deconvolve the signal with the electronic response and the field response beforehand, which is discussed in Chapter 6, Section 6.1.1. The LArASIC can operate at four gain settings (4.7 mV/fC, 7.8 mV/fC, 14 mV/fC and 25 mV/fC) and four shaping time settings (0.5  $\mu\text{s}$ , 1.0  $\mu\text{s}$ , 2.0  $\mu\text{s}$  and

$3.0\ \mu s$ ). For the SBND simulation, the gain is set to 14 mV/fC and the shaping time to  $2.0\ \mu s$ .

The signal is then passed to a 12 bit, 1.8 V peak-to-peak, AD7274 Commercial-Off-The-Shelf (COTS) ADC [210, 211]. Having the ADC in the cryostat aids in the reduction of noise and reduces the number of channels exiting the cryostat. This digitised signal is then passed to a Field-Programmable Gate Array (FPGA) where the signal is multiplexed and sent to the Serializer/Deserializer (SerDes) to prepare the data for transmission out of the cryostat. The LArASIC, COTS ADC and a fraction of the FPGA and SerDes boards make up the Front End Mother Board (FEMB) assembly [17]. The LArASIC reads up to 16 channels, and therefore four LArASICs and 4 COTS ADCs make up a FEMB. 176 FEMB are required for the full TPC readout. 124 FPGA and SerDes boards are used in a full FEMB assembly. The signals are passed out of the cryostat via copper cable to a Warm-Interface Board (WIB) which distributes signals in 64-channel chunks of uncompressed data to a TPC readout module via optical fibre. The WIB also controls the clock synchronisation. The TPC readout crate then compresses and stores the data in a beam data stream and a continuous supernova data stream [17]. 11 TPC read-out crates are used to readout the data and then send the data to DAQ (Data AcQuisition system) PCs via optical links.

This COTS ADC electronics design is of particular interest for future experiments such as DUNE and is a new setup in liquid argon neutrino physics. This is because the DUNE experiment will have a large number of readout channels which will be difficult to extract from the cryostat efficiently. Placing the ADC within the cryostat will reduce the number of channels leaving the detector. Tests of the SBND electronic readout is discussed in Chapter 5.

## 4.4 The MicroBooNE Detector

The MicroBooNE detector has been running since 2015 and has produced several physics analyses separate from SBN. Many of the analyses have motivated the procedures, commissioning practices, simulation and reconstruction methods implemented by the other detectors of the SBN programme. Of particular note for this thesis are the noise characterisation studies and practices [212] that have been implemented. This has been drawn upon in Chapter 5]. Also of note is the implementation of the Pandora pattern recognition software within liquid argon [27] as well as the development of deconvolution processes [157] used upon the waveforms from the TPC, both of which are described in Chapter 6. MicroBooNE has also performed analyses in order to determine space charge effects in the detector [213], develop signal processing using 2D deconvolution techniques [214] as well as creating techniques to

identify and reconstruct cosmic rays [215]. These techniques are not taken advantage of in this thesis and are an essential addition for future iterations of the analysis.

MicroBooNE is also capable of independently performing cross-section measurements [216, 20] which are useful for tuning current Monte Carlos for the SBN programme as well as for future experiments. With the addition of the intrinsic electrons from the NuMi beam the number of intrinsic electrons is comparable, although MicroBooNE is significantly further away from the BNB than SBND. Also, although SBND will have more statistics for cross-section data, the methods developed by MicroBooNE pave the way for effective analyses with the SBND data.

The main physics goal of MicroBooNE was to investigate the low energy excess observed at MiniBooNE. Hence, MicroBooNE is capable of performing independent sterile searches. Figure 4.11 [22] shows the possible sensitivity of an electron neutrino excess given a set of sterile oscillation parameters following the 3+1 model [134], discussed in Chapter 2. As shown in the Figure, it is hoped that MicroBooNE will compare favorably to that of MiniBooNE. In fact, in some areas of the LSND allowed region of the sterile oscillation parameters, MicroBooNE is able to identify an excess with a confidence level that is greater  $5\sigma$ . The SBN programme improves upon MicroBooNE due to the reduction in flux systematic errors because of the multi-detector approach.

The design of the detector is similar to the SBND detector and various similarities. These include: include the liquid argon detector medium, the readout wires and the wire pitch. In addition the detector uses a version of the LArASIC chip for pre-amplification and shaping of the signal [207]. Therefore the inherent electronics noise of a similar form. One of the main differences is the geometry of the cryostat [22]. The MicroBooNE cryostat is cylindrical, see Figure 4.12, with a length of 12.2 m and an inner diameter of 3.81 m. The cryostat houses a rectangular TPC with 170 tonne capacity with dimensions of 2.6 m in drift length, 2.3 m in height and 10.4 m in the direction of the beam [13]. Due to the change in the shape of the cryostat, the amount of liquid argon between the TPC and the edges of the cryostat varies. This is unlike the SBND and ICARUS detector geometries. The change in volume affects the rate of dirt events which are events, which interact outside the active volume but deposit energy within the TPC. As is discussed in Chapter 8, the dirt events are significant background in the sterile oscillation analysis.

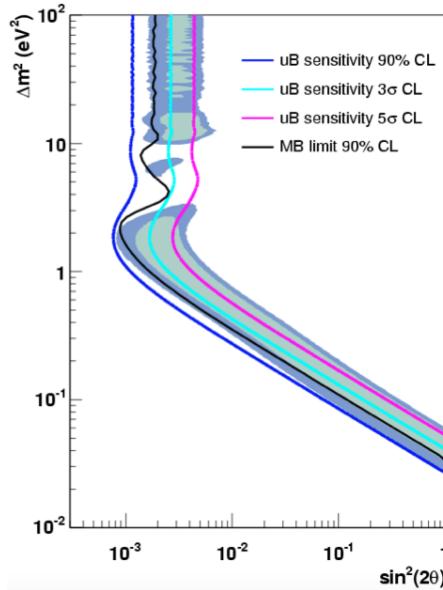


Fig. 4.11 The MicroBooNE sensitivity using Monte Carlo to an excess of electron neutrino events given the sterile oscillation parameters in the x and y axes. The MiniBooNE sensitivity 90% confidence limit is also shown and the LSND 90% confidence limit. Area to the right of the contour has a higher confidence value than the contour value. See Chapter 8 for further detail. From Reference [22].

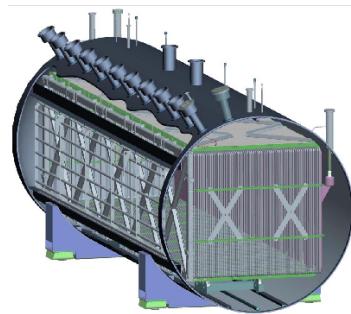


Fig. 4.12 Schematic of the MicroBooNE cryostat and TPC. From Reference [13].

Another difference is that the ADCs are outside the detector which changes the electronics noise.

The MicroBooNE anode is made up of three wire planes, U 2400 wires, V 2400 wires and Y 3456 wires, each with a 3 mm pitch. The two induction planes U and V are  $\pm 60^\circ$  with respect to the vertical, and the collection plane wires are in the vertical direction [13]. This geometry is the same orientation as the SBND detector. 32 PMTs are positioned behind the wire planes to read out the scintillation light. The MicroBooNE CRT system is similar to SBND CRT system, Section 4.3, and uses identical modules. MicroBooNE currently runs at

a drift electric field of 273 V/cm compared to the nominal 500 V/cm [158]. The difference in the field affects the time taken for electrons to drift to the readout anode and so affects properties such as the signal to noise level, see Chapter 5 Section 5.2.3.

## 4.5 The ICARUS Detector

The ICARUS detector was previously installed at the Laboratori Nazionali del Gran Sasso (INFN-LNGS) Laboratory [159]. ICARUS, like MicroBooNE, was a detector in its own right before SBN and was the first large scale LArTPC. As such, the experiment contributed significantly to the field in its early years. Of particular interest to this thesis is the electron lifetime method developed, which is used in Chapter 5, and the reconstruction methods ICARUS developed [217]. ICARUS also performed an independent analysis to search for the low energy excess. The results of which are shown in Figure 4.13 where the sensitivity to an excess of electron neutrinos due to the presence of a sterile neutrino with a range of 3+1 oscillation parameters is presented. Due to muon neutrino energy range being analysed (10–30 GeV), [23] and the distance travelled (730 km), the  $L/E \approx 36.5 \text{ m/MeV}$  [23] was much larger than the low energy excess experiments. Thus for oscillation parameters in the low-energy excess region, the oscillation approximately averages out to 0.5 [23]. This is why the curve becomes vertical.

As can be seen in Figure 4.14, the ICARUS detector consists of two identical modules, separated by  $\sim 1.36 \text{ m}$ , with each having a cathode in the centre supplying an electric field to two TPCs. Each TPC contains 476 tons of liquid argon in the active volume and has dimensions of 1.5 m in the drift length, 2.96 m in height and 18 m in the beam direction [159]. The anode of each TPC consists of three wire planes with a 3 mm wire pitch. However, unlike the SBND and MicroBooNE detectors, the middle induction plane and the collection plane are orientated  $\pm 60^\circ$ , whilst the outermost induction plane is orientated such that the wires are travelling horizontally in the beam direction. This is a disadvantage as hit and clustering reconstruction suffers when particles travel horizontal to the wire plane, see Chapter 6. The first induction plane wires are separated into two adjacent planes each with 1056 wires. The middle induction plane and the collection plane are separated into two adjacent planes of unequal size. The smaller plane contains 960 wires, and the larger plane contains 4640. 54248 wires are used in the detector in total [159]. The cathode is designed to provide an electric field of 500 V/cm across the drift length of the TPCs.

As part of the ICARUS relocation to Fermilab, some parts of the experiment underwent refurbishment. The refurbishment was completed at CERN and included the implementation of new readout electronics, updates in the slow control system, a new cryostat and purification

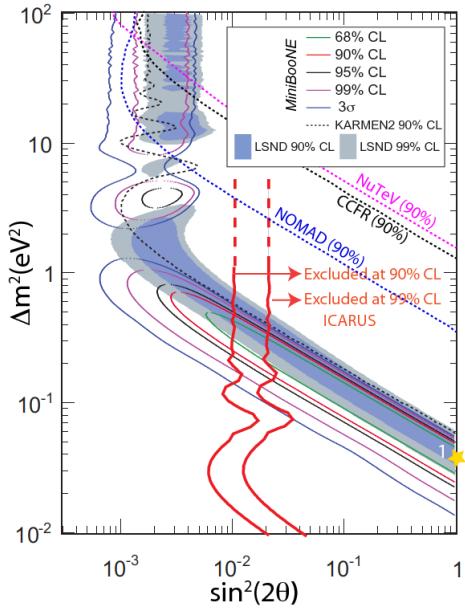


Fig. 4.13 The ICARUS sensitivity to an excess of  $\nu_e$  events given the sterile oscillation parameters in the x and y axes. Area to the right of the ICARUS contour has a higher confidence value than the contour value. See Chapter 8 for further detail. From Reference [23].

system and a new light collection system [17]. The upgrade also includes a CRT system. The light collection system is positioned 40 cm behind the wire planes and consists of 3 rows of 9 ETL 9357FLA PMTs [218] spaced by 2 m, adding up to 54 PMTs in total. The new light collection system is expected to resolve the bucket structure in the BNB discussed in Section 4.1. The CRT system is split into three subsystems: the top, sides and bottom. The top system is made of 125 scintillator strips with a SiPM based readout, while the sides are made from the Main Injector Neutrino Oscillation Search (MINOS) experiment scintillator shield modules [219]. The bottom panels are made from the Double Chooz experiment veto modules [220]. The CRT system provides almost full coverage of the detector.

Like MicroBooNE the electronics for ICARUS differ from SBND as the analogue to digital conversion of the signals occurs outside the cryostat. However, the electronic readouts are similar; all three detectors use versions of the LArASIC chip for pre-amplification and shaping of the signal [207]. Other similarities between the detectors include the liquid argon detector medium, the readout wires and the wire pitch. Therefore similar particles will deposit the same amount of energy within the per wire spacing, disregarding the detector effects discussed in Chapter 3. Therefore calibration efforts and reconstruction efforts are linked. These similarities will help to minimise detector systematics in future oscillation analyses.

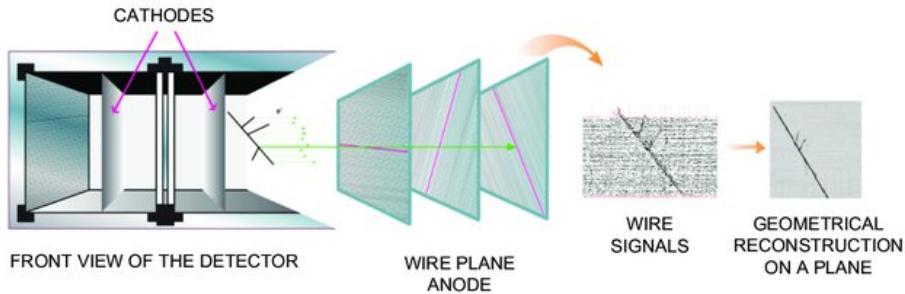


Fig. 4.14 Diagram showing the geometry of the ICARUS TPCs. From Reference [17].

## 4.6 Physics Capabilities

The SBN programme can perform searches for several physics phenomena predicted by theories beyond the standard model. As well as this, it can produce high statistical cross-section inclusive and exclusive measurements. This Section will discuss the physics capabilities of the programme with a particular focus on sterile searches.

### 4.6.1 Sensitivity of Sterile Neutrino Searches

The SBN programme is capable of performing searches of  $\nu_\mu \rightarrow \nu_e$  appearance,  $\nu_e \rightarrow \nu_x$  disappearance and  $\nu_\mu \rightarrow \nu_x$  disappearance. To define the capabilities of the programme a sensitivity analysis is performed using simulated data. To do this the original SBN experiment proposal, The analysis simulated events and used Monte Carlo truth information to perform selections of events and analyses [17]. This was instead of reconstructing information from the energy depositions using the simulated detector data. The analysis is therefore defined as the truth-based analysis in this thesis. This Section will discuss the analyses performed in the proposal [17] which predict the sensitivity of the SBN programme to sterile neutrino searches (which is not the work of the author). The topic of the this thesis recreates the below analysis and develops the selection criteria using reconstructed simulated data in Chapters 7 and 8. The analyses undertaken were an electron neutrino appearance analysis and a muon neutrino disappearance analysis. The results of the electron neutrino appearance analysis, described in this Section, are then compared to an updated version of the analysis as well as analyses which use the detector readout information, rather than truth information, in Chapters 7 and 8.

The common theoretical framework used for the sterile oscillation is the "3+1" model, see Reference [134] for further information and described in Chapter 2, Section 2.2.5. Further

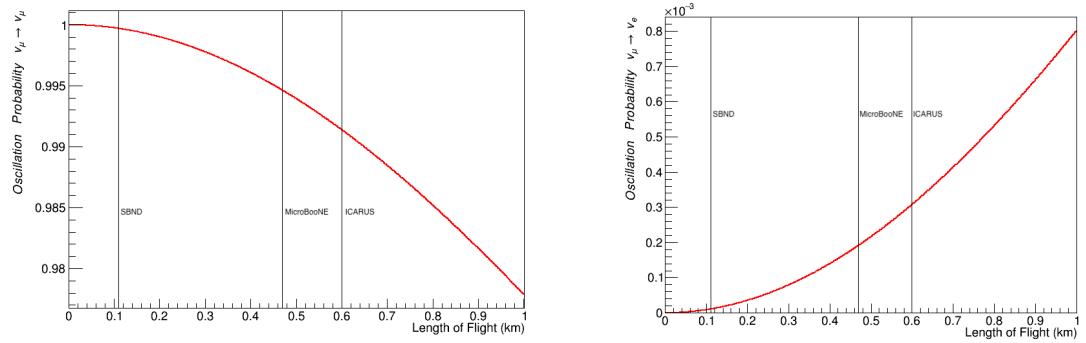
developments since the proposal, which were presented in Chapter 2, Section 2.2.6, have placed constraints on the "3+1" model. Therefore further sterile models will be evaluated in the future. However, in this thesis, the "3+1" model is used for comparison.

The probability of oscillations in the "3+1" model for  $\nu_e$  appearance and  $\nu_\mu$  disappearance are:

$$P_{\nu_\mu \rightarrow \nu_e}^{3+1} = \sin^2 2\theta_{\mu e} \sin^2 \left( \frac{\Delta m_{41}^2 L}{4E_\nu} \right), \quad P_{\nu_\mu \rightarrow \nu_\mu}^{3+1} = 1 - \sin^2 2\theta_{\mu \mu} \sin^2 \left( \frac{\Delta m_{41}^2 L}{4E_\nu} \right), \quad (4.1)$$

where  $L$  is the length of flight of the neutrino,  $E_\nu$  is the energy of the neutrino and  $\theta_{\mu e}$ ,  $\theta_{\mu \mu}$  and  $\Delta m_{41}^2$  are oscillation parameters. Further details are found in Chapter 8.

Figures 4.15a and 4.15b show the oscillation probabilities for the global best fit values for  $\nu_e$  appearance of  $\Delta m_{41}^2 = 1.2 \text{ eV}^2$  and  $\theta_{\mu e} = 0.003$  [9] and disappearance of  $\Delta m_{41}^2 = 1.32 \text{ eV}^2$  and  $\theta_{\mu \mu} = 0.07$  respectively [221]. The three detectors are identified on the Figures and show that the position of SBND allows for minimal oscillation. This allows the SBND detector to provide a good constraint on the beam content and thus reduce the effects of the flux systematic error, described in Section 4.1, on the sensitivity of the programme.



(a) Oscillation probability of a  $\nu_\mu \rightarrow \nu_\mu$  for best fit oscillation parameters  $\Delta m_{41}^2 = 1.32 \text{ eV}^2$  and  $\theta_{\mu \mu} = 0.07$ . (b) Oscillation probability of a  $\nu_\mu \rightarrow \nu_e$  for best fit oscillation parameters  $\theta_{\mu e} = 0.003$  and disappearance of  $\Delta m_{41}^2 = 1.2 \text{ eV}^2$ .

Fig. 4.15 Oscillation probability as a function of length travel for a 0.7 GeV neutrino. Global best-fit parameters for sterile  $\nu_e$  appearance excess and  $\nu_\mu$  disappearance are used to calculate the probability. SBND (110 m), MicroBooNE (470 m) and ICARUS (600 m) are indicated on the plot by their distance from the beam target. Note that the hadrons decay between the beam target and up to 50 m in the decay pipe, therefore the neutrino length of flight need not be the full distance to the detectors from the beam target.

The sensitivity of the SBN programme to sterile oscillations in the proposal is parameterised by the following  $\chi^2$  definition

$$\chi^2(\Delta m_{41}^2, \sin^2 \theta) = \sum_{ij} [N_i^{\text{null}} - N_i^{\text{osc}}(\Delta m_{41}^2, \sin^2 \theta)] (E_{ij})^{-1} [N_j^{\text{null}} - N_j^{\text{osc}}(\Delta m_{41}^2, \sin^2 \theta)], \quad (4.2)$$

where  $N_i^{\text{null}}$  and  $N_i^{\text{osc}}$  are the number of events in the energy bin  $i$  under the null and oscillated hypotheses for the analysis respectively.  $E_{ij}$  is the covariance matrix element between the energy bin  $i$  and energy bin  $j$ .  $E_{ij}$  describes the combination of the flux, interaction and detector systematic error matrices along with the uncorrelated statistical error matrix. The flux and interaction error matrices are calculated by reweighting events based on changes applied to the flux and interaction systematics. More information can be found in Chapter 8. Currently, no detector systematics are applied. A 15% normalisation variation was allowed on the dirt events.

Selection of inclusive Charge Current (CC)  $\nu_\mu$  events and inclusive CC  $\nu_e$  are required to maximise the sensitivity for the  $\nu_\mu$  disappearance and  $\nu_e$  appearance analyses. The typical signal in the detector for a  $\nu_\mu$  CC event is the muon arising from the interaction, as discussed in Chapter 2, Section 3.6. The muon travels a significant distance and creates a track-like energy deposition. For a  $\nu_e$  interaction, the lepton is an electron and creates a shower-like energy deposition. Therefore, one background to the analyses include track-like particles misidentified as muons for the  $\nu_\mu$  disappearance and shower-like particles misidentified as electrons for the  $\nu_e$  analysis. It is also the case that a track-like particle can be misidentified as a shower and vice versa by automated reconstruction. Other backgrounds include, neutral current interactions, misidentified cosmic interactions and charged particles which propagates into the detector from neutrino interaction that occur outside the TPC.

Figure 4.16 shows the truth-based selection spectra for the  $\nu_\mu$  analysis within the SBN proposal. For the  $\nu_\mu$  analysis, the main backgrounds include cosmic rays which stop in the detector and resonant neutral current interactions where a charged pion is misidentified as a muon. However, cosmic events (and also dirt events) were not included in the analysis. The resonant interactions can be removed from the selection via topological and calorimetric cuts. Muons travel the furthest on average in the detectors because, at the relevant energies, they are minimum ionising and thus deposit (or lose) the least amount of energy per unit length. Therefore, track-like particles that travel small distances, less than 50 cm, were identified as charged pions or protons. These were removed. In the future, a particle identification analysis can be applied using the  $dE/dx$  of track-like particles to remove the pions and protons, discussed in Chapter 6, Section 6.1.8. The sensitivity of the programme for such a selection is shown in Figure 4.17. The sensitivity curves exclude parameter space to the right of the curves and show a significant increase in sensitivity from MiniBooNE.

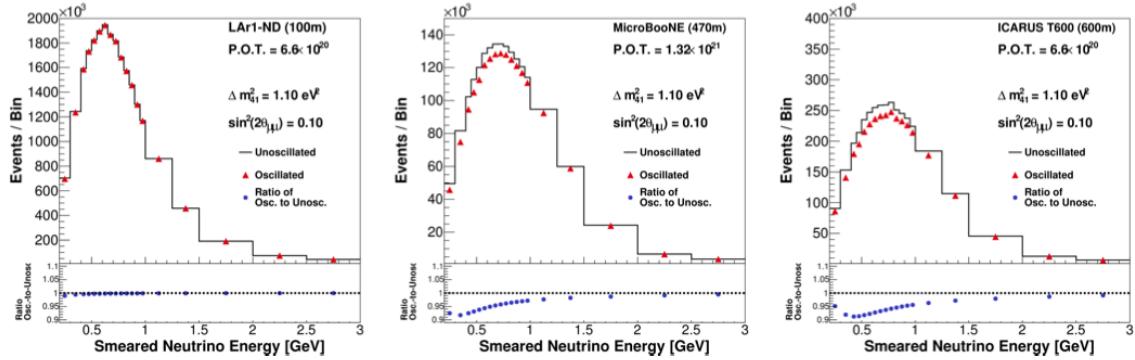


Fig. 4.16 SBN muon neutrino spectra after the  $\nu_\mu$  CC inclusive truth-based selection. Note that no cosmic background was considered in the analysis. From Reference [17].

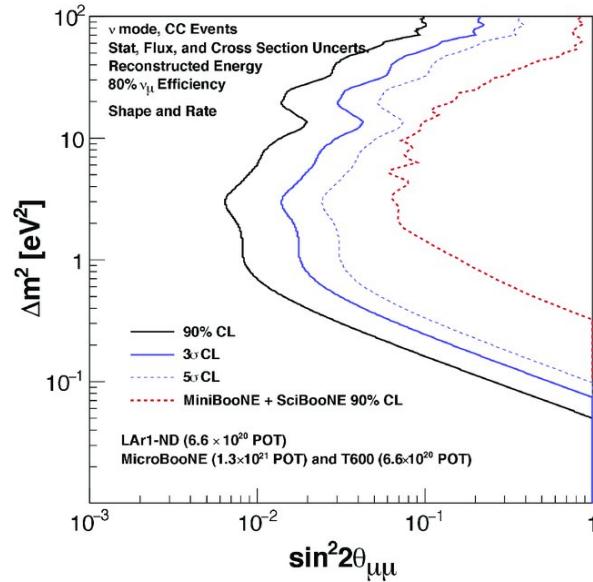


Fig. 4.17 The SBN sensitivity to muon neutrino disappearance for the proposal truth-based analysis as well as the sensitivity contour for a joint MiniBooNE and SciBooNE analysis. Area to the right of the contour has a higher confidence value than the contour value. From Reference [17].

For the  $\nu_e$  analysis, the main background is photon showers which are misidentified as electrons. Photons arise from neutral and CC events originating directly from the vertex interaction and the decay of neutral pions arising from the vertex interaction. The background photons from beam events can occur inside the active volume of the detectors or outside where the photon propagates into the TPC (these are dirt events). Cosmic photon events also can be misidentified as CC  $\nu_e$  interactions.

Figure 4.18 shows the spectra for the proposal selection procedure with the main background compilations. To obtain the spectra, topological and calorimetric cuts to reduce the beam backgrounds and fiducial volume cuts to reduce the dirt and cosmic background were applied. Assumptions of the cosmic background removal using the CRT system and PDS to take advantage of the bucket structure were also applied in the analysis.

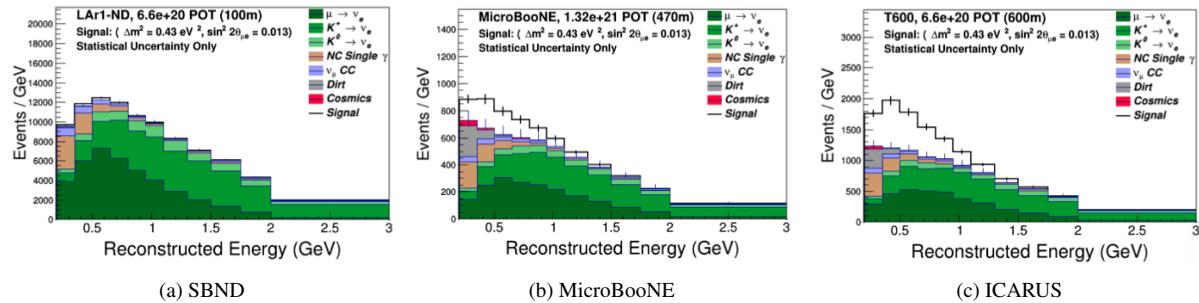


Fig. 4.18 SBN electron neutrino spectra after the  $\nu_e$  CC inclusive truth-based selection. From Reference [17].

A further 94% of photon events were removed, assuming this was the separational power of a  $dE/dx$  cut. Calculating the median  $dE/dx$  of the initial track stub of the shower is a powerful tool to remove photon events and is discussed in Chapter 6, Section 6.2.2. The intrinsic  $\nu_e$  beam events are an irreducible background. Other rare events such as electron scattering and neutral pion Dalitz decays are also backgrounds to the analysis. A  $dE/dx$  cut can not remove such events as the showering particle is an electron.

Furthermore, events where two showers were present were removed, assuming these were the result of  $\pi^0$  decays arising from resonant  $\nu_\mu$  events. Additionally, events where muons which travelled further than 1 m in CC  $\nu_\mu$  events were identified and removed. Finally, showers below 200 MeV shower were removed. A detailed description of the selection procedure is in Chapter 7, Section 7.2. The sensitivity to the sterile oscillation parameters can be seen in Figure 4.19 for the proposal truth-based  $\nu_e$  analysis. As can be seen in this Figure, the analysis can confirm or refute to  $5\sigma$  the  $\nu_e$  appearance seen by the LSND.

In Chapter 7 the above procedure for the  $\nu_e$  selection is re-performed on a more recent Monte Carlo sample. This was performed in order to evaluate the changes in sensitivity of the programme due to changes in the physics models and changes in the programme. Furthermore, additional metrics and methods, presented in Chapter 7, were designed to improve the selection and thus the sensitivity of the SBN programme. The updated sensitivities using truth information and reconstructed information from simulated detector responses are presented in Chapter 8.

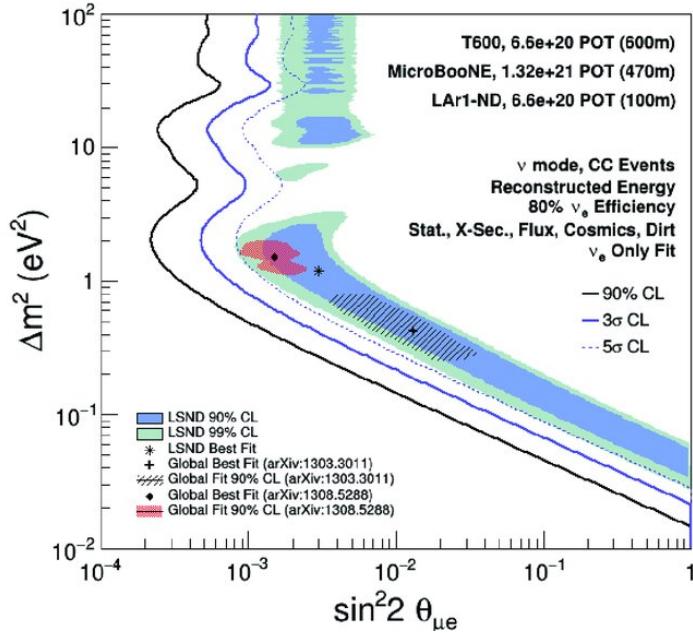


Fig. 4.19 The sensitivity of the SBN programme to electron neutrino appearance signal for the proposal truth-based analysis. The Figure shows the allowed regions for LSND and some global best fits at the time. Area to the right of the contour has a higher confidence value than the contour value. See Chapter 8 for further detail. From Reference [17].

#### 4.6.2 Other Physics Searches

Due to the high number of neutrino interactions in the SBN programme, the programme is also capable of performing high-statistics cross-section measurements for neutrino interactions in the BNB energy range. This will include inclusive cross-section measurements as well as dedicated channel cross-section measurements. Rare events such as electron scattering can also be studied, see Reference [197]. All the interaction studies will prove useful for future experiments, such as DUNE. The SBN programme will also be able to search for sub-GeV dark matter in decay products of the BNB during a "beam-dump" mode.

### 4.7 Software Used in the SBN Programme Simulation

The SBN programme uses the common C++/ROOT [222] Art [223] framework to process and simulate events. A common framework for liquid argon technology, LArSoft [224], is used within the Art framework to provide more specific support to the coding base of the experiments. This enables the collaboration to keep up to date with the latest developments within the field and easily share common simulation and reconstruction software.

The neutrino interactions are simulated using the Monte Carlo GENIE [225] (Generates Events for Neutrino Interaction Experiments) software. GENIE is common in many neutrino experiments and uses data from existing experiments to tune physical models. GENIE accounts for various theoretical models allowing for a range of scattering mechanisms within the BNB energy range. GENIE also models the creation of the Final State Interaction particles (FSI). The SBN programme simulates Electron Scattering, Quasi-Elastic scattering, Resonant, Meson Exchange Current (MEC), Deep Inelastic Scattering (DIS) and Coherent Scattering events. See Chapter 3, Section 3.6 for information on the neutrino cross-section model. GENIE uses flux information from MiniBooNE and the total neutrino cross-sections to simulate the correct number of interactions. GENIE then uses the cross-section model to simulate the correct number of each interaction type. At the few GeV energy range, GENIE uses the impulse approximation and the relativistic Fermi gas model to simulate the nuclear effects during the interactions. Pauli-blocking is also applied to Quasi-Elastic events [157].

In addition to the neutrino interaction events generated by GENIE, there is also events arising from cosmic-ray interactions. This the SBN detectors are on the surface of the earth, and therefore cosmic events regularly occur, whilst energy depositions from neutrino interactions drift to the readout. Cosmic events can mimic neutrino interactions and are backgrounds in analyses. As well as this, cosmic interactions can overlap with neutrino interactions. This makes reconstruction of events more difficult. Therefore, in the simulation, neutrino events are overlaid with cosmic events using CORSIKA [226] (COsmic Ray SImulations for KAscade). CORSIKA generates high energy cosmic ray particles and propagates the particles through the atmosphere to the detector until an interaction or decay occurs. CORSIKA uses several models, described in Reference [227], to generate the cosmic rays.

The FSI particles from neutrino interactions and cosmic particles are then propagated in the simulation by Geant4 [228–230]. Geant4 handles the energy depositions within the detector as well as particle propagation and particle decay. Then the simulation calculates the deposited charge which arrives at the wires using the modified box model to account for recombination. The simulation uses models of diffusion and electron lifetime to account for losses from the deposition region to the wires. The detector response to the charge depositions is then calculated using a convolution of the signal, detector and field responses which are parameterised by a unit charge in the MicroBooNE simulation.

The scintillation light from recombination is propagated to the PDS using a semi-analytical method which accounts for Rayleigh Scattering.

Alternative particle generators are available. There is the ability to generate specific particles with known characteristics in the simulation. This is useful to provide high statistics of specific interactions which are useful for reconstruction development, see Chapter 6.

# Chapter 5

## Studies on the SBND Electronics Chain

Prior to the installation of the Short-Baseline Near Detector (SBND) detector, extensive efforts were made to evaluate the performance of the electronics chain. One specific test of the performance of the electronics is known as the Vertical Slice Test (VST). This was performed to evaluate the entire electronic readout in a test stand Liquid Argon Time Projection Chamber (LArTPC). The VST experiment was housed in the 550 L LArIAT LArTPC at the Fermilab Test Beam Facility. The experimental setup used will be discussed in Section 5.1. To evaluate the performance of the electronics, one of the primary aims of the VST was to perform a signal to noise ratio measurement to ensure the SBND electronics meet the requirements for the physics goals of the collaboration [17]. This Chapter will discuss the details of the signal to noise ratio measurement in Section 5.2. Furthermore, to ensure the electronics meet the physics goals of SBND, the signal to noise ratio measurement was extrapolated to account for the differences between SBND and the VST. This extrapolation procedure is discussed in Section 5.3.

In addition to the testing of the SBND electronics, a hybrid readout system of the SBND front end and the ICARUS warm electronics was tested at CERN prior to the VST. This electronics chain was thought of as a potential backup for the Cold Commercial Off The Shelf (COTS) ADCs [210] SBND has implemented. A signal to noise ratio analysis was undertaken for a cosmic muon sample taken by the test stand to evaluate the performance. This analysis will be discussed in Section 5.4.

In order to calculate the signal to noise ratio of these test stands, an electron lifetime analysis is required. Prior to the signal to noise ratio analyses above, three electron lifetime analysis measurements were studied using the SBND detector simulation. These methods were used in the signal to noise ratio measurements and online monitoring of the VST. The methods are also to be employed in SBND software. These methods will be discussed in

Section 5.5, along with the reasoning for choosing specific methods in the signal to noise ratio analysis.

## 5.1 VST Overview

The test of the SBND electronics chain, known as the Vertical Slice Test (VST), used the LArIAT [231, 232] detector. This Section describes the experimental set-up and the software used during the VST.

Shown in Figure 5.1 is a schematic of the beamline used for the VST programme. This beam facility consists of several sub-detectors which are used to identify charged particles that enter the 550 L LArTPC. The beam itself is made by impinging a pion-proton beam, with energies between 8-32 GeV, onto a copper target. This produces a mixture of charged particles [231]. The momentum and composition of the generated beam of charged particles is altered by a configurable dipole magnet. The momentum of the beam entering the TPC is in the range 0.2-2 GeV [231]. This is an ideal energy range to perform a comparative study on known particles with energies similar to that of charged particles in neutrino interactions from the BNB. Upstream of the detector is a set of multi-wire proportional chambers which measure the particles' momenta. There is also a time of flight measurement system, a collimator, a muon range stack and a Cherenkov detector which help to identify the flavour of particles. The beam, triggers and subdetectors were used during the first run and part of the second run of the VST.

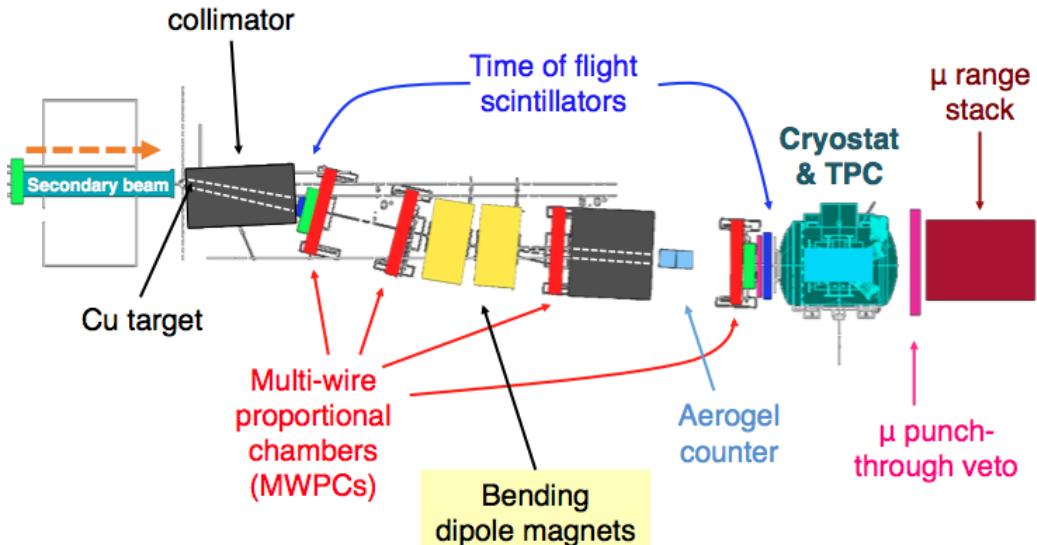


Fig. 5.1 The LArIAT tertiary beamline with the sub-detector and TPC elements. From Reference [24].

In addition to the beamline sub-detectors, a cosmic ray tagging system was used during the VST runs in order to obtain a sample of cosmic events [232]. The system is made up of two towers, each including two horizontal (horizontal, veto up/down) and two vertical (trig up/down, vert 1,2) scintillator paddles. The scintillation light is detected by PMTs. Refer to Figure 5.2 for a schematic of this system. There are two scenarios whereby a trigger is passed to the trigger board: 1) vert 1 and vert 2 are triggered and only one of the horizontal boards are triggered; or 2) trig up and trig down are triggered, but neither of the veto up or veto down paddles is triggered. The veto boards exist to remove cosmic showering events. The triggering is determined by a bespoke logic nim board which combines the signals from the PMTs [232]. This triggering process also ensures that the cosmic particles pass through the TPC within a finite angular distribution. The Figures 5.3a and 5.3b show the distribution of the azimuthal and zenith angles of the cosmic triggered events during the second VST run. Cosmic events in the second run of the VST were used in the signal to noise ratio analysis as they provide the minimum ionisation deposition in liquid argon.

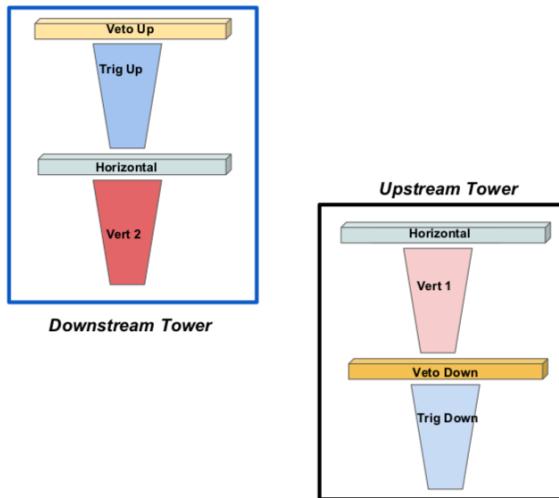


Fig. 5.2 The LArIAT cosmic muon tagging system. From an internal presentation, credit: J Asaadi.

The LArIAT LArTPC has a 170 l active volume ( $90 \text{ l} \times 47 \text{ w} \times 40 \text{ h cm}^3$ ) and a drift distance of 47 cm [24]. This drift distance requires 23.5 kV at the cathode for a 500 V/cm electric field. The TPC has two readout planes  $\pm 60^\circ$  from the vertical and one shield plane of vertical wires. For the VST, planes of 240 wires with a 4 mm wire spacing were used. The readout signal from the wires was passed via a tailored adapter board to the cold electronics on the Front End MotherBoard (FEMB). Each FEMB consisted of a set of LArASICs which amplified and integrated the signal. The integrated signal was then sent directly to a COTS 12-bit ADC, with a 2 MHz sampling frequency. The digitized output was then multiplexed

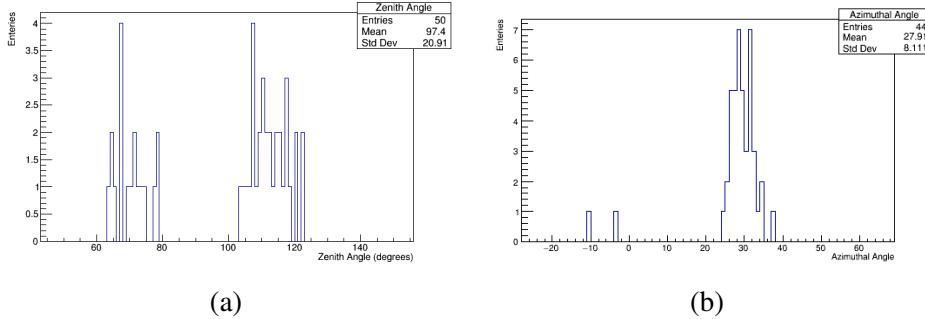


Fig. 5.3 The azimuthal (a) and zenith (b) angle distributions arising from events triggered from the cosmic muon tagging system.

and sent via miniSAS cables to the Warm Interface Board (WIB). The WIB also managed the power supply to the FEMBs, the input triggers as well as the gain and shaping time settings for the LArASICS. It should be noted that due to the addition of the FEMBs and the adapter boards, the distance from the high voltage feedthrough and the cold electronics was 7 cm. The FEMBs and the COTS ADCs were an addition to the existing LArIAT front end setup and therefore the 7 cm gap was significantly larger than in the LArIAT setup. An electrostatic shield made of copper was placed between the electronics and the feedthrough to prevent damage in the event of a high voltage breakdown.

As well as testing the SBND electronic readout, newly developed software was tested. New DAQ software called ArtDAQ v3.02.01 [233, 234], which is managed by the DAQInterface v03.02.01, was used [235]. ArtDAQ and DAQInterface are software packages used by multiple experiments and are maintained by the Art term at Fermilab. A new slow control software that is to be used in SBND was also tested: the Experiment Physical and Industrial Control System (EPICS) [236]. The EPICS system controlled the wire bias power supply, the TPC electronics power supply and the power distribution unit in the DAQ rack.

For the VST, online monitoring software was designed and ran during the data acquisition or post-acquisition. Due to issues with the DAQInterface, the binary data files were stored on the DAQ server and then separately converted into Art-Root human-readable files. The online monitoring then performed fast analyses of the following detector metrics: noise RMS, baseline, hit occupancy, hit peak height and electron lifetime. More detail about the hit finding and the peak height analysis can be found in Section 5.2. The method used to calculate the electron lifetime was based on method one described in Section 5.5. More information on the specific VST electron lifetime analysis can be found in Section 5.2.

The metrics were sent to a Redis [237] server where several averages of the metrics within set time periods were stored. The online monitor was designed to work within the 10-20 Hz peak rates that were expected from the configured beam settings. A wire-noise

covariance matrix and an FFT analysis were also calculated per event. A website based on the online monitoring of the SNO+ experiment [238] would then interact with the database via the Redis API and present the data in a user-friendly format. All metrics could then be accessed as an average value over a given time length of 1 s, 10 s or 100 s. As is shown in Figure 5.4, metrics were presented as cubisms [25] which are cubes that present how the averages of the metrics changed over time. Additionally, one could see the averages for each wire, ASIC chip or FEMB. The online monitor for the VST has now been used to develop the monitoring architecture for SBND.

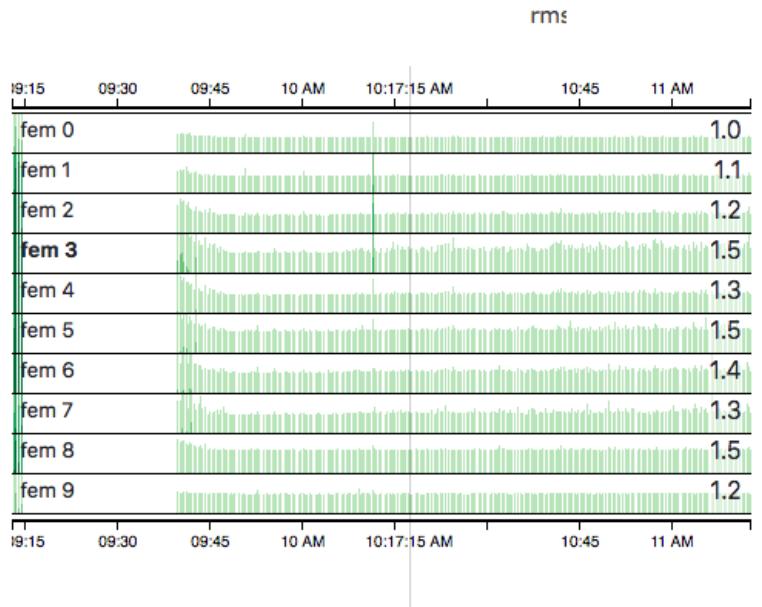


Fig. 5.4 Online monitor cubisms [25] for the noise RMS for the first 8 channels. The current RMS value is shown on the right hand side of the cube in units of ADC. The rest of the cube shows how the noise RMS has changed with time.

In addition to new software, existing LArIAT [232] software was adapted for the VST. The LArIAT DAQ software was used to manage the acquisition of the sub-detectors data. An existing PostgreSQL [239] manager for slow control data was edited for the VST and recorded over 100 million measurements. The slow control server was useful during the second run as it identified when the cryostat was being filled with liquid argon, via evaluating the argon flow valve status, which created high levels of noise. Archiver software was also reformatted for the VST data. This archiver software included a set of python scripts which monitored the DAQ directory for new files and created meta-data for the files, such as a timestamp, a checksum and run/sub-run numbers. This meta-data was then stored in another PostgreSQL database. Studies of the archiver tools were undertaken to check the viability of the software for the SBN programme. The study showed the software would achieve a

rate above the rate required for beam events in SBND: 0.22 Hz. The meta-data was then used in a SAM [240] (Sequential Access via Metadata) .json file for storage onto the tape. The storage was controlled by the File Transfer System (FTS) [241] software, maintained by the Fermilab computing division. After the data was stored onto tape, it was then copied to a local disk where an event display would loop over the events. The event display was a powerful tool and clearly showed when the argon was filling. This is demonstrated in Figure 5.5, where the noise (the vertical lines) on the wire-tick readout and shift in baseline in the waveform only occurs during filling.

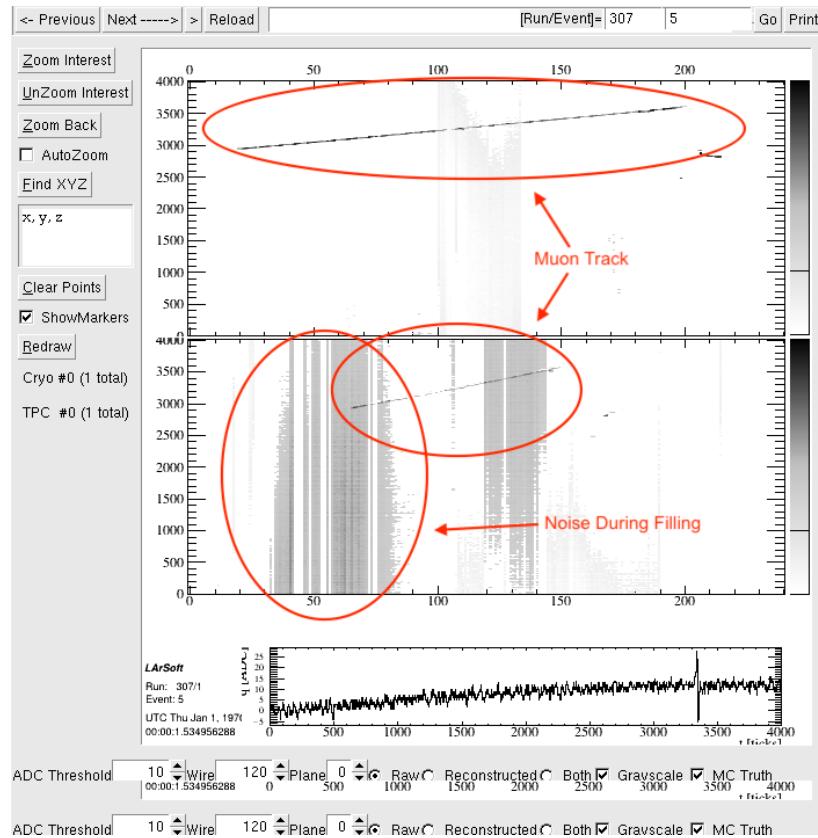


Fig. 5.5 A triggered event whilst the liquid argon was filling during the second run of the VST. The top two images show the charge deposition in wire-time space. The bottom is the waveform on a particular wire.

The VST was carried out on two runs: the first was with cosmic and beam data and the second took just cosmic data. Both lasted two weeks. The second run included many successful sub-runs with varying electric field strengths totalling  $\sim 13$  K events. A signal to noise ratio analysis was undertaken on the second run for cathode voltages of 15 kV, 17 kV and 20 kV. When the cathode voltage was set to 23.5 kV, the nominal drift field, several of the FEMBs were damaged irretrievably. This was due to the small distance between the

cathode feed-through and the electronics boards. This will not be an issue for SBND as the geometry of the system is different. In addition, the electronics have performed successfully in other tests. The two remaining FEMBs were at opposite ends of the TPC, on the two different planes. After the FEMBs were damaged, due to geometry of the cosmic tagging system, track reconstruction was not possible, and therefore the signal to noise ratio analysis was not possible. The FEMBs were recovered once the TPC reached room temperature. The next Sections will discuss the signal to noise ratio analysis as well as parts of the online monitoring relating to it.

## 5.2 Signal to Noise Ratio Measurement in the VST

The signal to noise ratio offline analysis was undertaken on the second run data where successful stable runs were taken. An incorrect wiring of the wire plane voltage biased meant the first run could not be used in this analysis. The signal is defined as the most probable peak height of a waveform from a Minimum Ionising Particle (MIP) muon which is travelling parallel to the wire plane and perpendicular to the wires, across all wires. This ensures the signal is defined as the smallest possible signal expected in the TPC. Corrections are applied such that the signal is defined as a particle which is traversing the TPC at the anode. This is to remove the effects which occur during transportation of the charge depositions and therefore is independent of the detector argon purity. To obtain the signal to noise ratio, the signal is divided by the RMS noise of the wires. The analysis used cosmic events triggered via the cosmic muon paddles.

In order to perform the signal to noise ratio analysis, reconstruction is required. An example of the first stages of the reconstruction is shown in Figure 5.6. Firstly, charge depositions need to be identified in the waveform. Once identified, these depositions are known as hits, highlighted in the Figure as red Gaussians. After the hits are identified they can be combined into collections, known as 2D clusters, in the wire-tick planes. The Figure shows an example of a cluster in green. The clusters combine hits which are topologically related, e.g. a straight line. Finally, a 3D tracking reconstruction is performed, where the 2D clusters in the wire planes are matched together using the timing information.

This hit finding was performed by the RawHitFinder [242] and is discussed in Section 5.2.1. The LineCluster module from LArSoft [243] performed the 2D clustering and the 3D matching was applied using the LArSoft module PMAlgTrackMaker [244]. To maximise the clustering performance, the preexisting LArIAT tuning of the modules was applied.

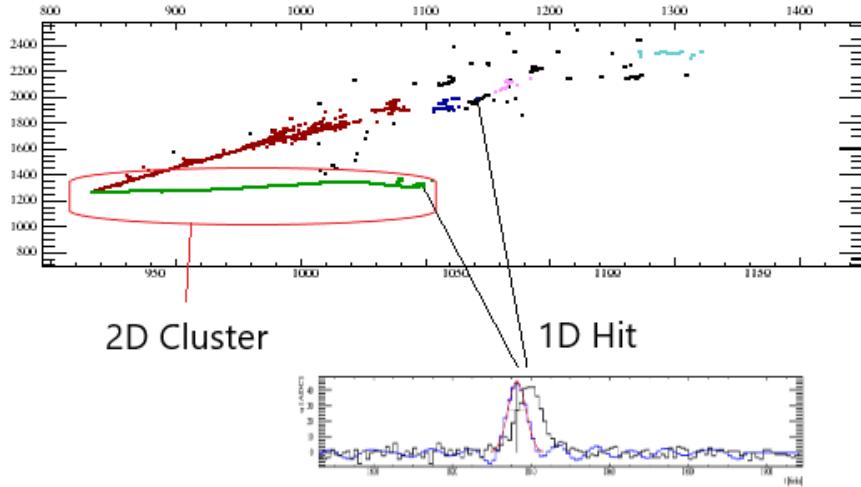


Fig. 5.6 Example of the reconstruction process. Hits (black and coloured dots) are identified, and presented in the wire-tick (wire-time) space. Hits are then clustered together based on their topology (coloured dots). An event such as this would be removed during the signal to noise ratio analysis.

Once the particles were reconstructed an electron lifetime analysis was performed. This is to ensure the signal strength is not diminished by impurities and thus only the performance of the electronics is evaluated in the signal to noise ratio analysis.

In order to isolate MIP events for the signal to noise analysis ratio, a selection procedure was undertaken. This is outlined in Section 5.2.3. The main selection goals are to remove showering particles, such as in Figure 5.6, and events where the particle does not cross the TPC. This is to ensure that MIP-like muons, such as in Figure 5.7, are the only particles being assessed. From the reconstruction, the most probable peak height of the charge deposits can be extracted and the electron lifetime correction can be applied, as well as corrections to account for differences in the trajectories from the defined signal above. The noise RMS can then be calculated from the waveform outside of the signal regions, and the signal to noise ratio can be defined.

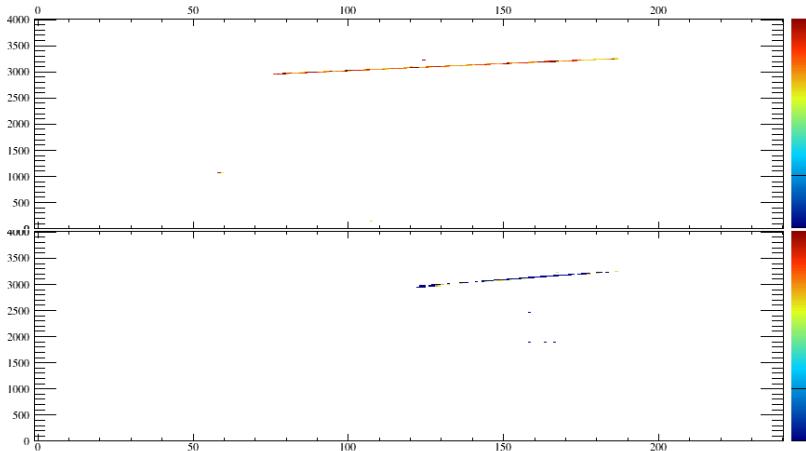


Fig. 5.7 Event display of raw data from the VST showing a through-going cosmic muon. Hits on the waveforms are visible and the 2D cluster in each of the views can be easily performed. There is an offset of 2560 ticks in the readout.

### 5.2.1 Raw Hit Finding

Hit finding is the first step towards the signal to noise ratio analysis. The hit finding described here is the process of finding charge depositions, known as hits, in the waveform. An example of such a charge deposition is shown in the simulated event display in Figure 5.6. One algorithm that is used to find such a deposition is the RawHitFinder.

The RawHitFinder [242] is an existing LArSoft module designed originally for the 35t experiment [179]. In addition to the signal to noise ratio analysis, it was also implemented in the online monitoring to receive hits for the electron lifetime online analysis. The hit finder was also used to estimate the average peak heights of identified hits and the occupancy of hits. Initially, the RawHitFinder searches for a point on the waveform greater than the threshold. Once this occurs, the charge is integrated by adding the ticks until a point on the waveform below the threshold is reached. This method is fast and ideal for online monitoring purposes.

With some minor adjustments and tuning of the induction and collection thresholds to 5 ADC and 10 ADC respectively, the module was successfully used in the VST. Figure 5.8 shows an event where the raw hit finder was implemented. The black points are hits identified by the raw hit finder.

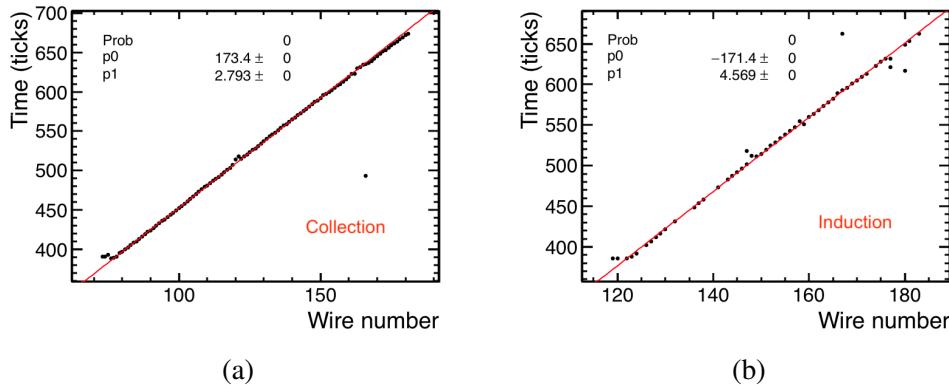


Fig. 5.8 The output of the raw hit finder for the collection (a) and induction (b) planes with initial linear fits performed. Credit: Tom Brooks

### 5.2.2 Electron Lifetime Analysis

As described in Chapter 3, the electron lifetime describes the purity of the liquid argon. The amount of charge that reaches the wires decreases with the drift time due to electronegative impurities in the argon. The losses can be measured by evaluating the exponential Equation 3.12 and extracting the electron lifetime as a fit parameter. This calculation requires various most probable charge deposition measurements with differing drift times. Once the electron lifetime is calculated, the losses due to impurities can be corrected for and the signal to noise ratio of the electronics can be defined without this detector dependent effect. The electron lifetime is also a useful measure of the status of the detector and can indicate purification system failures. This Section describes the electron lifetime analysis used in the signal to noise ratio and its use in the online monitoring.

The monitor used cosmic muons, like the event in Figure 5.7, that triggered the muon paddles. The analysis starts with the hits identified on the collection and induction wire planes using the raw hit finder as shown in Figure 5.8. Using the hits, a recursive linear fit was performed to identify hits from the cosmic track. Hits further than 30, 10 and finally 3 ticks from the fit, were removed to improve the fit. This helped to remove noise hits and hits from delta-rays, which are electrons with energies from the Landau tail which ionise additional particles. Cutting on the  $\chi^2/\text{ndof}$  of the linear fit removed non-track like events. An example of selected tracks can be seen in Figure 5.9. Cuts on the total number of remaining hits, the size of the track and ensuring a time overlap between the collection and induction tracks improved the lifetime measurement. The cuts used in this analysis are summarised in Table 5.1.

Cut	Value
Total hits	60
Collection hits	40
Induction hits	20
Max $\chi^2/\text{ndof}$	40
Wire extent	50
Time extent (ticks)	100
Min collection/induction track overlap (%)	80
Max lifetime from exponential fit (ms)	100

Table 5.1 Cut values used to select straight through-going cosmic muon tracks. Credit: Tom Brooks

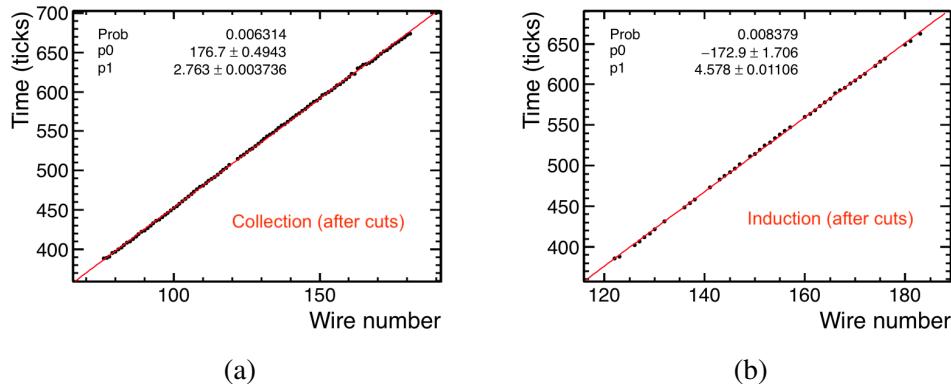
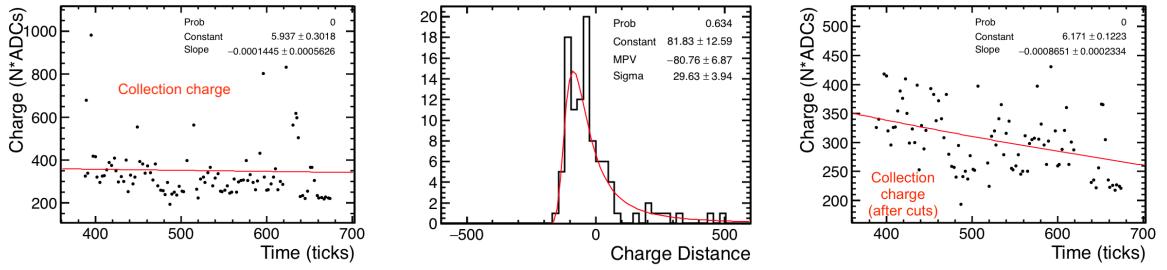


Fig. 5.9 Tracks on the collection (a) and induction (b) planes after selection cuts. The cuts were chosen by evaluating the results on a sub-sample of the data. Credit: Tom Brooks

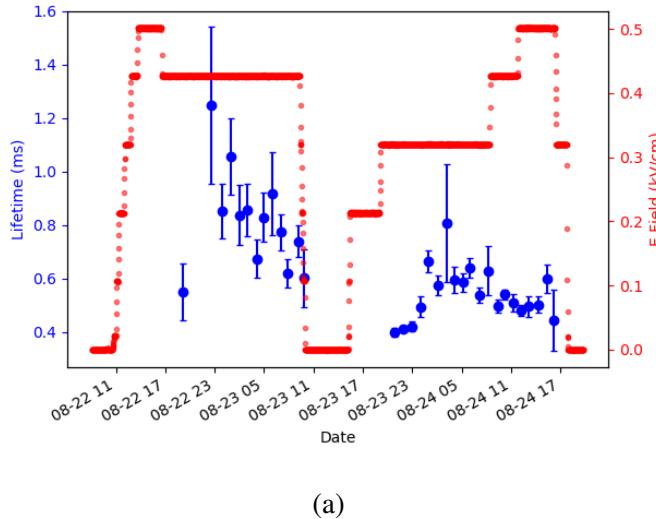
For tracks which pass the cuts, the hit charge against time is fitted with an exponential, as in Figure 5.10a. Hits which have a charge too far away from this fit are removed using a Landau fit on the hit charges, which is shown in Figure 5.10b. A log-likelihood minimization is then performed to extract the lifetime, where the charge vs time is modelled as an exponential modified by a Landau in charge. An example of the hits used in this minimization is shown in Figure 5.10c.

The purity calculation ran stably as part of the online monitoring presenting result on the website. The website would average the purity in a given time to present the result. Once the FEMBs died, track reconstruction could no longer be performed. All of the successful measurements, were collected and averaged over long periods to reduce variation in the results. Figure 5.11a shows this reduction in variation.



(a) Initial exponential fit of charge vs time.  
(b) Landau fit of charge histogram.  
(c) Charge vs time after Landau cut.

Fig. 5.10 Cuts performed on the charge vs time to select hits for the minimization. Credit: Tom Brooks



(a)

Fig. 5.11 All lifetime measurements (a) and the hour average with the changes to the drift field (b). Credit: Tom Brooks

The averages were then used for the signal to noise ratio analysis. As can be seen in Figure 5.11a, these values ranged from 0.4 ms to 1.25 ms. When the cathode is at full voltage a 1 ms lifetime results in a ~30% loss in charge from depositions which occur at the cathode (corresponding to a drift time of 300  $\mu$ s).

### 5.2.3 Signal to Noise Ratio Analysis

Once data taking was completed the signal to noise ratio analysis was performed offline. The basis of the analysis was the 3D reconstruction briefly mentioned above. For more detail on

reconstruction techniques see Chapter 6. The analysis is split into two measurements: the noise measurement and the signal measurement. The noise measurement will be discussed first. The signal measurement will then be described. This involves identifying the MIP events within the reconstruction and applying lifetime corrections as determined by the means described above. Direction corrections are also required and will be discussed. The signal is then extracted as the most probable peak height of charge depositions for the corrected MIP events. The extraction was completed by performing a Landau-Gaussian fit. The signal to noise ratio was defined as the signal divided by the noise RMS.

For the noise analysis, linear fits of the raw hits associated to the reconstructed 3D track in the wire vs time frames in the collection and induction planes were undertaken. Raw data that is within 20 ticks of any of the fitted lines in an event was removed from the noise RMS analysis to ensure no signal information was processed. Furthermore, the events that occurred during argon filling and up to 15 minutes after were excluded. Cuts were applied to the events to remove any showering particles. If the event had above 1000 hits or the density of hits per wire was above 1.1, the event was removed. All cuts in the analysis were chosen by evaluating the best parameters for a sub-sample of events.

The first 2000 ticks of the waveform were used to calculate the noise RMS for the event. This data corresponds to pretrigger information as there was an offset of 2560 ticks. This does not introduce a bias unless there is the presence of an interaction in the TPC. Such events are identified in the reconstruction and the fraction of the waveform in the regions that correspond to the event is not used. As the TPC contains several wires of varying size only the largest wires (length of 46 cm) were used in the analysis. As the largest wires have largest capacitance and hence have a largest noise RMS the result is conservative. The RMS for each sub-run before run 343 was calculated by fitting a Gaussian around the mean noise RMS values of the largest wires and taking the mean of those values as the RMS of the run. After sub-run 343, only 2 FEMBs were functioning. As a result of this, the mean RMS was defined as the mean RMS of all the wires for all events in the run. The mean noise RMS for the evaluated runs can be found in Figure 5.13 and as a function of the wire identification number (ID) for sub-run 322 in Figure 5.12. Due to channel mapping and the fact the readout wires go across the TPC frame at an angle, the wire ID of 1 corresponds to the smallest wire (10 cm). The wires then increase in size until the central longest wires are reached then the length decreases again as the wires get closer to the opposite corner. Hence the form of Figure 5.12 is due to the change in wire size as larger wires have a larger noise RMS. It also explains the decrease after sub-run 343 in Figure 5.13 where the smaller wires are used.

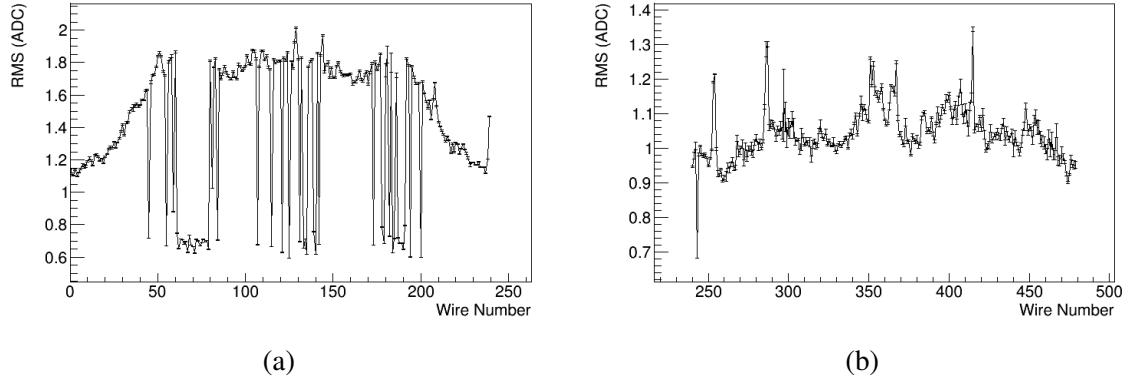


Fig. 5.12 RMS of each wire for sub-run 322 for the induction (a) and collection (b) planes. The error bars are the standard error on the mean. Dead wires, which can be noticed by a sudden drop in RMS in the distribution, were removed from the analysis. Only the longest wires, where the RMS is the largest, were used in the run RMS analysis. Errors in y correspond to the error on the mean value.

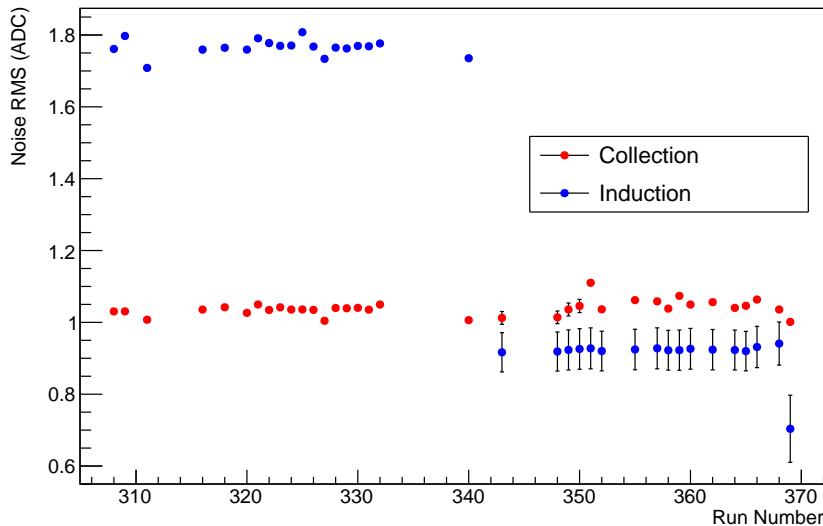


Fig. 5.13 The noise RMS in ADC for several sub-runs in the second run of the VST. The error bars are the standard error on the mean. The drop in the RMS in the induction plane is the result of the remaining FEMB readout wires existing at the end of the induction plane where they are smaller in length. Errors in y correspond to the error on the mean value.

For the signal analysis, further cuts were applied to events. A Principal Component Analysis (PCA) [245] was undertaken on all the hits in the event. The PCA analysis consists of an orthogonal transformation whereby the first principal component has the largest possible variance on the data. Hence a cut was applied on the first Principal component [246]. If

the Log10(1-FirstEigenvalue) of the event was less than 1.7, the event was assumed to be a shower and removed. Cuts on the track length ( $L < 60$  cm) and the number of hits (<50) were also applied. The two cuts remove stopping muons, which are not minimum ionising.

A cut on the azimuthal angle (angle into the wire plane) was also applied. This cut removed events where charge is lost due to the length of the shaping time of the pre-amplifier being less than the time of the hit. In addition, due to the geometry of the cosmic muon panels, the zenith angle of tracks was a bi-modal distribution. Therefore, one part of the distribution is nearly parallel to the collection wire direction and the other is nearly perpendicular. For the collection plane analysis, events were removed where the zenith angle was greater than 90 degrees. For the induction plane analysis, events where the zenith angle was less than 90 degrees were removed. These cuts ensured that events where the track travelled in parallel with the wires were removed. Such events have poor hit finding efficiency, track reconstruction and charge reconstruction. This is because the charge on the event is deposited on only a few wires, making the hits large. The time of these hits is then larger than the shaping time of the ASIC, meaning charge is lost. Finally, as only a few wires are hit, tracking reconstruction is difficult as the 2D track projection holds little information of the 3D track. The cuts were chosen by performing a brief handscan of events to ensure the selected events were tracks. As there was sufficient data remaining after the cuts the efficiencies of each cut was not evaluated.

Once a track has passed all the cuts, the peak values of the raw hits associated to each track in the event are then found. The peak heights were then normalised using the cosine of the azimuthal angle, the electron lifetime and the wire spacing divided by the track pitch [158]. The track pitch is the effective distance between two wires along the direction of the track. Combining these normalisations extrapolates the signal to that of a parallel muon at the anode. A Landau-Gaussian fit [26] was then applied to the normalised peak heights, which is shown in Figures 5.14. A diffusion normalisation was not considered for this analysis due to small drift distance.

The signal to noise ratio was then calculated for runs with significant statistics. The runs had different cathode voltages between 15-20 kV. As discussed in Chapter 3, this affects the amount of recombination and hence the signal strength in each run. The RMS for these runs can be found in Figure 5.12 and the signal to noise ratio can be found in Figure 5.15. The signal to noise ratio was calculated as  $57.55 \pm 0.31$  (stat) for the collection plane and  $15.58 \pm 0.31$  (stat) for the induction plane for the 20 kV runs. The statistical error is calculated by propagating error on the most probable value of the Landau-Gaussian signal distribution and the error on the mean of the noise RMS. The error on the fit is extracted using the MINOS

technique [26]. The technique finds the error on the by taking the difference of the most probably values for which the  $\chi^2$  is +1 from the minimised value.

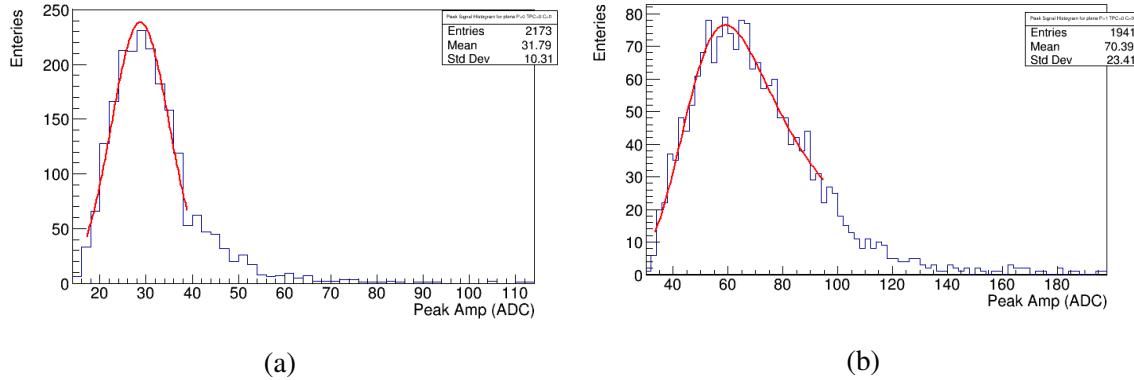


Fig. 5.14 The normalised peak heights distributions of tracks that pass the cuts for sub-run 322. The fit to the induction (a) and collection (b) distributions is a Landau-Gaussian fit.

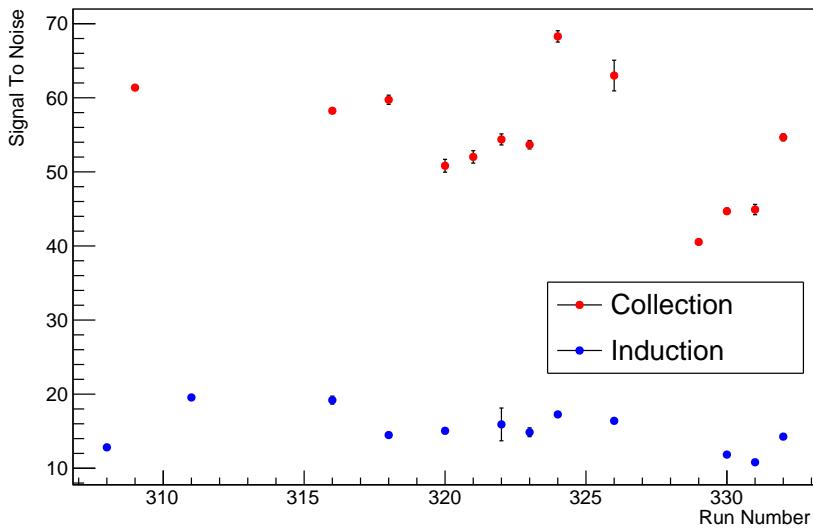


Fig. 5.15 The signal to noise ratio for several sub-runs in the second run of the VST. The error bars are the propagation of the statistical error on the Landau-Gaussian fit and the error on the noise RMS. Note that sub-runs 310-315 were held at the nominal cathode voltage of 23.5 kV, sub-runs 309,317-327 and 332 were held at 20 kV and sub-runs 308 and 329-331 were held at 15 kV. The errors in y are calculated by propagating the statistical errors on the noise RMS and signal value.

## 5.3 VST Signal to Noise Ratio Extrapolation to SBND

An essential use of the signal to noise ratio measurement of the VST is its comparison to the signal to noise ratio values required by the collaboration in order to achieve the physics goals of the SBN programme. The collaboration stated that the noise RMS must meet the requirement of being <700 ENC for the induction plane and <580 ENC for the collection plane. The Equivalent Noise Charge (ENC) is defined as the number of instantaneously collected electrons required so that their peak ADC count is equal to the root RMS of the noise measured [212]. The conversion factor from ENC to ADC can be theoretically found from knowing the ADC to voltage conversion and the gain of the ASIC. Thus theoretically the conversion factor is

$$1 \text{ ADC} = 1 \text{ ADC} \times \frac{2 \text{ V}}{4096 \text{ ADC}} \times \frac{1 \text{ fC}}{0.014 \text{ V}} \times \frac{6241 \text{ } e^-}{1 \text{ fC}} = 218e^-.$$

However, as we have shown in Section 5.4.5, the conversion factor can be calculated by comparing the theoretical charge deposited by a MIP in the detector that is identified on a wire and the reconstructed charge distribution of all the hits on the wires from a sample of MIPs, which is measured in ADC. The difference between reconstructed charge distribution and the signal peak height distribution, shown in Figure 5.14 is that the reconstructed charge distribution is proportional to the integral of the hits whereas peak height distribution is the peak value of the hits. This analysis resulted in a conversion factor of  $1 \text{ ADC} = 236 \pm 0.3 \text{ (stat)} \text{ ENC}$ . The statistical error is the result propagating the Landau-Gaussian fit error (calculated using the MINOS technique [26]). Although the theoretical value is not within the statistical error of the theoretical values losses due to diffusion and changes in recombination due to electric field are not considered. This however results in a higher ENC conversion factor and hence the final result is a conservative value. A systematical treatment should be considered in the future.

Equivalently, the ENC requirements correspond signal to noise ratio of 5/1 for the induction plane and 12/1 for the collection plane for an event anywhere in the TPC is also the benchmark set by the collaboration.

These requirements are essential because if the signal to noise ratio is too low, then hit finding becomes difficult as the classic threshold methods struggle to distinguish between noise and signal. If hits are not detected then particles are not identified which leads to the mis-identification of events, a reduction in efficiency and unfeasible physics analyses. This Section describes the methods used to extrapolate the VST results to the SBND detector design such that the VST signal to noise ratio electronics results can be compared to the SBND detector requirements.

In order to extrapolate the performance of the electronics to the size of SBND, a factor of 0.75 was applied to the signal size. This was to account for the difference in the wire pitch of both detectors. This is not applied to the noise RMS as only fewer drift electrons are seen per wire in the 3 mm SBND detector design compared to the 4 mm VST wire spacing. The inherent noise is independent of the spacing to first order. An additional correction factor to the signal was applied to account for the difference in the drift field strength, which affects the amount of recombination.

To calculate the size of the signal expected in SBND, the analysis also accounted for diffusion. This was undertaken by simulating 1000 MIP like muons, with and without diffusion effects. The events had the same distribution as that of the cosmic sample. The samples were simulated with an infinite electron lifetime such that there were no lifetime effects. The ratio between the signal (calculated using the signal analysis above) measured from simulated events, with diffusion and without diffusion, was taken as a normalisation term. This normalisation increased the signal size by 1%.

Additionally, the difference in wire size must be considered as this affects the noise RMS. To extrapolate the noise RMS to the wire lengths of SBND, the results from the MicroBooNE publication on noise levels [158] were used. In the paper, after employing software noise filtering to remove external noise sources, the noise RMS values are plotted against wire length. The noise contributions from internal electronics can be approximated as a contribution of the ASIC series noise and white parallel noise from other front-end readout components. SBND has the same form of noise contributions, and hence using this MicroBooNE result is justified. The noise in the paper was modelled using the following equation

$$\text{Noise}_{\text{RMS}} = \sqrt{x^2 + (y + zL)^2} \quad (5.1)$$

where  $x, y, z$  are fit parameters and  $L$  is the wire length. For large wire lengths this equation can be approximated as a linear slope,  $\text{Noise}_{\text{RMS}} = y + zL$ . The fit parameters to the MicroBooNE data were found to be  $x = 0.90$  ADC,  $y = 0.7$  ADC and  $z = 0.22$  ADC/m. Note that there is approximately a 10% deviation from the measured MicroBooNE value at wire lengths of 4 m using this approximation on the MicroBooNE data. Furthermore, this approximation assumes that the noise from external sources is independent of wire length.

The VST average noise levels were found to be  $398 \pm 34$  (stat) ENC for the induction plane and  $254 \pm 12$  (stat) ENC for the collection plane, before noise filtering using the noise RMS analysis described in Section 5.2.3. The error is calculated by propagating the error on the mean noise and the error on the ENC conversion value. It is assumed the noise-filtered MicroBooNE data models the noise from the VST wire capacitance and the SBND FE ASIC. Hence, for large wires, the gradient of the MicroBooNE data can be used to extrapolate the

VST noise RMS to wire lengths of the SBND size. However, as the size of the  $x$  term is unknown for the SBND electronics, the validity of the linear approximation at 4 m is not known. Therefore, the average noise RMS for the largest wires in the VST was used as the intercept of the linear approximation. As the largest wires have contributions from both the external noise and from noise due the wire capacitance (0.48 m), this ensures this is an upper limit for the noise RMS. The fact that the induction plane noise is larger than the collection plane for the same wire length suggests the presence of non-negligible external noise in the VST. In addition, placing the intercept as the RMS of the largest wires ensures the approximated RMS is higher than the true value when considering Equation 5.1. Figures 5.16a and 5.16b show the extrapolation of the noise RMS as a function of wire length using this method.

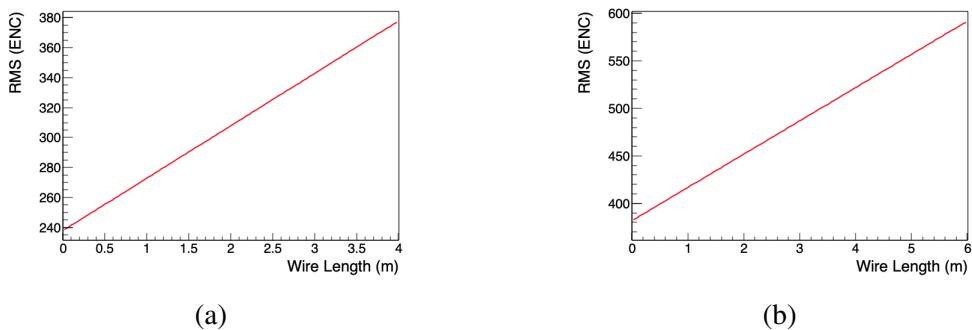


Fig. 5.16 The extrapolated noise RMS as a function of wire length in the collection plane (a) and induction plane (b).

Using the noise of the largest wire as the intercept from the VST and the gradient from the MicroBooNE data, the noise RMS in the SBND detector is then  $630 \pm 52$  (stat) ENC for the induction plane and  $412 \pm 36$  (stat) ENC for the collection plane where errors on the fit are not considered. Therefore, the cold electronics used in the VST meets the SBND noise requirement of <700 ENC for the induction plane and <580 ENC for the collection plane.

The SBND extrapolated signal to noise ratio was then calculated using the extrapolated noise and signal. This resulted in an expected signal to noise ratio measurement for an event at the anode for SBND. However, the SBND collaboration suggested the minimal signal to noise ratio must be met in all parts of the SBND detector. To consider the signal to noise ratio within the rest of the SBND detector, MIP like muons, travelling parallel to the wire plane, were simulated. A sample with diffusion and 3 ms lifetime was created and a sample without diffusion and an infinite lifetime was also created. This was done every 10 cm in the drift direction. The signal for both sets of simulated muons was then calculated using the same analysis above. The ratio of the two signals was then used to define the signal to noise

ratio at that specific drift time. The signal to noise ratio at a specified drift was calculated by multiplying the ratio at the specified drift time by the extrapolated signal to noise ratio at the anode. Figure 5.17 shows how lifetime and diffusion affects the signal size at a varying distance in the drift time. As can be seen in the Figure the requirements set by SBND of a signal to noise ratio of 12/1 on the collection plane wires and 5/1 of the induction plane wires is met across the entire drift region. This is true even before noise filtering.

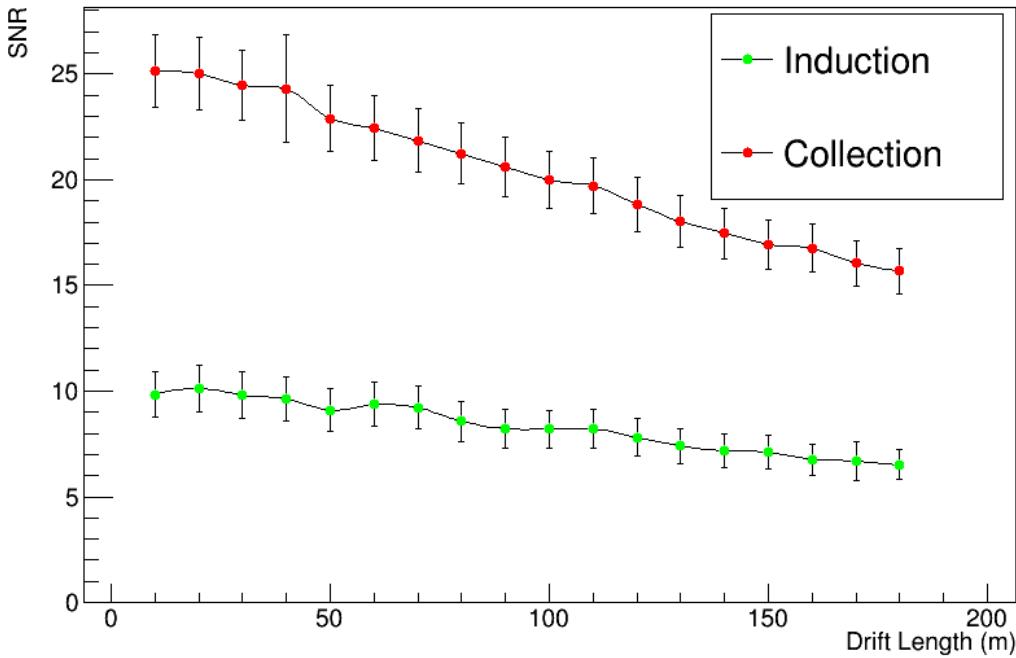


Fig. 5.17 The signal to noise ratio as a function of drift distance in the SBND detector. The value at a drift of 0 cm is the extrapolated value discussed above. The errors are calculated by propagating the error of the original SNR to the extrapolated value and thus equal to the original error scaled by lifetime and diffusion effects.

Therefore, both the noise RMS levels and the signal to noise ratio requirements set by the collaboration are met by the COTS ADC electronics.

## 5.4 Warm Icarus Electronics Alternative chain for SBND

A potential backup to the cold electronics ADC readout was an ad-hoc mixture of the SBND readout and the warm ICARUS electronics. This setup, which is discussed in Section 5.4.1, was undertaken in a test stand TPC at CERN using cosmic muons. The events underwent

a selection procedure to ensure MIP-like particles were used in the signal analysis. The selection is discussed in Section 5.4.2. From the selected events, a lifetime analysis was carried out in order to extrapolate back to the deposited signal. This removed detector effects and therefore the signal to noise ratio defines the performance of only the electronics. A noise study, discussed in Section 5.4.3, was also carried out to identify any abnormalities in the spectra. Using the selected events and the noise analysis, the signal to noise ratio of the system was calculated. This is shown in Section 5.4.4. Finally, an Electron Noise Count (ENC) to ADC conversion was also calculated to verify the procedure. This calculation is discussed in Section 5.4.5.

### 5.4.1 Test Stand Electronic Design

The test stand at CERN used for the assessment of the alternative electronics is a liquid argon TPC with a drift distance of 52 cm and a drift field of 500 V/cm. This corresponded to a drift time of 330  $\mu$ s. The anode consists of two perpendicular wire readout planes of 128 wires, with a wire spacing of 0.254 cm. Signals from the wire planes were readout using LArASIC chips [206, 209, 158]. The signal was passed to a CAEN A2795 [247], a 12 bit ADC with a full-scale range of 3.3 V rather than the COTS ADC. The signal to noise ratio was found for a run with a gain of 14 mV/fC and a shaping time of 3  $\mu$ s. With these particular LArASIC parameters, the ADC to electron conversion is

$$1 \text{ ADC} = 1 \text{ ADC} \frac{3.3 \text{ V}}{4096 \text{ ADC}} \frac{1 \text{ fC}}{0.014 \text{ V}} \frac{6241 e^-}{1 \text{ fC}} = 359 e^-.$$

Two scintillator counters, approximately 50 cm long and 10 cm wide, were placed on the sides of the chamber such that coincidence signals are detected when muons travel approximately 35° in the drift direction and with a zenith angle of approximately 135°. The DAQ was triggered using coincidence signals from the muon counters. The wires were given a capacitor to simulate a longer wire. Two induction wires given a larger capacitor. The first wire and the last two wires were disregarded in the analysis as the electric field is not well defined at these wires. The wires which were given a larger capacitance were also removed to ensure the results were not biased.

### 5.4.2 Event Selection

The peak signal to noise ratio was defined as the most probable peak charge deposited by a minimum ionising particle, travelling parallel to the wire plane and perpendicular to wire direction, divided by the RMS of the noise. Electron lifetime corrections were also

applied. This definition is the same that was used in the VST analysis. An event selection criterion was designed to obtain a sample of minimum ionising particles. At first, events were only processed when there was a coincidence in the triggering between the two cosmic ray scintillators. Hence most of the events cross the entire TPC.

Firstly, charge deposits over 15 ADCs from the software set baseline were labelled as hits. The 15 ADC cut off was chosen by considering the typical values of signals on the collection plane. The hits on the collection planes were then used in a Principal Component Analysis (PCA) [246, 245] cut. The PCA quantifies how dispersed the hits are with respect to the linear best fit. For track-like particles the first eigenvalue of the PCA transformation matrix is very close to 1. Therefore, a cut was applied to separate tracks and showers by defining an event as a track if the first eigenvalue  $E_1$  satisfied  $\log_{10}(1 - E_1) < -1.8$ . The value  $-1.8$  was chosen by evaluating a track/shower categorisation of a subsample of events.

Once tracks were identified, a hit cut was applied. If the event had less than 40 hits on the collection plane, it was assumed that it was a stopping particle or that the coincidence between cosmic ray scintillators was not due to one particle. In either case, the event was removed from the selection. This is because the TPC has 128 wires and, due to the orientation of the scintillators, the charged particles should traverse the entire detector. Therefore, generally, 128 hits above the 15 ADC threshold for MIP events were expected. The analysis was performed iteratively by identifying charge deposits, calculating the baselines of the waveform from the segments of the waveform and reidentifying hits. Events were removed if less than 70% of the wires on the collection plane were hit during the second iteration. A third hit cut is also applied that removes any events that have over 8000 hits. All cuts were chosen by evaluating a subsample of events and choosing parameters which removed the most showers whilst keeping the most tracks. Finally, if there are no hits on the induction plane the event is removed.

A linear fit to the hits was then applied to events which passed the selection cuts above. A typical event fit is shown in Figures 5.18a in the collection view and in Figure 5.18b for the induction plane. This fit was done using the robust regression fitting algorithm provided by the CERN ROOT package [26, 222] in the wire vs time plane. The algorithm uses at least 75% of the data to perform the fit. Data that was within  $40 \mu\text{s}$  of the best fit was then defined as signal and data  $4000 \mu\text{s}$  away from the line, corresponding to value outside the drift time, was defined as noise.

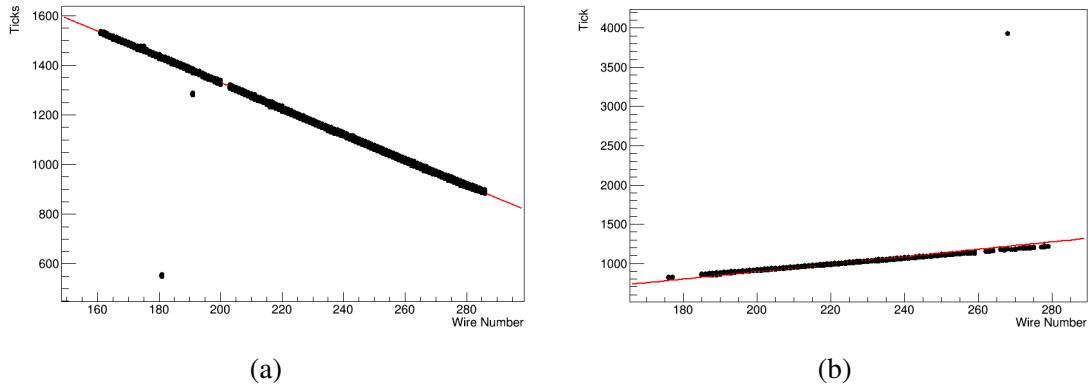


Fig. 5.18 The regression fit in the collection-tick plane (a) and the induction-tick plane (b).

The noise analysis described below is then performed to extract the pedestal of each wire. Once extracted, the selection procedure is performed iteratively and if the pedestal differs from the previous value by more than 1 ADC the selection procedure is repeated.

The selection resulted in a sample of MIP like muons which traversed the TPC. These events were used for the signal to noise ratio measurement. A typical event which passes the cuts is shown in the fit Figures 5.18a and Figure 5.18b.

### 5.4.3 Noise Characterisation

In parallel with the selection, which is required to define the signal element of the signal to noise ratio, the noise was also evaluated. The distribution that resulted from segments of the wire waveforms, which were  $4000 \mu s$  from the best fit line, were defined as the noise distributions. The noise distribution for the collection and induction planes for all wires can be seen in Figures 5.19a and 5.19b respectively. To define the noise RMS that is used in the signal to noise ratio analysis, a Gaussian fit on the noise distribution for each wire is performed. Values that are  $3\sigma$  from the mean of the fit are removed. This procedure removes data that do not correspond to electronic noise, e.g. radioactive decay. The fit is then repeated and the RMS of the fit is used as the noise RMS of the wire. The mean of the fit defines the pedestal of the wire. The noise RMS per wire can be seen in Figures 5.20a, 5.20b. For every wire, the RMS is always below 1 ADC for the collection plane.

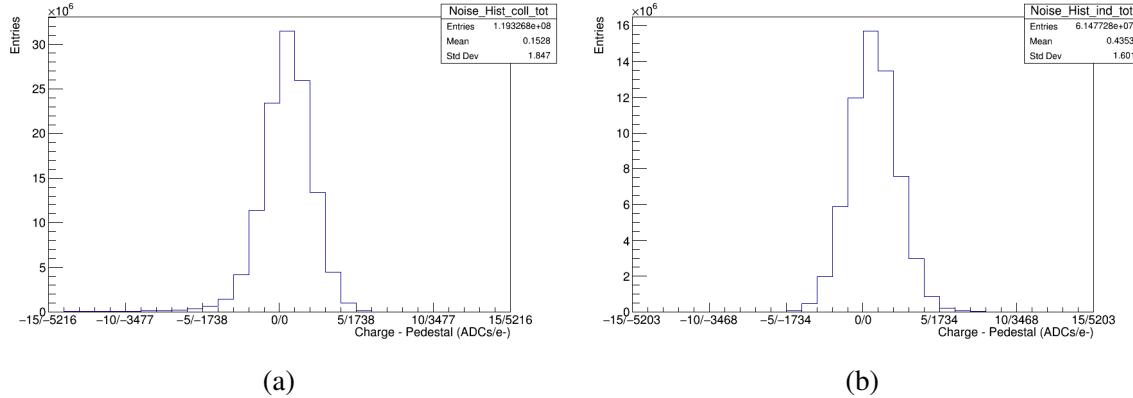


Fig. 5.19 The noise distribution for the collection plane (a) and induction plane (b). This corresponds to all ADC counts 2000 ticks away from the linear signal fit.

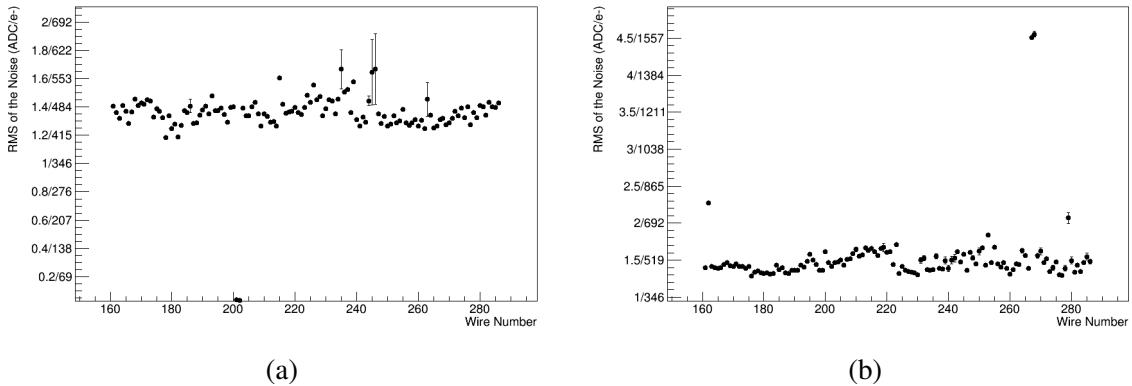


Fig. 5.20 The noise RMS for each wire the collection plane (a) and induction plane (b). Note: the two higher capacitance wires on the induction plane are not used in the analysis. The errors in y are the standard error on the mean values.

The noise RMS of the system is then defined as

$$N_{RMS} = \frac{1}{N_{events}} \sum_{events} \frac{\sum_j \sigma_j}{N_{wires}}, \quad (5.2)$$

where  $\sigma_j$  is the standard deviation of the wire  $j$  extracted from the second Gaussian fit,  $N_{wires}$  is the number of wires and  $N_{events}$  is the number of events.

Events with less than ten hits underwent a Fast Fourier Transform (FFT) analysis. For wires where there were no hits greater than 10 ADC, the FFT was then calculated using the CERN ROOT TVirtual FFT class [248]. An average of the frequency spectrum was then

taken. Figure 5.21 shows a typical wire frequency distribution. Note that there is a cut off at 2 MHz in the spectrum as this is the limiting frequency of the DAQ.

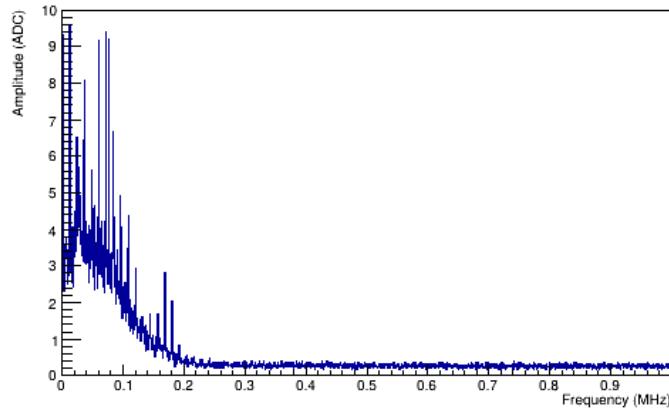


Fig. 5.21 The FFT for a wire on the collection plane.

#### 5.4.4 Signal to Noise Ratio Measurement

Using the selection above to calculate the signal strength and the noise RMS, in Figures 5.20a and 5.20b, the signal to noise ratio can be defined. As defined above, the signal to noise ratio is the most probable peak charge deposited by a minimum ionising particle travelling parallel to the wire plane and perpendicular to wire direction divided by the RMS of the noise. MIPs that follow the signal to noise ratio criteria are the smallest signals that are detected.

From the signal data, the peak charge was extracted for each wire. To prevent charge depositions that were longer in time than the shaping time of the ASIC, events are removed where the angle between the parallel direction and the direction of the track is greater than  $\tan^{-1}(\frac{\tau v}{0.254})$ , where  $\tau$  is the shaping time of pre-amplifier, 0.254 is the wire spacing and  $v = 1.5 \text{ mm } \mu\text{s}^{-1}$  is the drift velocity.

As most tracks are not perfectly parallel to the wire plane and perpendicular to the wires, a normalisation factor  $n$  was calculated to extrapolate back to the peak signal of a parallel MIP. The normalisation factor is the ratio of the wire spacing and the track pitch [158]. This can also be thought of as the ratio between the distance travelled by a MIP compared to a MIP travelling with an azimuthal angle  $\theta$  and zenith angle  $\phi$  between two wires. Therefore the normalisation for the collection plane is  $n_{coll} = \cos(\theta) \sin(\phi)$  and for the induction plane is  $n_{ind} = \sin(\theta) \sin(\phi)$ . The angles are calculated from the gradients of the linear fittings in each plane described above. For positive gradients, the zenith angle is  $\phi = \tan^{-1}(\frac{1}{m_{coll-tick}})$ , where  $m_i$  is the gradient of the fit in plane  $i$ . For negative gradients,  $\phi = \pi - \tan^{-1}(\frac{1}{m_{coll-tick}})$ . The azimuthal angle is calculated by  $\theta = \tan^{-1}(\frac{m_{coll-tick}}{m_{ind-tick}})$ . The gradients are scaled such that

they are dimensionless. The angular distribution of the events can be seen in Figures 5.22a and 5.22b. Events that are  $\pm 2\sigma$  from the mean angles are removed.

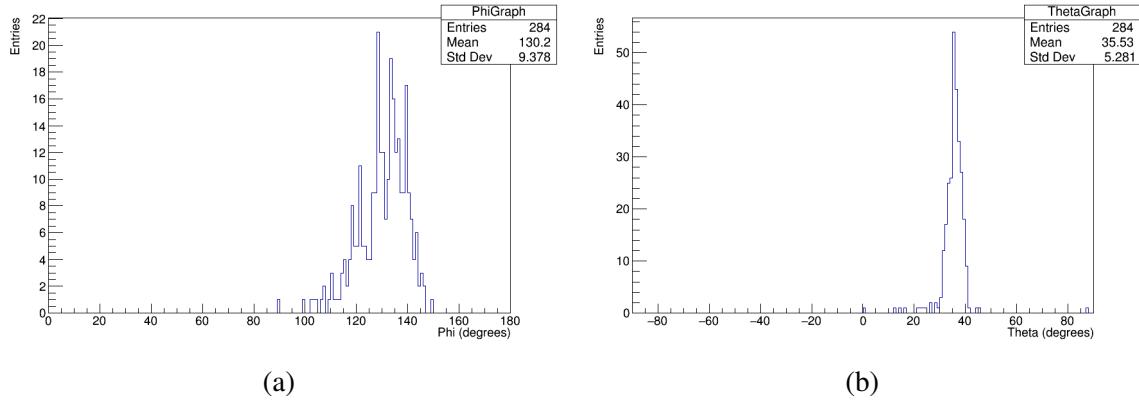


Fig. 5.22 The zenith (a) and azimuthal (b) angular distribution for all events in the run

The electron lifetime [177] was calculated by splitting all charge deposition data into  $35 \mu\text{s}$  time segments. A Landau-Gaussian fit was applied to segments with significant statistics and the most probable charge deposited was extracted. The mean time was then fitted against the most probable charge deposited to extract the lifetime, shown in Figure 5.23, using the exponential in Equation 3.12. This method was initially tested in the SBND framework and is method 2 in the lifetime Section 5.5. The lifetime,  $\tau$ , was found to be  $311.9 \pm 3.0 \mu\text{s}$ . This value was then used to extrapolate back to the charge deposited.

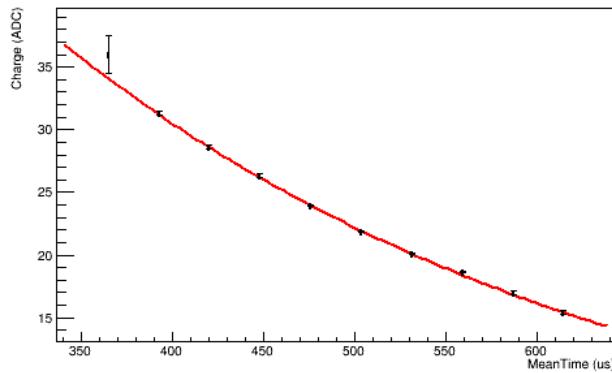


Fig. 5.23 The exponential plot  $Q = Q_0 e^{t/\tau}$ . The errors are the errors on the most probable value that arises during the fitting procedure. The errors are calculated using the MINOS technique [26].

Due to the small size of the TPC, diffusion effects were not considered. A Landau-Gaussian distribution, which can be found in Figures 5.24a and 5.24b, was fitted to the

normalised charge distribution and the most probable value is extracted as the peak signal,  $Q_{mp,peak}$ .

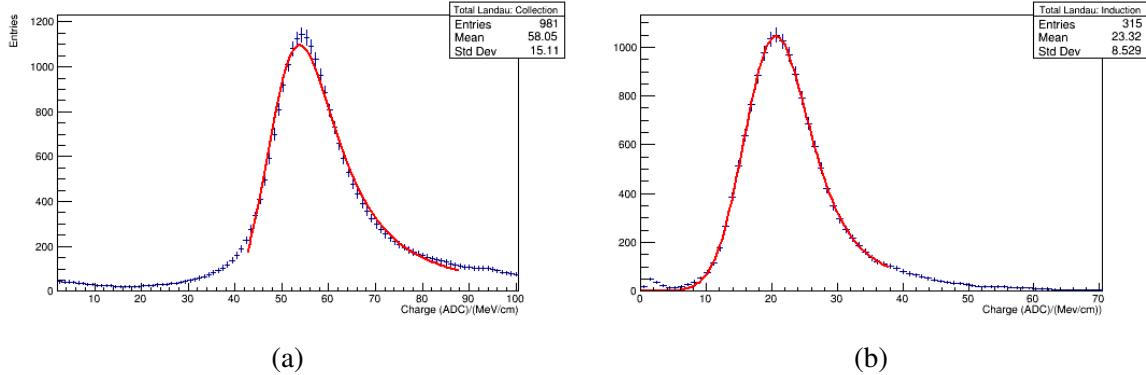


Fig. 5.24 The normalised charge deposition of track like particles in the TPC on the collection plane (a) and induction plane (b). Bins corresponding to less than 42 ADC for the collection plane and 8 ADC for the induction plane are not considered in the fitting to remove noise effects. Errors are Poissonian.

The peak signal to noise ratio is then calculated as

$$SNR_{peak} = \frac{Q_{mp,peak}}{N_{RMS}}. \quad (5.3)$$

The signal to noise ratio of the collection was found to be  $34.60 \pm 0.21$  and for the induction plane  $12.08 \pm 0.08$ . The signal to noise ratio is high and thus this electronics setup would have likely been relatively successful in SBND if the COTS electronics had failed.

### 5.4.5 ENC Measurement

In order to validate the signal to noise ratio analysis an ADC to ENC conversion analysis was performed. This conversion could then be compared to the theoretical value for verification. The ENC is defined as the RMS of the electronic noise in electrons and therefore is a universal scale for comparison between other electronic setups.

The selected normalised signal data was integrated to calculate the total charge deposited on each wire. Assuming that the charge deposited is point-like, the Gaussian height of the shaping amplifier response function corresponds to the charge deposited. Hence the charge deposited was calculated by multiplying by the Gaussian coefficient

$$Q_{dep} = Q_{norm} \frac{1}{2.5\sqrt{2\pi}\tau}, \quad (5.4)$$

where  $\tau$  is the shaping time ( $3\mu\text{s}$ ), and 2.5 is a conversion factor from ticks to microseconds.

A Landau-Gaussian convolution was fitted to the deposited charge distribution, shown in Figure 5.25, and the most probable charge was extracted in ADCs. One can then convert from ADCs to electrons by calculating the number of electrons deposited by a MIP on a wire using

$$N_e = 1.7 \text{ MeV/cm} \times p \times \frac{1}{23.6 \text{ eV}} \times 0.68, \quad (5.5)$$

where 1.7 MeV/cm is the minimum ionisation for a relativistic particle which obeys the Bethe-Bloch equation.  $p$  is the wire pitch (0.254 cm), 23.6 eV is the energy required to ionise a liquid argon atom, and 0.68 is the recombination factor.

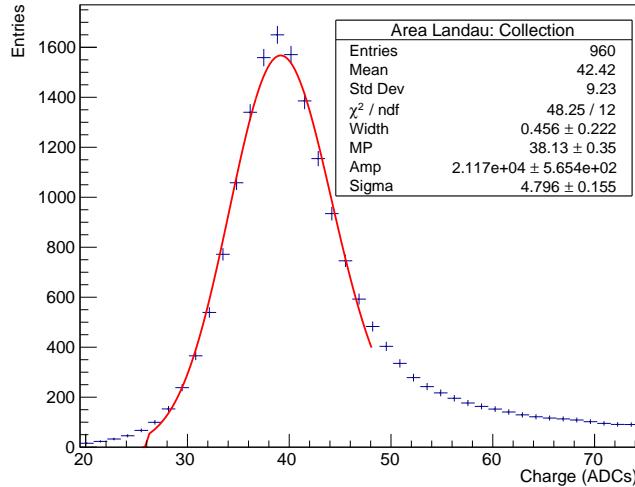


Fig. 5.25 The normalised charge deposition of track-like particles in the TPC multiplied by the Gaussian coefficient on the collection plane. Bins corresponding to less than 26 ADC are not considered in the fitting to remove noise effects. Errors are Poissonian.

The ADC to electron conversion was found to be  $346.9 \pm 1.2$ . This is comparable to the theoretical value of 359, which is derived in Equation 5.4.1. However, the value is not within the statistical error and an extension of this work would be to perform a systematic error analysis to see if the two numbers were in agreement. Using the conversion factor, the ENC for the collection plane is  $499.8 \pm 3.4$  and induction plane is  $517.2 \pm 3.5$ . These numbers cannot be directly compared to the VST as there was no effort to filter noise in the analysis, and so the contribution from external noise sources sets a bias in the comparison. Despite this, the signal to noise ratio is still higher than the SBND requirements and thus this electronics setup is likely to be sufficient.

## 5.5 Electron Lifetime Studies in SBND

Three modules were created, each using a different method for calculating the electron lifetime. The "LArIAT" method [232], method 1, was implemented in the VST online monitoring and used in the VST signal to noise ratio analysis. The "ICARUS" method [177, 175], method two, was used in the CERN test stand data. The third "parallel" tracks method is to be used in the SBND online analysis. Each method is described below. Each method possesses characteristics which are appealing for certain implementations. The theoretical background of electron lifetime calculations can be found in Chapter 3. Firstly, the methods will be described and the results from simulated data will be discussed. Finally, the implementation of the modules will be considered.

### 5.5.1 Methods to Calculate the Lifetime

#### Method 1: Single Track Method

The first method, based on the LArIAT method [232], calculates the electron lifetime for each muon event and the average of the data sample is taken as the electron lifetime of the detector. The data sample only contains simulated muons which cross through the anode and cathode. Such an event can be seen in Figure 5.26. This ensures enough hits to make an accurate estimate of the lifetime. Selection cuts are then implemented to ensure the event has sufficient statistics to perform the lifetime fit. The selection cuts can be found in Table 5.2. The charge deposited is recorded along with the timing data. The data is then fitted on an exponential curve, described by Equation (3.12), using a log-likelihood estimation that assumes an underlying Landau distribution. The lifetime parameter is then extracted for a single track. Finally, the lifetime is found by taking the mean of the lifetime values calculated for all events.

Simulated muons with random orientation and with an energy of 3 GeV were used to test the performance of the method. The performance for calculating various simulated electron lifetimes correctly is shown in Figure 5.29a.

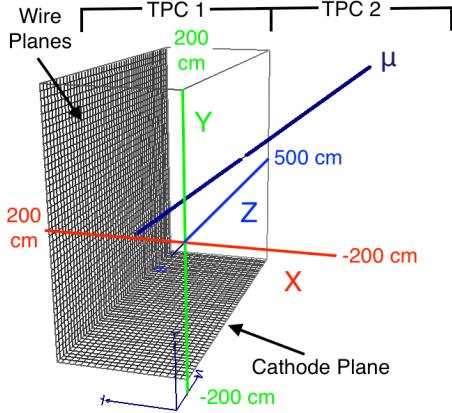


Fig. 5.26 A example of an anode to cathode crossing muons in the SBND detector. Credit: Tom Brooks.

Selection Cuts
Hits associated with delta rays removed
Recursive linear fit performed and hits are removed out of the tolerance
Tracks with less than 80 hits removed
Tracks with $\chi^2/Ndof < 2$ are removed
Tracks close to parallel or perpendicular to wire plane are removed
Tracks cut if they cover less than 94 cm in the drift distance

Table 5.2 Selection cuts used in the lifetime selections.

## Method 2: Multiple Track Method

This method, based on the ArgonNeuT [177]/ICARUS [175] methods, uses the same data set as method 1 and undergoes the same selection procedure. However, this method uses all of the data at the same time. The  $dQ/dx$  data is split into 20 time slices and a Landau-Gaussian convolution is fitted to each time slice. One of the fits can be seen in Figure 5.27a. The most probable charge is then extracted for each time slice and the mean time is taken as the drift time. The data is then fitted to the exponential described by Equation (3.12) and shown in Figure 5.27b. This is achieved by using a  $\chi^2$  minimisation and the lifetime is extracted as a fit parameter.

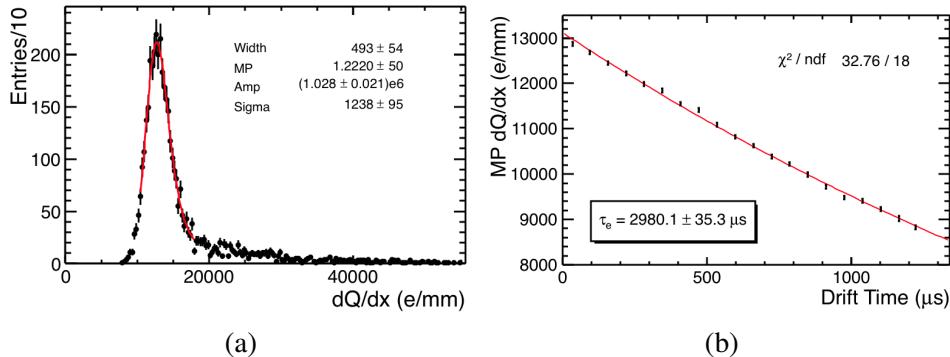


Fig. 5.27 a) The fitted Landau-Gaussian convolution to the charge deposited on the wire plane for a specific time slice. The exponential fit of the average time of each time slice against the most probable charge. The error bars are extracted using the MINOS technique [26]. The technique finds the error on the lifetime by taking the difference of the lifetime values for which the  $\chi^2$  is +1 from the minimised value. Credit: Tom Brooks

### Method 3: Parallel Track Method

This method uses a data set that consists of simulated muons that travel parallel to the wire plane. Figure 5.28 shows a schematic of such an event. In this case, the CRT system described in Chapter 4 is used for triggering. The analysis below considers a sample of 180 muons at 3 GeV, evenly distributed between 19 CRT strip pairs that lie between  $-180$  cm and  $180$  cm from the cathode, separated by 20 cm. The direction of the particles is such that the particles pass through two opposite side CRT strips. A Landau-Gaussian is then fitted to muon charge depositions in the CRT strips using the log-likelihood method. This method uses the same technique and error analysis as method 2. The most probable charge is then extracted as well as the average hit time. These values are then fitted to the exponential (3.12) using a  $\chi^2$  minimisation.

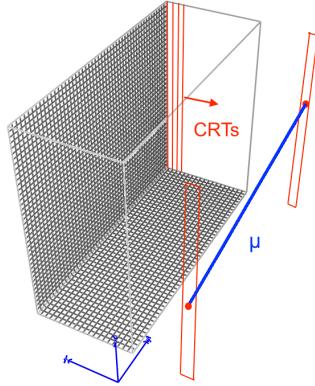


Fig. 5.28 An example of a muon crossing parallel to the wire plane. Credit: Tom Brooks

### 5.5.2 Efficiency and Accuracy of the Methods

The results for different simulated lifetimes can be found in Figure 5.29a. As shown, all three methods broadly concur with the simulated lifetime for small lifetimes. As the lifetime increases, the results vary from the simulated lifetime, notably for method 1. This effect can have a noticeable impact on the charge losses during the correction for the maximum drift distance. As can be seen in Figure 5.29b, the effect has a difference of about 10% on the calculated percentage of charge lost during the full drift for a lifetime of 10 ms for method 1.

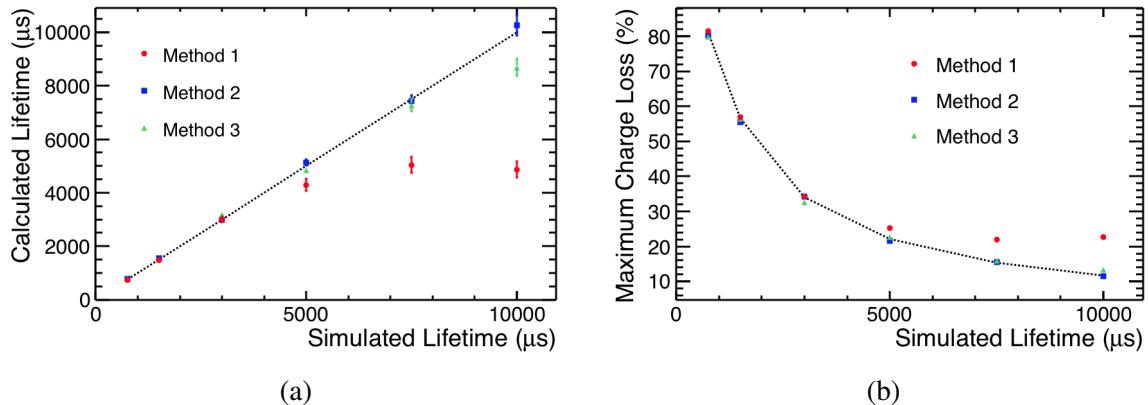


Fig. 5.29 a) A plot of the calculated lifetime against the simulated lifetime for each module. The error bars are calculated using the MINOS technique [26]. The technique finds the error on the lifetime by taking the difference of the lifetime values for which the  $\chi^2$  is +1 from the minimised value. Credit: Tom Brooks. b) A plot of the calculated percentage of charge lost across the full drift distance of the TPC.

### 5.5.3 Comparison of the Methods

Methods 1 and 2 are more advantageous than method 3 due to the cosmic flux rate. Assuming the muon flux is described in terms of a  $\cos^2 \theta$  distribution,

$$N = \int I_0 \cos^2 \theta d\Omega dt dA \quad (5.6)$$

where  $I_0$  is the flux vertically ( $\theta = 0$ ),  $A$  is the area, and  $\Omega$  is the solid angle. The rate of anode to cathode muons compared to muons travelling parallel to the wire plane is approximately 2:1. This simple analysis does not consider muons that arise due to interactions of beam particles within the dirt before entering the detector. A full simulation is required to acquire the appropriate rates. However, method 3 could require less reconstruction and hence less processing time if the CRT system is exploited. This characteristic would make method 3 useful as an online or near-online module.

Method 1 only requires one event to produce a lifetime measurement and the processing time is proportional to the size of the data. It is also the fastest module and this makes it useful as an online purity monitor which measures the lifetime quickly but not accurately. Such a monitor would protect against disaster type situations. Methods 2 and 3 require a sufficiently large data sample to make the Landau-Gaussian fitter efficient. As the data set size increases, the processing time of the Landau-Gaussian fit falls exponentially. After the number of events reaches approximately 150, other analytical processes in the modules start to dominate the processor and the processing time starts to increase with sample size linearly. Therefore, 150 events is the optimum sample size for reducing the processing speed.

In terms of accuracy, the statistical error decreases rapidly for methods 1 and 2 as a function of event number, saturating at 600 events for method 2 and 200 events for method 1. Method 2 has the best statistical error.

All three lifetime analyses have their strengths and hence all three methods have been adapted for use on the test stand data or the online monitoring of SBND. In conclusion, method 1 provides the least accurate evaluation of the lifetime but is the fastest method. Method 2 and 3 require a data sample of approximately 150 muons in order to run at their optimum efficiency, which is a lengthy process. However, all three methods require more events than their optimum run time sample size to reduce the statistical error. Therefore, for the level of accuracy required, this difference in time is not significant. Furthermore, methods 1 and 2 require additional reconstruction in order to perform the analysis which increases the run time of the analysis. The CRT system could be exploited to reduce the reconstruction in method 3. On balance, method 3 has the most advantages although it does take longer to obtain the sample required.

## 5.6 Concluding Remarks

Prior to the construction of the SBND detector, extensive efforts were made to test the COTS electronics readout chain in the VST experiment. The experiment was able to run successfully in order to perform a signal to noise ratio analysis during which the signal to noise ratio was found to be  $57.55 \pm 0.31$  (stat) for the collection plane and  $15.58 \pm 0.31$  (stat) for the induction plane for the 20 kV runs. The signal to noise ratio was then extrapolated for the SBND detector and it was shown that the noise RMS was within the minimum values required for the collaboration to reach the physics goals set in the proposal. It was also shown that the requirements set by SBND of a signal to noise ratio of 12/1 on the collection plane wires is met across the entire drift region.

A successful backup for the COTS electronics was demonstrated at the CERN test-stand with the ICARUS-SBND hybrid electronics. The signal to noise ratio of the collection plane for the CERN test stand was found to be  $34.60 \pm 0.21$  and for the induction plane it was found to be  $12.08 \pm 0.08$ . This is likely to be sufficient for the requirements of the SBN programme. Furthermore, no external noise removal was undertaken on either setup and an extension of this work would attempt to remove noise and recalculate the signal to noise ratio value.

In both analyses, an electron lifetime analysis was required. After formulating three different approaches to calculate the electron lifetime in the SBND simulation, two of the approaches were used in the test-stand analyses. The final parallel track lifetime measurement will be adapted to work as a near-line method of calculation for the SBND experiment.

# Chapter 6

## Reconstruction of Particle Interactions in the SBND Detector

As the amount and rate of data taken by the Short-Baseline Near Detector (SBND) is large, a sophisticated and reliable automated reconstruction chain is required. This Chapter will discuss the reconstruction chain in SBND which takes advantage of preexisting LArSoft algorithms [224, 249] developed by previous experiments. Algorithms developed in LArSoft are easily accessible and designed to run out of the box for any liquid argon experiment. Because of this, all the Short-Baseline Neutrino (SBN) experiments could also easily use the same base reconstruction, minimising systematic errors for an oscillation analysis. Focus will be placed on the reconstruction of showers. The reconstruction of showers is of particular interest as it is required for identifying  $\nu_e$  CC events. This identification is essential for the main analysis of this Thesis which is a  $\nu_e$  sterile appearance analysis described in Chapter 8.

This Section will go onto describe the current reconstruction chain in SBND. The wire data reconstruction starts with the deconvolution of the charge signal discussed in Section 6.1.1. After the deconvolution of the signal with the detector response, 1D hits are found on the wires. The hit finding process is discussed in Section 6.1.2. Using the timing information and the location of the hit wire, the hits are then combined into 2D clusters in the wire planes, so that the clusters resemble particle trajectories. This is primarily performed by the Pandora pattern recognition software which is discussed in Section 6.1.3. However, alternative LArSoft algorithms also exist and one in particular, the Blurred Clustering algorithm which focuses on clustering electromagnetic showers, is discussed in Section 6.1.4. The images in the planes are then combined to form 3D objects. The Pandora software also performs this task and the process is discussed in Section 6.1.3 and 6.1.6.

Once particles have been identified, characterisation processes occur which provide further information on the particle. For track-like particles dedicated calorimetric software is

used and this is discussed in Section 6.1.7. The information from the calorimetry is used to identify the particles from the energy deposition. This particle identification is discussed in Section 6.1.8.

As indicated above, there will be a particular focus on the shower reconstruction pathway and a new framework for characterising showering particle properties. The software and performance will be discussed in Section 6.2.

## 6.1 Reconstruction Chain in SBND

### 6.1.1 Charge Reconstruction

The current induced on a wire at the anode of the TPC can be described approximately by Ramo’s Theorem [250, 251]. The theorem describes the field response to a unit charge deposited at the cathode. The current induced is then received by the front-end LArASIC [208] which amplifies and shapes the signal. A 1D deconvolution is performed to disentangle the signal from the field and the electronic effects. The detector response also accounts for noise which is independent of the signal and is currently modelled as a white noise background. Hence the detector response can be described as

$$y(t) = h(t) * x(t) + n(t), \quad (6.1)$$

where  $t$  is time,  $y(t)$  is the convolved detector signal response,  $x(t)$  is the original signal response,  $h(t)$  is the impulse response of the detector which includes the field and electronics response and  $n(t)$  is the noise distribution [251]. In order to reconstruct the observed signal, a Wiener-like filter has been created. A filter in this context takes the form  $g(t)$  with

$$\hat{x}(t) = g(t) * y(t), \quad (6.2)$$

where  $\hat{x}(t)$  is the closest estimate to the real signal. The Wiener filter has the following form [251]

$$G(\omega) = \frac{H^*(\omega)S(\omega)}{|H(\omega)|^2S(\omega) + N(\omega)}, \quad (6.3)$$

where  $\omega$  is frequency and  $G(\omega)$  and  $H(\omega)$  are Fourier transforms of  $g(t)$  and  $h(t)$  respectively. Taking  $X(\omega)$  and  $V(\omega)$  as the Fourier transforms of  $x(t)$  and  $n(t)$  respectively, then  $S(\omega) = \mathbb{E}|X(\omega)|^2$  and  $N(\omega) = \mathbb{E}|V(\omega)|^2$  are the mean spectral power densities of  $x(t)$  and  $n(t)$  respectively.

It has been proven that the Wiener filter minimises the mean square error  $\varepsilon(t) = \mathbb{E}|x(t) - \hat{x}(t)|^2$  [252, 251] and therefore provides the best estimate of the true signal. However, as explained in Reference [251], the form of the filter is problematic for LArTPC signal processing as it leads to large distortions, and therefore a Wiener-like filter is used

$$F(\omega) = \begin{cases} e^{\frac{1}{2}(\frac{\omega}{a})^b} & \omega > 0 \\ 0 & \omega = 0, \end{cases} \quad (6.4)$$

where a,b and c are positive fit parameters, This form means  $\lim_{\omega \rightarrow 0} F(\omega) = 1$  ensuring the smearing in the time domain is local. Other devolution techniques, such as 2D deconvolution using the WireCell package, are currently in use in the MicroBooNE detector and could be used in SBND in the future [253, 254].

### 6.1.2 Hit Finding

After the deconvolution and noise removal, the hit finding algorithm *GausHitFinder\_module* [255, 256] is applied to the readout waveforms. The algorithm identifies peaks above a threshold and within a collection of ticks then attempts an N-Gaussian fit to describe the charge depositions, where N is the number of peaks.

The method first finds hit candidates by identifying the highest peak in the waveform above a set threshold (for SBND = 10 ADC (after deconvolution)). The algorithm then iterates through the time ticks, prior to the peak in the waveform, until a minimum or inflection point is found. The same procedure is done for finding the endpoint of the hit. This method is particularly effective when the Wiener-like filter, described in Section 6.1.1, is used in the deconvolution. The reason is two fold: firstly, the deconvolution procedure creates two minima at the ends of the hits; secondly, the bipolar induction signals become unipolar after deconvolution removing the minima. These effects can be observed in Figure 6.1 which shows an example of a typical waveform on the induction and collection planes. Using the procedure above, the algorithm recursively finds hits on the waveform prior to the start position of the hit and after the end position of the hit.

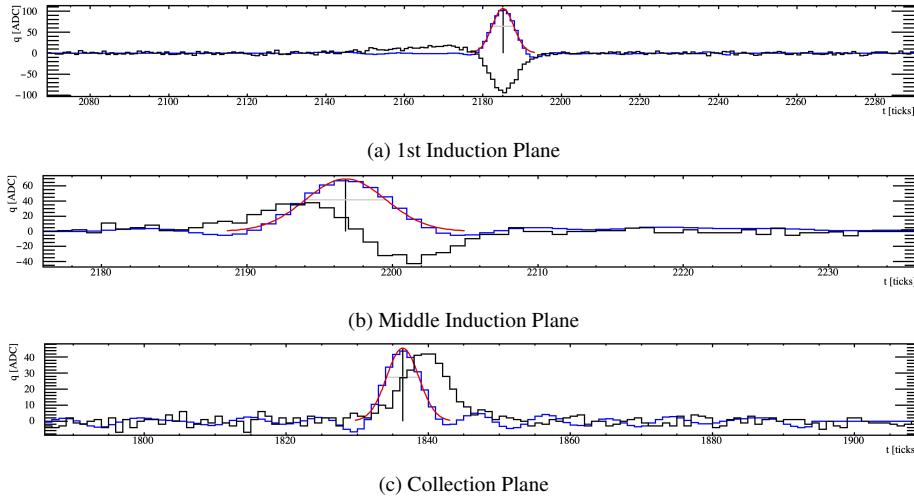


Fig. 6.1 The raw (black) and deconvolved (blue) waveforms. A Gaussian (red) has been fitted to peaks using *GausHitFinder\_module*. The x-axis corresponds to the time in ticks ( $0.5 \mu\text{s}$ ), and the y-axis corresponds to amplitude in ADC.

Candidate hits then undergo a merging procedure if they border other candidate hits. Then, for a hit group made of  $N$  merged hits, an attempt is made to fit  $N$  (for SBND  $N_{Max} = 10$  (set for computational reasons)) Gaussians to the group to find the properties of the hits. This is done with the TF1 class in ROOT [257], using a  $\chi^2$  minimisation.

A `recob::Hit` is created using the hit properties and the charge of the hit can be calculated from the peak height as

$$Q = \frac{1}{\sqrt{2\pi}} \frac{A\sigma}{C}, \quad (6.5)$$

where  $A$  is the amplitude of the fitted Gaussian (or  $\frac{1}{\sqrt{2\pi}} \sum_i \frac{Q_i}{\sigma}$  as above),  $\sigma$  is the width of the fitted Gaussian and  $C$  is a user-defined constant to convert the area to the peak height. Alternatively, the charge of the hit is calculated as the integral of a fitted Gaussian,

$$Q = \sum_{i=\text{first tick}}^{\text{last tick}} A e^{\frac{i-\mu}{2\sigma^2}}, \quad (6.6)$$

where  $\mu$  is the mean tick of the fitted Gaussian (or the peak position if the fit failed). SBND uses the former method to calculate the charge. The charge is calculated like this because, as was explained in Section 4.3.2, the impulse response from the front-end ASIC is a Gaussian where the area of the hit is the charge deposited.

After creating the `recob::Hit`, each hit undergoes a filtering procedure which removes any hits where the amplitude and width are less than a user-defined value (SBND minimum

amplitude = 5 ADC, SBND minimum width = 1 tick). This stage attempts to remove small charge depositions which are more likely noise.

### Hit Finding Efficiency

In order to test the efficiency of the implementation of the *GausHitFinder\_module* in SBND, a truth-hit finding study was undertaken. The study involved creating hits using the true charge deposited on the wires without the inherent noise. Hits in the reconstruction were then compared to these true hits to identify the efficiency of the hit finder.

In the analysis, muons were simulated within the SBND detector using the SBND simulation chain described in Section 4.7. Samples containing 100 muons with energy of 2 GeV were produced for every  $10^\circ$  from the zenith ( $\phi = XZ$ ) and azimuthal ( $\theta = YZ$ ) directions. For the zenith sample, the particles were simulated with an azimuthal angle of  $0^\circ$ , corresponding to the beam direction and hence parallel to the wire plane. For the azimuthal sample, events are simulated with a zenith angle such that the muons travelled perpendicular to the wire direction of the plane being analysed. This setup minimises the energy deposited on the wires.

Monte Carlo (MC) hits for each wire were created by identifying maxima, above a set threshold (MC height threshold), in the true charge depositions for the muons and the daughter delta rays separately. Neighbouring points (or voxels) on the waveform that contained charge were combined and considered one packet due to the electronics response. Multiple packets can exist on one wire. For  $N < 10$  maxima in a packet fits using permutations of  $1 \rightarrow N$  Gaussians were performed. The fitting is done using a  $\chi^2$  minimisation in the ROOT TF1 [257]. Figure 6.2 shows a packet with three maxima and the possible permutations of the fit. The fit with the minimum  $\chi^2/\text{Ndof}$  (Ndof = Number of degrees of freedom) was used to create MC hits definitions using the fit parameters of the Gaussians.

The fitting procedure is done for the muons and delta ray charge depositions independently. Most delta ray depositions overlap with the muon depositions in time, therefore in reconstruction they are indistinguishable. Hence, once the MC hits are created, the hits are merged if the peak times of a muon and delta ray MC hit occur within a set tolerance ( $dt$  Threshold).

MC hits are then compared to reconstructed hits. Reconstructed hits that are more than five ticks ( $2.5\mu\text{s}$ ) away from a MC hit were assumed to be created from noise. The remaining hits were matched to the closest MC hit in time.

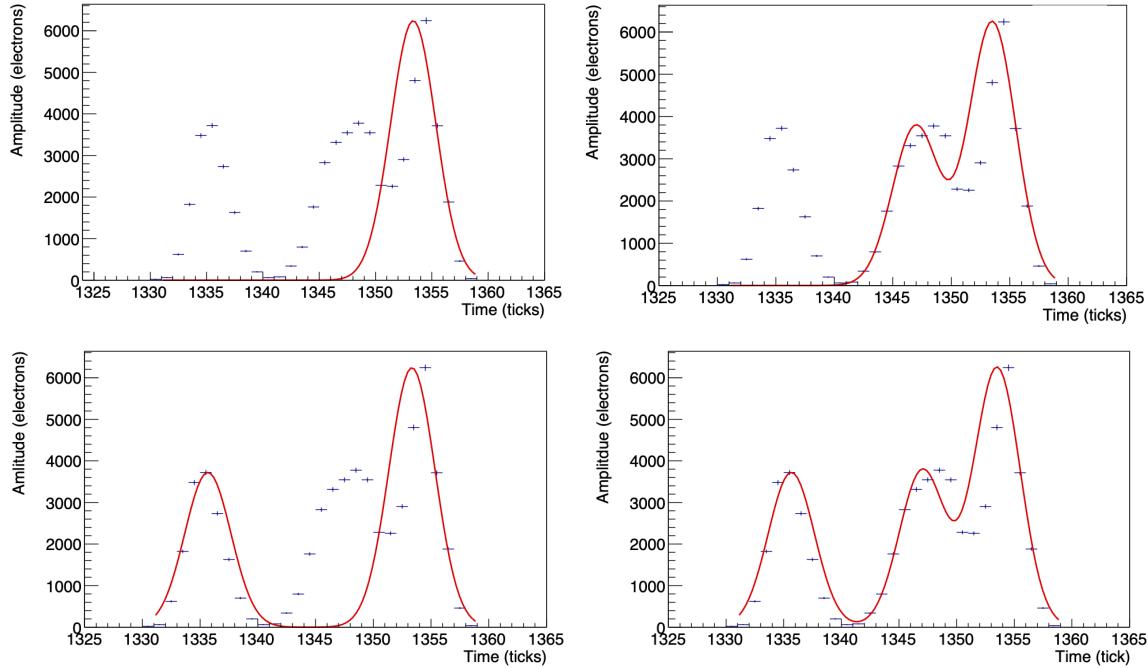


Fig. 6.2 A Monte Carlo waveform with corresponding Gaussian fit attempts. The fit with the lowest  $\chi^2/\text{Ndof}$  is chosen to model the wavepacket.

An efficiency analysis using various MC  $dt$  merging thresholds, MC amplitude thresholds and *GausHitFinder\_module* height thresholds was undertaken. Below a subset of the Figures are shown and the full results can be seen in Appendix A. A nominal white noise was convolved with the detector response to produce the waveforms used in this analysis. Therefore, the creation of fake noise hits was minimal, and the study should be repeated when a more realistic model is implemented in SBND.

Three definitions of efficiency were defined to evaluate the hit finding efficiency:

- $\frac{\text{The number of wires with one reco hit}}{\text{number of wires crossed} - \text{number of wires with more than one hit}}$  (noise hits are not considered)
- $\frac{\text{The number of reco hits} - \text{number of noise hits}}{\text{number of MC hits}}$  (excluding wires with delta ray hits)
- $\frac{\text{The number of MC hits} - \text{number of MC hits not matched to a reco hit}}{\text{number of MC hits}}$

The first definition does not use the MC procedure and therefore can be adapted to use on real data. The second removes the effect of defining the merging procedure by ignoring delta-rays whilst the third definition includes the delta-rays. The third definition therefore gives an indication of the fundamental efficiency level of vertex reconstruction in the detector.

As can be seen in Figures 6.3, 6.4 and 6.5 the *GausHitFinder\_module* has a very high efficiency at identifying muon hits when the muon momentum is parallel to the wire plane.

As expected, the efficiency drops as the threshold on the *GausHitFinder\_module* is increased. The efficiency decreases as the muon momentum becomes more in line with the face of the wire plane. There is also a drop in efficiency when the muon travels parallel to the wire direction. In these cases, only a few sense wires detect the energy deposition meaning statistical fluctuations are more prominent and effects such as diffusion have a greater impact. Furthermore, in these topologies, a point source poorly describes the charge induced on the wires and instead the charge is called a pulse train, shown in Figure 6.6. In this case, the analysis becomes ill-defined.

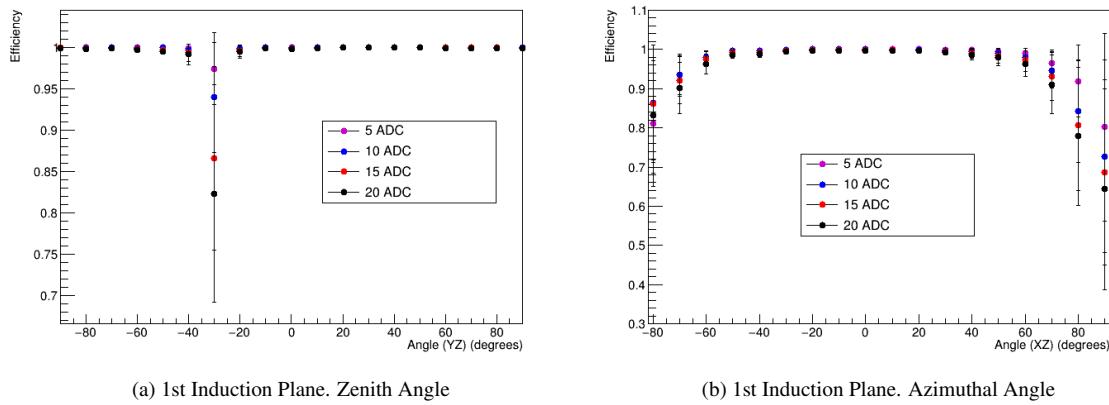


Fig. 6.3 The 1st efficiency definition in Equation 6.1.2 for various ADC thresholds set in the *GausHit\_Module* as a function of the azimuthal (XZ) and zenith (YZ) angles. The MC ADC threshold is set to 5 ADC and the  $dt$  merging threshold is set to 2. Errors are the standard deviation of the efficiency for the corresponding angle.

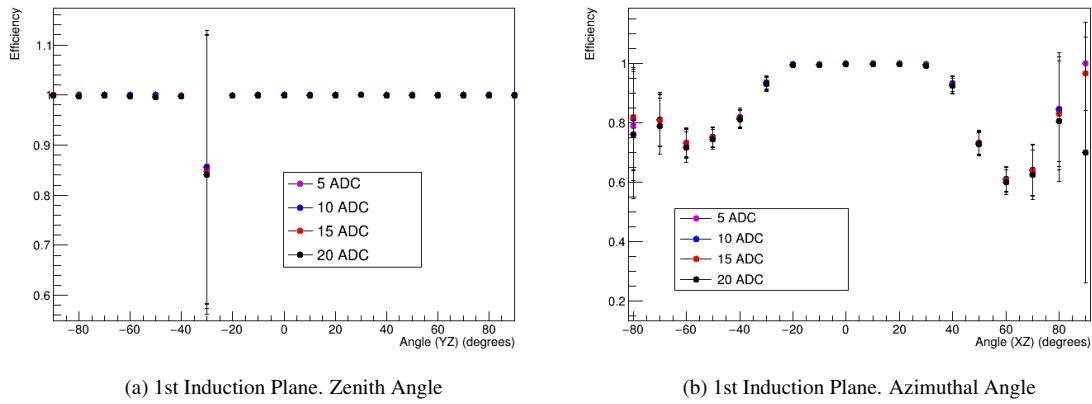


Fig. 6.4 The 2nd efficiency definition in Equation 6.1.2 for various ADC thresholds set in the *GausHit\_Module* as a function of the azimuthal (XZ) and zenith (YZ) angles. The MC ADC threshold is set to 5 ADC and the  $dt$  merging threshold is set to 2. Errors are the standard deviation of the efficiency for the corresponding angle.

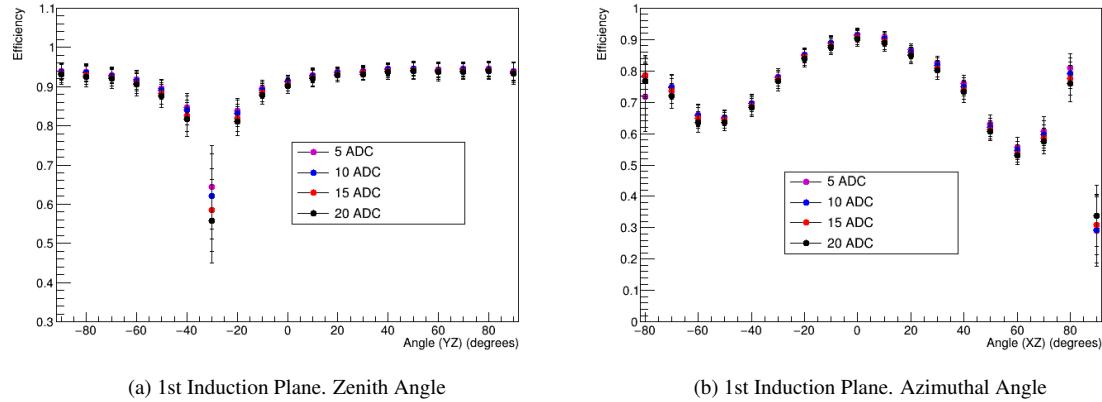


Fig. 6.5 The 3rd efficiency definition in Equation 6.1.2 for various ADC thresholds set in the *GausHit\_Module* as a function of the azimuthal (XZ) and zenith (YZ) angles. The MC ADC threshold is set to 5 ADC and the *dt* merging threshold is set to 2. Errors are the standard deviation of the efficiency for the corresponding angle.

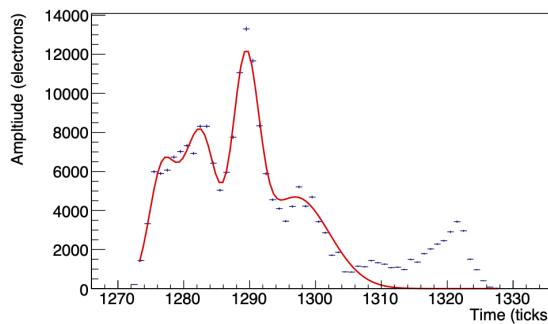


Fig. 6.6 Example of a waveform pulse train where the energy is deposited along several time ticks. In this case a multi-dimensional Gaussian fit becomes unsuitable

The efficiency increases as the MC merging threshold increase, as expected. Also, as expected, increasing the MC amplitude threshold also increases the efficiency as more true energy depositions are not considered in the analysis since they are too small to see in the detector. This is further discussed with the Figures in Appendix A.

Although the efficiency for hit identification is high, for MIP-like particles, on average 11% of the energy is not detected within the hits for the collection plane. This is visible in Figure 6.7 where the fractional reconstructed energy peaks at 96% but there is a long tail on the distribution. The tail is primarily due to losses in charge due to an increase in the azimuthal angle where the hits are poorly reconstruction.

Figure 6.7 also compares the fraction of energy reconstructed in the hits for the muon sample to a vertex sample where a charged pion and electron with energies corresponding to

the expected BNB energies are created at the same position. There is a significant difference between the electron shower sample (peaking at 83%) and the muon track MIPs (peaking at 96%). This difference has implications for the clustering, discussed in Section 6.1.3 and the shower energy calibration, discussed in Section 6.2.2, which uses MIPs to calibrate the shower energy.

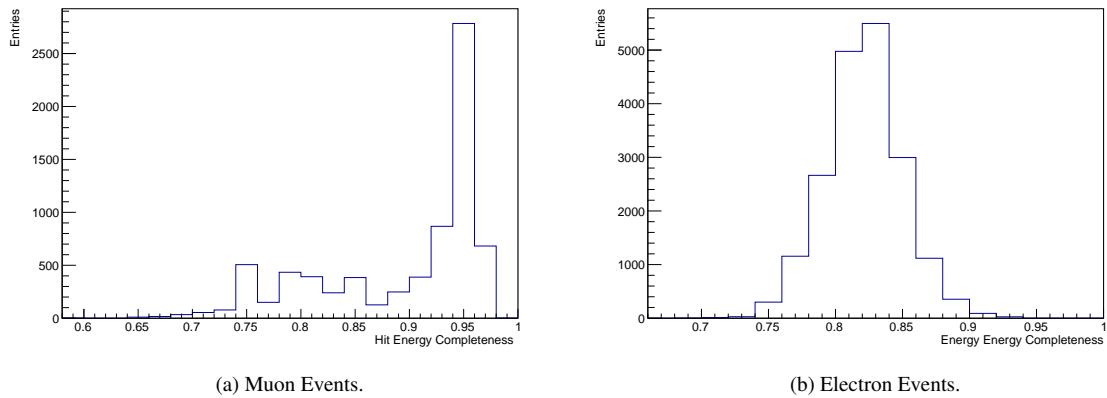


Fig. 6.7 The fraction of energy reconstructed in the hits in the muon sample and electron sample.

### 6.1.3 Pandora Pattern Recognition

Once the hits have been reconstructed, they are passed to the Pandora pattern recognition software [27]. This is the primary clustering approach currently used in SBND. All the shower characterisation methods discussed in Section 6.2 use the output of Pandora.

In order to identify particles, the pattern recognition software firstly reconstructs 2D clusters of hits that correspond to particles. Clustering is initially done tentatively and clusters are only made for continuous lines of hits with no overlapping features. This creates clusters where the fraction of energy deposited by a particle within the cluster is low, but the energy within the cluster is predominately from one particle. There are then several cluster merging algorithms which match, grow and break clusters depending on the topology of the cluster. For shower-like clustering, shower spines are initially identified and then branches are recursively added to the spine.

Once well-formed 2D clusters are made, clusters are matched between planes to form 3D clusters. This step is done for the track-like particles using a sliding fit algorithm which performs a linear fit at several points along a cluster. The fits are then combined to form the track trajectory. Shower-like particles undergo a similar procedure, but the hit positions

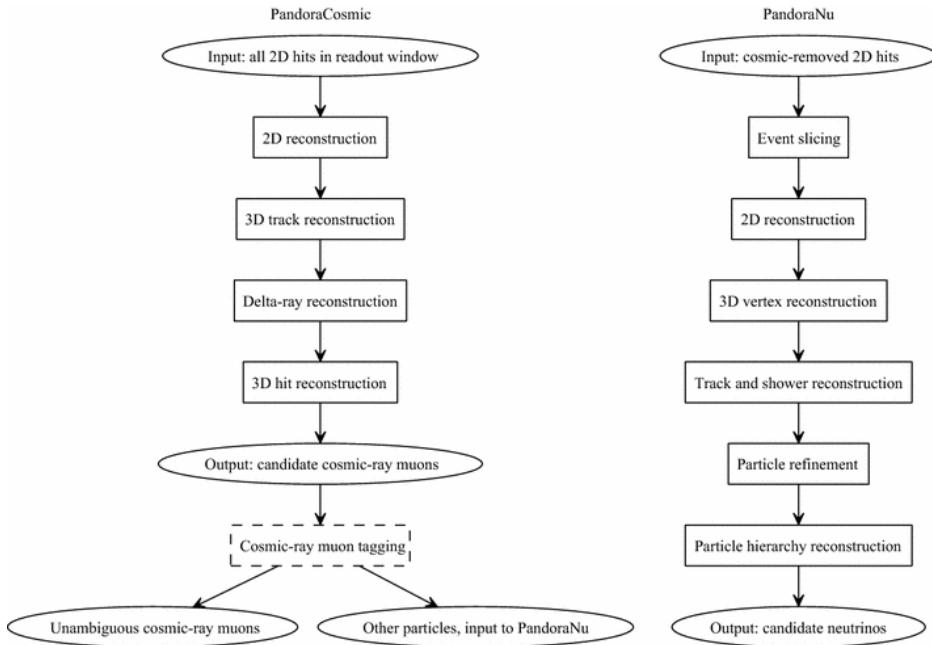


Fig. 6.8 The reconstruction chain for the Pandora pattern recognition. From Reference [27].

are used rather than the sliding fit. 3D clustering is an iterative procedure where the 2D reconstruction is altered to remove any ambiguity in the 3D reconstruction.

3D hits, known as spacepoints, are also created for each hit. Therefore if the reconstruction works correctly, there are three spacepoints for every charge deposition in the TPC. For track-like particles, the 2D hits are combined with a 3D fit to create the spacepoint. This is done by evaluating several trajectory points along the sliding fit and choosing the candidate with the minimum  $\chi^2$  to make the 3D hit. For the shower-like particles, there is no sliding fit and therefore hits within a small region in the drift direction are matched instead.

Events that are unambiguously defined as cosmic particles due to their topology are removed, and then the track-like clustering is repeated. The 3D hits created are then divided into slices which attempt to isolate neutrino interactions. In SBND, a Boosted Decision Tree (BDT) [258, 259] is used to identify slices. Once the slice area is identified, vertex reconstruction is undertaken. Vertex candidates are identified using two 2D clusters from different planes. Using the endpoints of the clusters and allowing extrapolation enables candidate vertices to be made. The input parameters to the BDTs include parameters such as the asymmetry of the event and a beam deweighting score which, for example, favours downstream vertices to account for the neutrino direction correlating with the beam direction.

Once the vertex has been identified, the clusters in the slices are classified into track-like and shower-like objects. This classification depends on parameters such as the length and RMS of the cluster. This procedure is performed by a cut based method, or a Support Vector

Machine (SVM) [258, 260]. Finally, after all particles in the neutrino interaction have been reconstructed, the hierarchy scheme is identified which describes the relationship between the identified particles. Each particle is reconstructed as a PFParticle with a 3D vertex position. As will be discussed in Chapter 7 and Appendix D, the characterisation SVM that was originally tuned for MicroBooNE was used in reconstruction to improve the performance of proton and shower identification. Note: the shower characterisation analysis in Section 6.2.2 does not use the SVM in the downstream reconstruction.

### 6.1.4 Alternative Shower Clustering

Instead of the pattern recognition the *BlurredClustering\_module* [261, 262] can be used to merge hits into 2D clusters. The *BlurredClustering\_module* was initially developed by Mike Wallbank [261] in the Deep Underground Neutrino Experiment (DUNE) framework and was originally from the T2K experiment. A full description can be found in Reference [263]. Firstly, the directionality of all the 2D hits on a plane is found by using a principal component analysis [245]. A Gaussian kernel, dependent on the direction of the hits, is then created which smears the hits in the wire and tick space. This smearing is shown in Figure 6.9. The kernel creates new fake hits that allow for real local hits to merge via a nearest neighbour method and form clusters. The fake hits are then removed to form the final 2D cluster.

The *BlurredClustering\_module* was tuned to maximise the completeness  $\times$  purity of clusters in a single  $\pi^0$  interaction in SBND using a grid search method. The completeness describes the fraction of energy deposited by the particle that was reconstructed in the cluster. The purity describes the ratio of the energy deposited by the particle in the cluster against the energy of other particles.

The results of this clustering method are given only to the *EMShower\_module* [261, 264, 265] shower characterisation in Section 6.2. As the *BlurredClustering\_module* only performs clustering in two dimensions the projection matching is performed by the *EMShower\_module* [261, 264, 265]. During this several clusters are combined to form the shower. Hence, metrics that describe the performs of this reconstruction path way are shown as "EMShower Using Blurred Clustering" in figures.

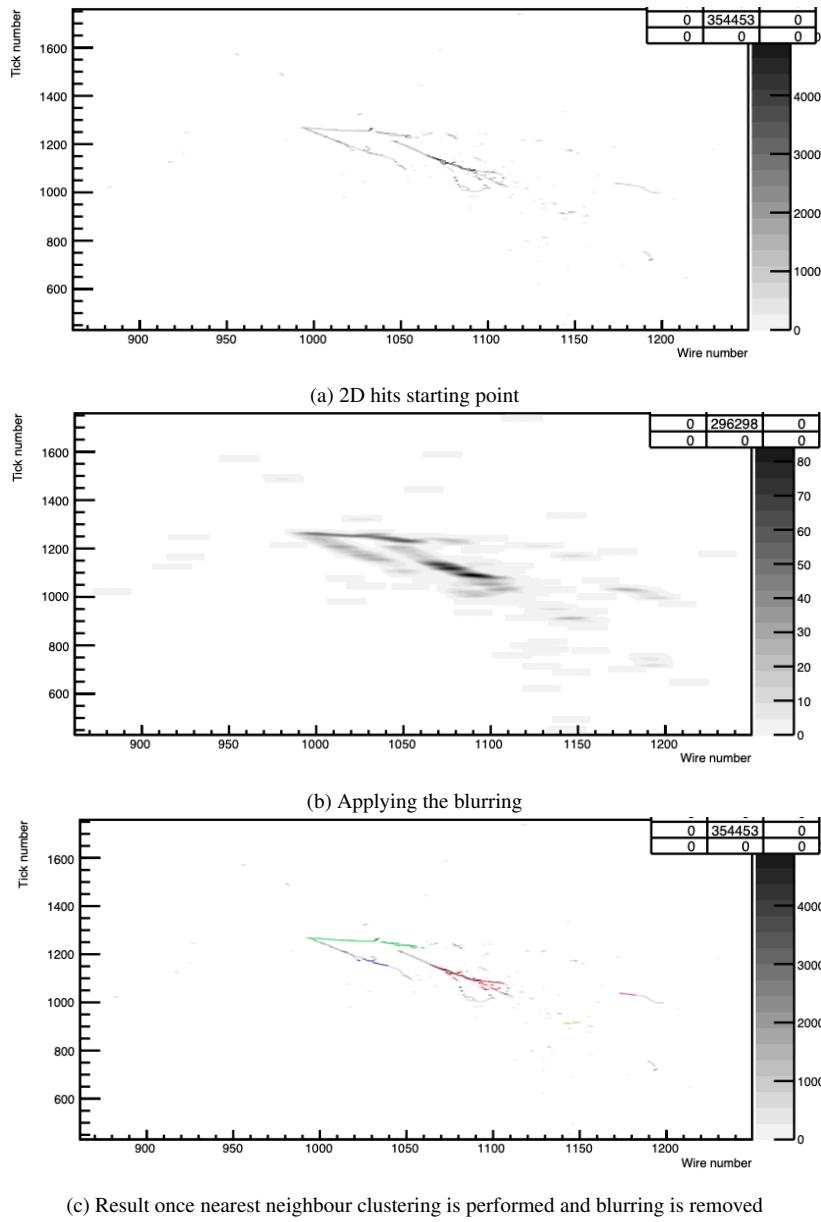


Fig. 6.9 The process of the *BlurredClustering\_module*. The 2D hits are blurred using a Gaussian kernel then a nearest neighbour algorithm is run to merge 2D hits together.

### 6.1.5 Comparison of Shower Clustering Methods

As reconstructing and characterising EM showers is essential for a  $\nu_e$  analysis the above clustering algorithms were compared. The clustering ability was examined using simulated samples of electrons and photons with energies distributed in the form of electrons arising from charge current events from the BNB. A charged pion was also simulated (with energy

corresponding to pions in BNB interactions) at the same starting position in order to make a vertex. Only reconstructed showers which satisfied the criteria in Table 6.1 were analysed.

Cut	Reason For Cut.
Remove the shower if the fraction of hits per wire in the reconstructed object is less than 1.1.	To remove non-shower like particles.
Remove the shower if the true energy of the shower particles is less than 40 MeV.	To remove delta-rays.
Remove reconstructed object with less than 20 hits.	To remove small showers which would be removed during the selection process in Chapter 7.
Only use the largest reconstructed shower in the event	To remove the the showers with a low fraction, which can be seen in Figure 6.10 of the overall energy. Such events skew the metrics used to validate the success.

Table 6.1 The quality cuts applied to the simulated data before analysis the success of the reconstruction

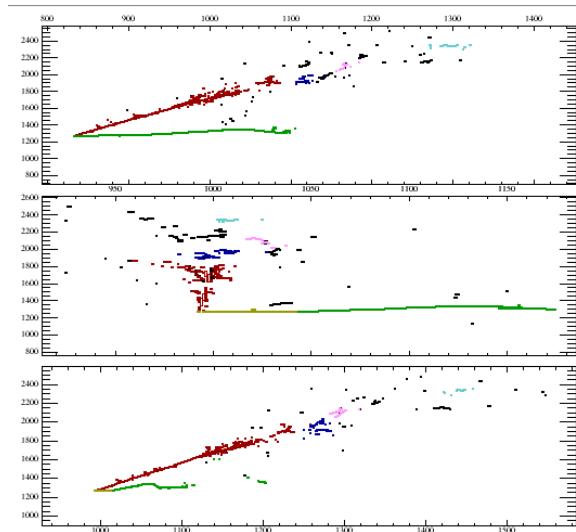


Fig. 6.10 Event display of a reconstructed electron shower (mostly red) and charged pion and its daughters (green and yellow) originating from the same vertex. The shower is slightly segmented, and therefore small shower clusters skew the distributions of completeness metric.

Four metrics were used to define the clustering success:

- The completeness: The fraction of hits (or energy) deposited by the EM particle that is correctly clustered into one object. This is shown in Figure 6.11.
- The purity: The fraction of hits (or energy) deposited by the EM particle within the clustered object. This is shown in Figure 6.12.
- Projection matching: The fraction of energy deposits where the 3D reconstructed position was within 2 cm of the true position. This is shown in Figure 6.13.
- Shower Segmentation:  $\frac{\text{The number of reconstructed showers}}{\text{The number of true showers}}$ .

As discussed earlier, for the *BlurredClustering\_module* reconstruction chain clusters are combined to form 3D objects in the *EMShower\_module* stage [261, 264, 265]. Hence the figures show the performance at the shower characterisations level. As all the shower characterisations except the *BlurredClustering\_module* reconstruction chain using the Pandora clustering, the results are essentially the same. Differences only arises due to the fact that one method the *PandoraShowerCreation\_module* [266] defines the hits of the shower as the hits that were reconstructed to 3D spacepoints rather than all the 2D hits.

The completeness for the pattern recognition degrades at low and high energy showers. This is due to the mischaracterisation of showers as tracks. If this occurs often smaller segments of the shower are only reconstructed as a shower. Hence, the majority of the hits caused by the showering particle are missed and the completeness is low. This is the cause of the low completeness peak in the histograms in Figure 6.11. When this mischaracterisation does not occur there is typically about a 90%. The purity also degrades as a function of the shower energy as the charged pion increasingly overlaps with the shower. However, the purity, shown in Figure 6.12 is usually rather high with the almost all the events being 80% pure or higher. For the *BlurredCluster\_module* case the completeness and purity improves as a function of energy, as merging becomes easier. However, the completeness is significantly lower than the Pandora chains with an average completeness approximately around 25%. The average purity is also less than Pandora chains apart from the higher energy regions where mischaracterisation occurs. Therefore, the conclusion of this analysis is that the *BlurredCluster\_module* is less successful than Pandora at clustering showers in SBND.

Figure 6.13 show the distributions of the geometric metric for the different clustering algorithms. As can be seen in the Figure, the Pandora pattern recognition performs better than the *BlurredCluster\_module*.

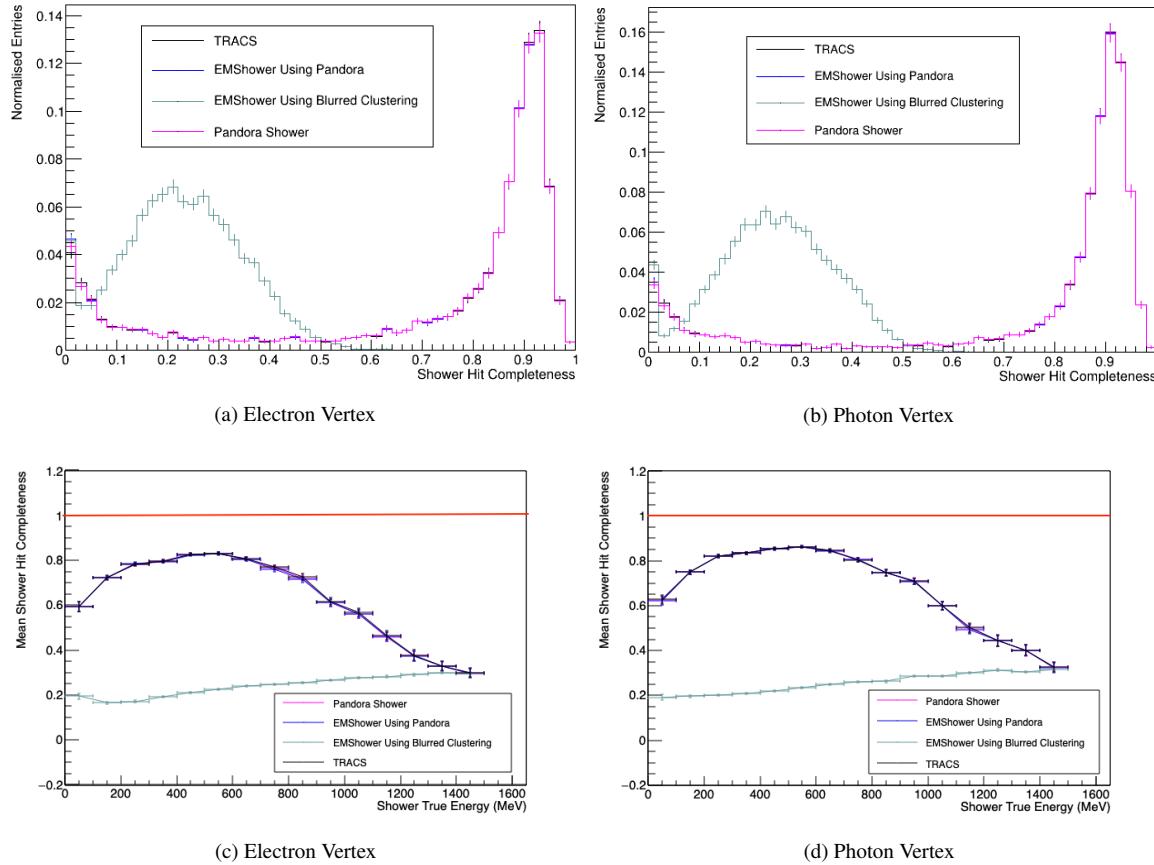


Fig. 6.11 The shower hit completeness. The errors on the histograms are Poissonian. The histograms are normalised to the number of reconstructed events for each method individually. The errors in  $y$  on the energy mean are the standard error on the mean value for events with true energy within the  $x$  error width. The red line corresponds to the ideal value.

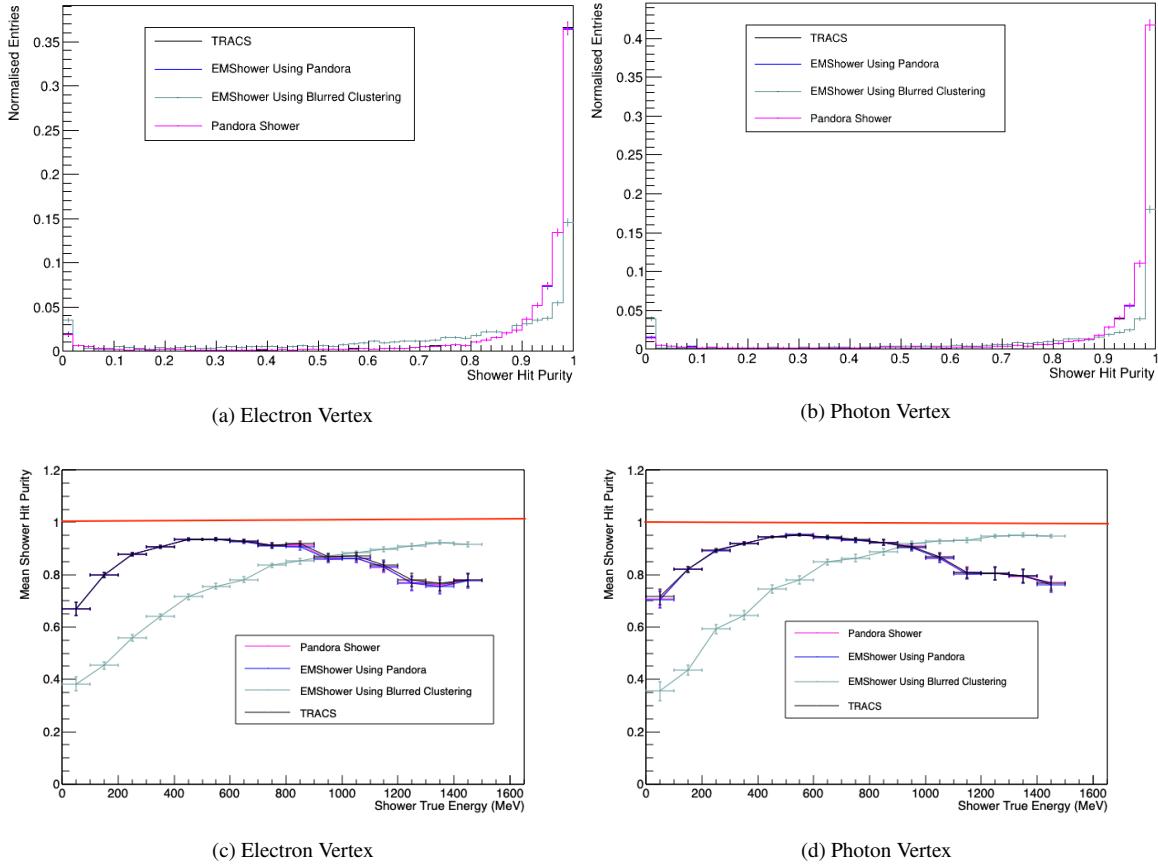


Fig. 6.12 The shower hit purity. The errors on the histograms are Poissonian. The histograms are normalised to the number of reconstructed events for each method individually. The errors in  $y$  on the energy mean are the standard error on the mean value for events with true energy within the  $x$  error width. The red line corresponds to the ideal value.

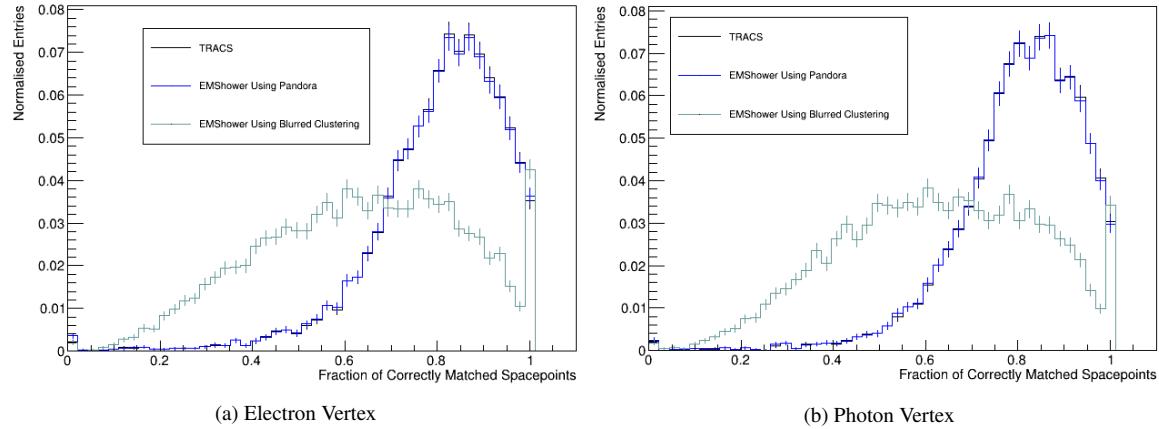


Fig. 6.13 The space-point geometry projection matching score for the collection plane. The induction plane results are similar. The errors on the histograms are Poissonian. The histograms are normalised to the number of reconstructed events for each method individually. The red line corresponds to the ideal value.

The segmentation identifies how often a single true shower is split into many reconstructed particles by the clustering reconstruction. A segmented shower does not necessarily have poor completeness as there could be one reconstructed shower which contains most of the energy. The segmentation of a true shower affects the selection procedures, such as removing events with two reconstructed showers in the event. This cut was performed in the SBN  $\nu_e$  proposal analysis to remove resonant  $\nu$  events with a  $\pi^0$ . The cut removes these events as the  $\pi^0$  decays into two photons but, as can be seen in Figures 6.14, where one true shower exists, performing this cut with no reconstruction quality control cuts can be detrimental to the signal efficiency. This cut is studied in further detail in Chapter 7. As can be seen in the Figures the various shower characterisation algorithms which use the Pandora pattern recognition perform better than the *EMShower\_module* using the *BlurredCluster\_module*.

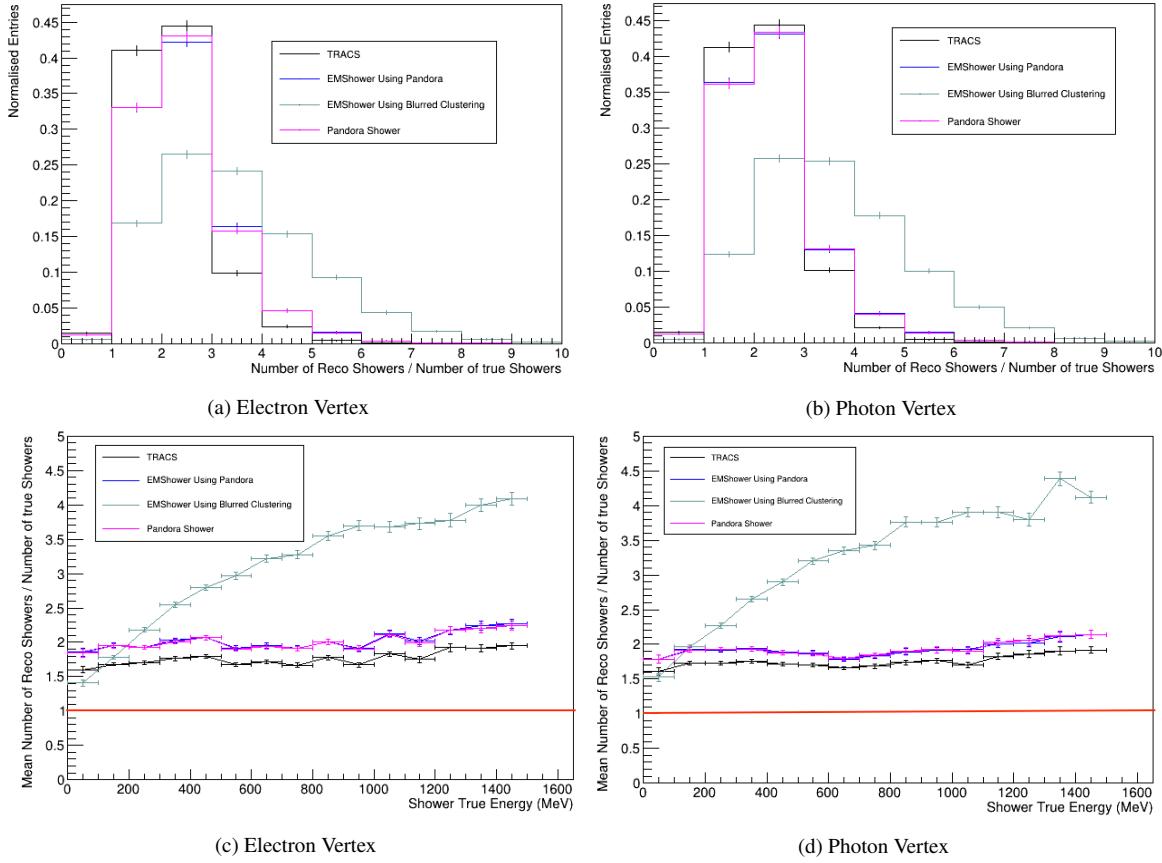


Fig. 6.14 The shower segmentation in the events. If shower reconstruction was perfect, the value should always be one. The induction plane results are similar. The errors on the histograms are Poissonian. The histograms are normalised to the number of reconstructed events for each method individually. The errors in  $y$  on the energy mean are the standard error on the mean value for events with true energy within the  $x$  error width. The red line corresponds to the ideal value.

The clustering ability can also have large impact on how the shower characterisation software, described in Section 6.2 performs. For example, if the clustering fails to cluster the initial track of a shower and the remaining bulk, the starting position and the  $dE/dx$  calculations that occur in the characterisation stage are incorrect. Thus the difference between the different clustering approaches is evaluated in Section 6.2 also.

### 6.1.6 Track Reconstruction

Once the Pandora pattern recognition software has identified track-like objects, an algorithm characterises the 3D direction and position of the track. The information is stored in a

`recob::Track` data object. SBND uses the *PandoraTrackCreation\_module.cc* [267] to make 3D tracks.

The PandoraTrackCreation algorithm takes a collection of 3D spacepoints associated with a track-like PFParticle and performs several linear regression fits to characterise the track trajectory. Initially, a principal component analysis is undertaken in order to identify the general direction of the track. The spacepoints are ordered with regards to their transverse direction along the principal axis. The spacepoints are also grouped into segments, with a width equal to that of the wire pitch (3 mm), along the transverse direction. Using the projected data upon orthogonal directions to the principal axis, two 2D sliding fits are performed. A 2D sliding fit is a series of linear regression fits performed at each segment in 2D space. The linear regression fit at a particular segment uses the ten segments before and after as additional fitting points. The gradients of the linear fits are used to calculate the direction of the track, whilst the gradients and intercepts are used to reconstruct the 3D position. The direction and positional information are stored in a `Track::TrajectoryPoint`.

### 6.1.7 Calorimetry

As described in Chapter 3, LArTPCs have a great ability to identify the flavour of charged particles and the energy of the particles. As will be discussed in Section 6.1.8, a measure of a particle’s  $dE/dx$  is key for particle identification. This Section will discuss the reconstruction of the  $dE/dx$  of particles as well as the energy reconstruction of track-like and shower-like particles.

#### $dE/dx$ Calorimetry

To calculate the  $dE/dx$  for track-like particles, a LArSoft module *Calorimetry\_module.cc* [268] is used. The  $dE/dx$  (MeV/cm) is calculated for each hit separately from the  $dQ/dx$  (ADC/cm) which is calculated as the area under the hit ( $dQ$ ) divided by the track pitch. The track pitch ( $dx$ ) defines the distance the particle travels within the wire pitch and is calculated from the direction of the track at the energy deposition using the `Track::TrajectoryPoint` as well as the wire pitch (3 mm). The  $dQ/dx$  (ADC/cm) is converted into a measure of electrons using a calorimetry constant. The constant is calculated by considering the detector response to a sample of MIP muons travelling parallel to the wire planes and perpendicular to the wires. The calculation of the constant also uses the most probable value of the  $dE/dx$  distribution for a MIP muon in liquid argon: 1.9 MeV/cm. After the calibration constant is applied, a lifetime correction is applied to account for losses due to impurities, see Chapter 5 for more detail on this correction. This lifetime correction allows for the number of electrons

deposited post recombination to be calculated. Finally, the modified box model, see Chapter 3, is applied to account for the recombination and to calculate the energy deposited in MeV.

For shower-like particles the  $dE/dx$  of the initial track stub, identified in Figures 6.24, is calculated. Once the track-stub is identified, see Section 6.2.2, the procedure to calculate the  $dE/dx$  is the same as for track-like particles. In the case of shower-like particles, the median  $dE/dx$  value of the hits is taken as the  $dE/dx$  of the shower for each plane.

### Energy Reconstruction

The energy reconstruction of track-like particles is trivially calculated as the sum of the  $dE/dx$  multiplied by the track pitch at the point of the track. A method based on multiple Coulomb scattering [269] can instead be used which will be especially useful for particles which exit the TPC. Investigations into employing this technique in SBND are being made.

For shower energy reconstruction, the total charge deposited by the shower is converted into an energy estimate using a linear calibration constant. More information on shower energy reconstruction can be found in Section 6.2.2.

The neutrino event energy is reconstructed as the sum of all the particle energies associated with the interaction. For simple topologies, such as charge current quasi-elastic scattering, kinematics [270] can be applied to reconstruct the energy of the neutrino. However, this is not currently performed in SBND. The reconstructed neutrino energy spectrum that was used in Chapter 7 and 8 is discussed in Appendix F.

#### 6.1.8 Particle Identification

Particle identification is based on the  $dE/dx$  of particles. Two methods which have been implemented in SBND for track-like particles are the  $\chi^2$  minimisation of the  $dE/dx$  as a function of the residual range and the PIDA method. Both identification methods are evaluated by the *Chi2ParticleID\_module.cc* [271] in the LArSoft framework.

The  $\chi^2$  minimisation method models how the  $dE/dx$  varies as a function of the distance from the end position of a particle, known as the residual range. It has been shown that stopping particles obey a power law [28]

$$\frac{dE}{dx} = AR^b, \quad (6.7)$$

where  $R$  is the residual range and  $A$  and  $b$  are fit parameters. Figure 6.15 shows the predicted  $dE/dx$  against residual range distributions for various particle types in liquid argon. The

particle is identified by comparing theoretical predictions to the  $dE/dx$  vs residual via a  $\chi^2$  minimisation best fit [28]. The particle identification is performed for each wire plane.

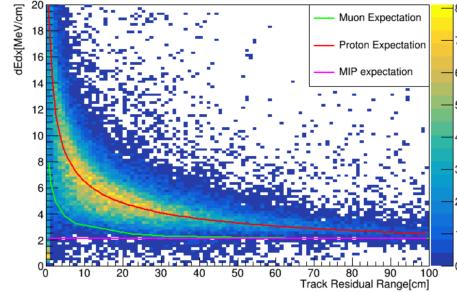


Fig. 6.15 Particle identification using the  $dE/dx$  vs residual range. From Reference [28].

The PIDA method also uses the power law, Equation 6.7, to identify particles. In this method a value,

$$A = \frac{dE}{dx} R^{0.42}, \quad (6.8)$$

is calculated for each wire hit [29]. 0.42 refers to the best fit parameterisation for protons measured at NIST [29].

Figure 6.16 demonstrates that a cut can then be placed on the average PIDA to identify particles as each particle will have a differing PIDA score. The Figure shows that muons and pions are the hardest to separate as they have a similar PIDA score. It also shows that protons have a significantly different PIDA score. This allows protons to be identified relatively easily, and thus prevents the removal of  $\nu_e$  events via a track length cut in the selection, discussed in Chapter 7, Section 7.4.4.

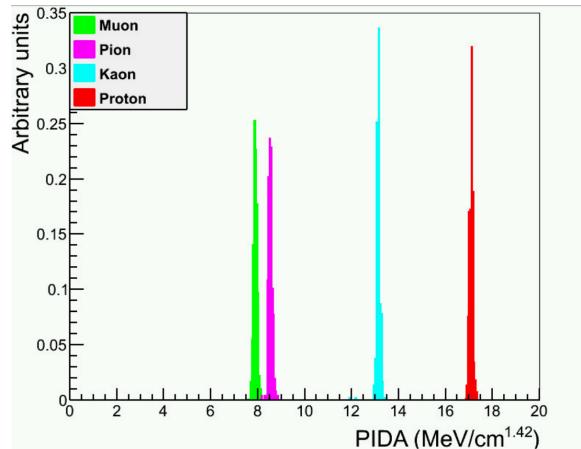


Fig. 6.16 Particle Identification using the PIDA method. From Reference [29].

Electromagnetic shower-like electron photon separation is performed using the median  $dE/dx$  of the initial track stub which is discussed in Section 6.2.2. However, other topological methods will be discussed in Chapter 7.

## 6.2 Shower Characterisation in Liquid Argon TPCs

Once the pattern recognition has characterised particles into track-like and shower-like particles, the shower-like particles undergo a shower characterisation stage. In this reconstruction stage, six key properties of the shower are calculated:

- The start position of the initial showering particle
- The direction of the initial showering particle
- The  $dE/dx$  of the initial track stub of the shower
- The energy of the shower
- The length of the shower
- The opening angle of the shower

The identified properties form the elements of a `recob::Shower` which is saved to the `art::Event` in LArSoft. To evaluate the abilities of shower characterisation software the following metrics were studied:

- **$|\text{True Shower Start Position} - \text{Reconstructed Shower Start Position}|$**
- **$\text{True Shower Direction} \cdot \text{Reconstructed Shower Direction}$**
- The electron photon separational power of the  $dE/dx$
- $\frac{\text{Reconstructed Energy}}{\text{True Energy}}$
- The electron photon separational power of the length.
- The electron photon separational power of the opening angle

The separational power was defined as the maximum value of the efficiency of selecting electrons  $\times$  efficiency of removing photons across all possible values of the metric for the method. This was used in cases where the true value of the property is ill-defined.

In the SBND software framework two existing shower characterisation modules, *EMShower\_module* [261, 264, 265] and *PandoraShowerCreation\_module* [266], were implemented. The former was developed in the DUNE framework by Mike Wallbank [261] and allows the user to input either 2D Cluster objects, a PFParticle or a collection of hits defined as track-like or shower-like by a neural network. The algorithm calculates from this a `recob::Shower`. The shower length and opening angle are the only properties of the shower not calculated by the *EMShower\_module*. The *PandoraShowerCreation\_module.cc* was developed by the Pandora software group and takes advantages of the pre-existing Pandora framework to define the shower. The algorithm, therefore, takes a PFParticle as input. The PFParticle along with its associated vertex and the 3D spacepoints are used to reconstruct the shower. Currently, *PandoraShowerCreation\_module.cc* does not include any calorimetric reconstruction. Both shower modules are discussed in Section 6.2.1.

In order to improve the reconstruction of the shower characterisation in the SBN programme, a framework was developed called the Tool-based Reconstruction Algorithm for Characterising Showers (TRACS), *TRACS\_module.cc* [272]. The TRACS framework was designed to modularise the shower characterisation using `art::Tools` [273], which calculate an element of the shower from a shower-like PFParticle. The user defines a list of tools that the algorithm runs from a configuration (`fcl`) file. This setup enables the tools to be swapped in the reconstruction chain efficiently and allows for developments to be quickly evaluated. More detail on the TRACS framework can be found in Section 6.2.2. A list of 31 tools have been created for TRACS and some of those tools have been developed to reproduce the results from the existing shower modules.

The reader should note that characterising the above properties is not performed in the Pandora pattern recognition software and software such as TRACS is essential for performing selection procedures, such as in Chapter 7. Since the authors contribution Ed Tyley has incorporated the TRACS software into the Pandora umbrella, where it is called *PandoraModularShowerCreation\_module*. The software is now implemented in SBND, ICARUS and the DUNE far detector main reconstruction chains.

To evaluate the performance of the shower characterisation, Monte Carlo samples of electrons and photons where the shower originates from a vertex with a charged pion were created. The energies of the showers are simulated with the typical BNB energies expected for  $e^-$  leptons in  $\nu_e$  interactions. The visible vertex and the forward nature of the events aid the pattern recognition when reconstructing the particles. This allows for the shower metrics to be less affected by the downstream reconstruction and the evaluation shower characterisation can be more easily studied. Again the cuts in Table 6.1 are used before analysing the reconstruction abilities.

### 6.2.1 Overview of Pre-Existing Shower Reconstruction Software in SBND

Two pre-existing shower characterisation modules were implemented in SBND, *EMShower\_module* and *PandoraShowerCreation\_module*. The *EMShower\_module* is flexible in its choice of downstream clustering whereas the *PandoraShowerCreation\_module* is designed to use the Pandora pattern recognition. In SBND the *EMShower\_module* uses the Pandora pattern recognition and the 2D *BlurredClustering\_module*. It should be noted that the preexisting reconstruction can be reproduced using the new characterisation framework TRACS, which will be discussed in Section 6.2.2. Therefore, the next Section will be linked to Section 6.2.2, where the individual calculations are discussed.

#### **EMShower Algorithm**

A detailed description of the *EMShower\_module* algorithm can be found at Reference [264]. The algorithm can run in three modes: using the output of a Convolutional Neural Network (CNN) which identifies hit patterns in the detector as either shower-like or track-like; using a set of 2D clusters and 3D tracks created by preexisting LArSoft algorithms; or using a Pandora 3D shower-like PFParticle object created. The CNN chain has not been currently analysed in the SBND detector.

In the case of providing 2D clusters, e.g. from the *BlurredCluster\_module* described in Section 6.1.4, and 3D tracks, e.g. from the *PandoraTrackCreation\_module* described in Section 6.2.1, 2D cluster objects are matched to the results of the 3D tracking algorithms to make a 3D shower. 3D spacepoints are then created by matching the hits of the 2D clusters on separate planes where the wires overlap and the time of the hits is comparable. After the initial creation of the shower the characterisation begins. The hits in each plane are ordered based on their projected position from the 2D shower centre and the axis of the shower is found using linear regression. The gradient of the RMS of the hit positions from the shower axis is then calculated to indicate the direction of the shower. From the direction of the shower, a rough start position of the shower can be identified. When the Pandora PFParticle is given to *EMShower\_module* instead, the 3D reconstruction has already performed. Therefore, for those cases, the previous stages are not required and a rough start position is given as the closest spacepoint to the most upstream vertex.

After a rough start position is found, the initial track stub of the shower is identified by using a 2D regression-based method, and using the 3D track Projection Matching Algorithm (PMA) method [244], both described in Section 6.2.2. The  $dE/dx$  of the shower is then calculated from the charge depositions in the initial track, the direction of the shower and the StandardCalorimetry algorithm [268]. This is discussed in further detail in Section 6.2.2

For the Pandora reconstruction pathway, the shower start position is defined as the 3D track spacepoint closest to the associated vertex. The direction is then defined as the direction between the 1st spacepoint and 5th spacepoint. For the Blurred Clustering approach, the start position and direction are defined as the start position of the initial reconstructed track and direction of the track. Figures 6.20 and 6.22 show the performance for both methods.

The energy of the shower is calculated using a linear conversion from the total deposited charge to the total energy, as described in Section 6.2.2. The comparison of the true energy to the reconstructed energy using this method is shown in 6.31.

### PandoraShower Algorithm

The *PandoraShowerCreation\_module* is used as an interface between the Pandora PFParticle objects and reconstructed `recob::Shower` object in LArSoft. The module only calculates the topological properties of the shower and not the energy or the initial  $dE/dx$ .

The direction of the shower is calculated via a 3D Principal Component Analysis (PCA) of the 3D spacepoints of the particle. The PCA calculation uses the Eigen package [274]. Figure 6.22 compares the performance of the direction calculation to the EMShower algorithm and the TRACS. The length of the shower is defined as three times the PCA principal eigenvalue defined by Eigen. This corresponds to 99.7% of the shower being confined within this length. Figure 6.33 shows the results using this method. The opening angle is defined as the inverse tangent of the two eigenvectors of the PCA analysis and the results of this definition is shown Figure 6.32. Finally, the start position of the shower is defined as the projected position of the associated vertex onto the primary axis of the shower. The result of which is shown in Figure 6.20.

### 6.2.2 The TRACS Reconstruction Chain and Performance Evaluation

To improve the shower characterisation, the new TRACS framework was developed. The premise of the framework is that the shower property calculations are separated into code blocks which can be easily interchanged. This setup is unlike the pre-existing shower reconstruction chains, which use fixed procedures to calculate the shower properties. This flexibility allows for the best reconstruction chain to be chosen for any experiment. The blocks are then combined into a list via a configuration file to create a flexible shower characterisation calculator. Information is then passed downstream in the tool list so that tools can use previously calculated information. More details of the framework can be found in appendix B. As TRACS can perform the same reconstruction as the pre-existing software it is guaranteed to at least match the ability of the existing software. In addition, the

added complexity of the framework does not increase the CPU requirements compared to the existing tools. The CPU time per event is in the 10-30 ms range for non cosmic events, which is the region of the other software.

Although the TRACS framework can calculate any of the shower properties at any stage of the reconstruction, the most successful tool chain relies on a set order such that previous tools can benefit from downstream information. Typically the start position of the shower is calculated first. This then allows the direction of the shower to be defined using the vector from the start position to the center of the shower. After this, it is then possible to identify hits from the initial track stub of the shower and construct a 3D track. It is then possible to calculate the  $dE/dx$  of the initial track. Therefore in simplest form, the TRACS module requires the pathway shown in Figure 6.17 to identify all the properties of a `recob::Shower`.

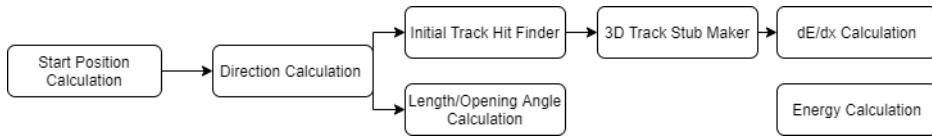


Fig. 6.17 Pathway for TRACS reconstructions

It is then possible to repeat the calculation for particular properties to improve the reconstruction performance. For example, a shower direction tool which specifically uses the initial 3D track direction as the shower direction can be used to redefine the shower direction. A list of the available tools is presented in Table 6.2 along with, the parameter it calculates, a brief description of the tool, a figure which presents the capabilities of the tool and if the tool was developed from the pre-existing shower software.

An extensive analysis was performed to identify the most successful reconstruction path for SBND. This is shown in Figure 6.19 and refers to the tools, presented in Table 6.2 used to calculate the properties of the shower at each stage. It can be seen that the flow follows the prescription above where multiple calculations of the shower properties are performed. The reason for this will be discussed in the next few pages. The tools with a red border were tuned independently of the blue border to maximise the efficiency  $\times$  background rejection for electron photon separation, using  $dE/dx$ ,  $\times$  the fraction of reconstructed events. The tools with a blue border were tuned to maximise the average dot product between the reconstructed and true direction of the showers multiplied by the number of reconstructed events. This tuning procedure is discussed in Section 6.2.2. The resultant metric distributions for the shower properties using the reconstruction chain in Figure 6.19 is shown in Figure 6.18.

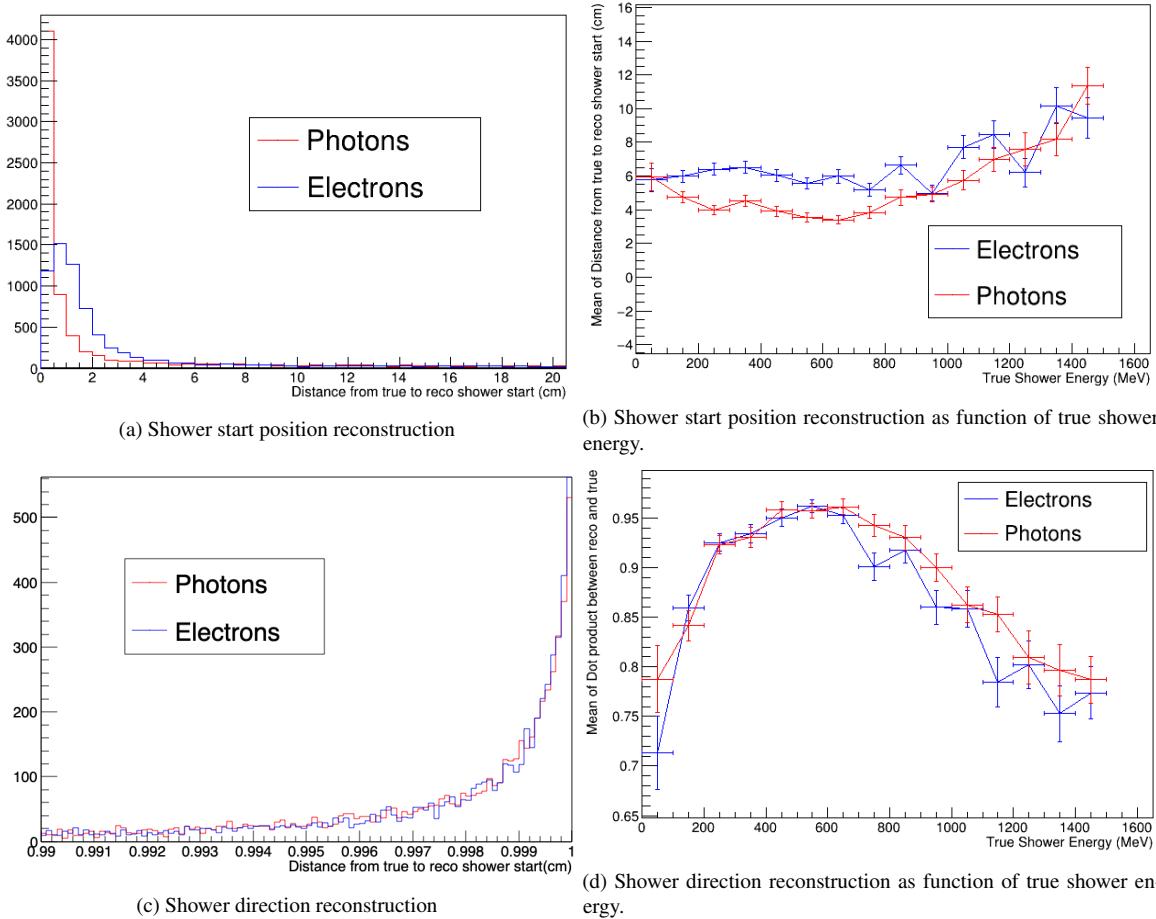
Tool	Property Calculated	Description	Figure Reference	Software Origin
<i>TRACS Pandora Vertex</i>	SP	Use vertex reconstructed by the pattern recognition.	6.20	Pandora
<i>PCA Propagation</i>	SP	Once a PCA analysis is complete, propagate the vertex position onto the PCA primary eigenvector axis of the shower and use this.	6.20	Pandora
<i>Shower Track Start Position</i>	SP	Once the initial track of the shower is defined, use the start of this as the start position of the shower.	N/A	EMShower
<i>Shower PCA Direction</i>	SD	Calculate the shower direction as the primary eigenvector calculated from the hits from a PCA analysis. Orientation is defined as going from the start position to the shower center. PCA analysis is charged weighted which is new in TRACS.	6.22	Pandora
<i>Shower Track Hit Direction, Shower Track Direction, Shower Track SpacePoint Direction</i>	SD	Once the initial track hits are defined, use the average value hits (or space points, or trajectory points of the track fit) to calculate the shower direction. Either use the shower start position or the initial track start position to define the direction to the points.	N/A	TRACS
<i>Shower Track Trajectory Point Direction</i>	SD	Rather than averaging all of the data points, as above, choose a singular data point.	N/A	EMShower
<i>Smart Trajectory</i>	SD	Increment along the trajectory points of the initial track until the direction vector of the three points are within a user defined tolerance. Then take the first point's direction as the shower direction. Either use the shower start position, the initial track start position, or the direction to the previous point to define the direction to the points.	6.22	TRACS
<i>Track PCA Decision</i>	SD	If the direction from the <i>Shower PCA Direction</i> tool and the <i>Smart Trajectory</i> tool are within a user define tolerance, use the <i>Smart Trajectory</i> direction otherwise use the <i>Shower PCA Direction</i> direction.	6.22	TRACS
<i>Shower 3D Cylinder</i>	ITH	Create a cylinder with user defined dimensions that starts from the start position of the shower and follows the shower direction. The hits associated to 3D spacepoints within the cylinder are defined as the initial track hits.	6.25	TRACS

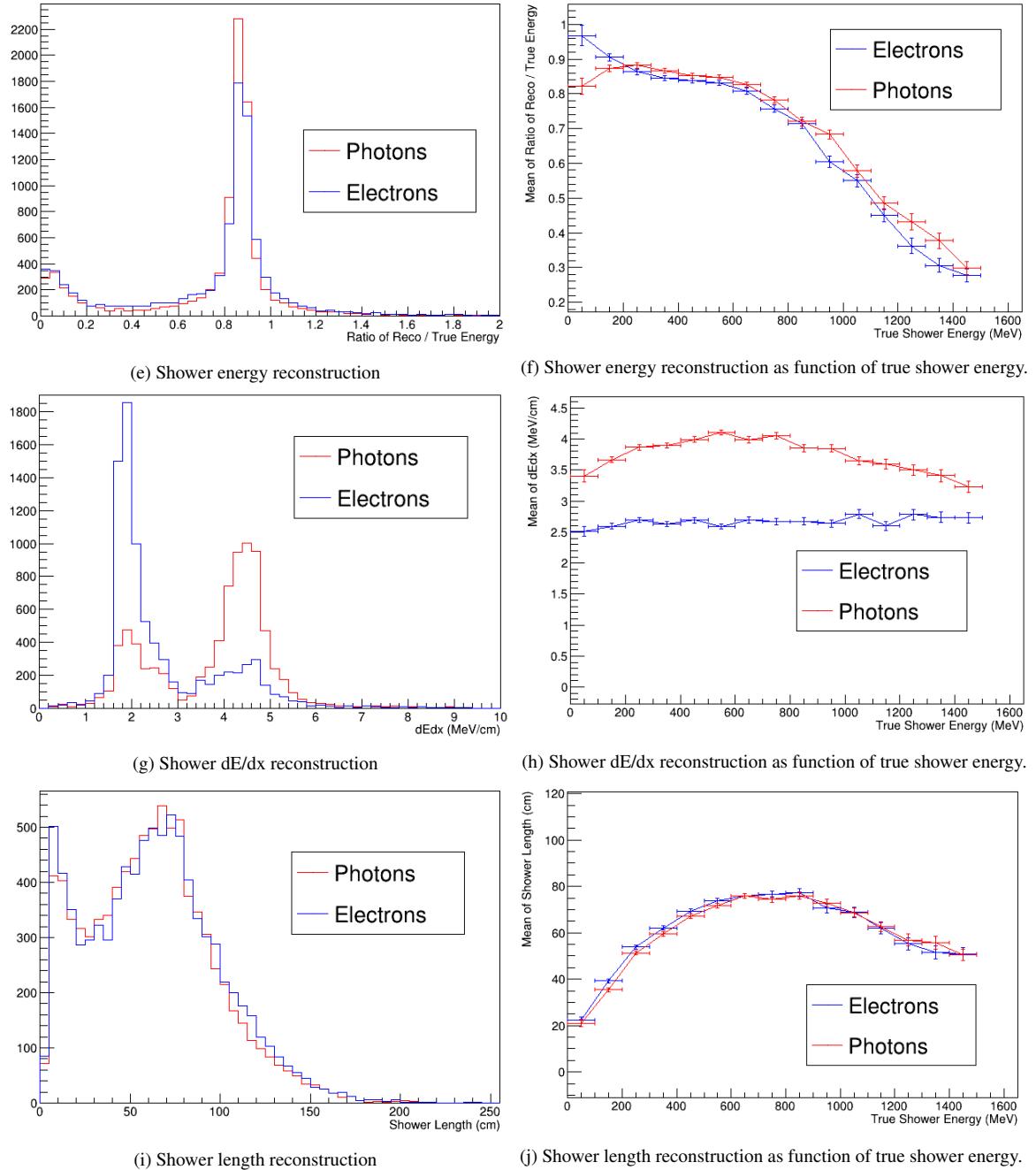
<i>Shower 2D Regression</i>	ITH	Take the initial $N$ hit (defined by the user) and iteratively perform a series of linear fits and remove hits that have a perpendicular distance from the fit greater than a user defined threshold. The threshold becomes more aggressive and the allowed number of hits decrease after each fit. Performed over 2D hit objects. Remaining hits are defined as the initial track hits.	6.25	EMShower
<i>Shower Residual</i>	ITH	Start with a $M$ 3D spacepoints seed and perform a PCA fit. Iteratively remove spacepoints points from the seed until the residual of the fit is below the user defined threshold. Then iteratively add a hit and perform the PCA fit again. Evaluate the average residual and the change in the residual. If both are below a user defined threshold, keep the spacepoint, otherwise skip. Allow $N$ skips and if this is passed try creating a new seed from the skipped spacepoints. If the residual of the new seed is above the user define tolerance stop. Prune spacepoints that are too far from neighbouring points. The hits associated to the spacepoints that remain in the bunch are defined as the initial track hits.	6.25	TRACS
<i>Shower PMA Track Finder</i>	IT	Take the hits from the two most populated planes and creates a straight line using a least-squares regression minimisation process. This line defines the initial track stub. The tool uses the existing <code>ProjectionMatchingAlgorithm</code> (PMA) [244].	6.26	EMShower
<i>Shower Pandora Sliding Fit</i>	IT	Define the 3D direction of the initial track sub using a set of two 2D linear regression fits. The spacepoints are split into several segments and the regressions fits are performed to reconstruct the track trajectory, using the method described in Section 6.1.6. Heavily leans on algorithms defined in LArPandora package [275].	6.26	TRACS

<i>Standard Calorimetry</i>	$dE/dx$	Calculate the $dQ/dx$ using the charge of the initial track stub hits and the track pitch. The pitch is the effective distance the particle travels within the projected wire pitch and is calculated using the shower direction. The first hit and hits above a user define threshold are removed. $dQ/dx$ is converted to $dE/dx$ using the modified box model and calorimetry constants via the <code>StandardCalorimetry</code> algorithm [268]. The median value of all the hits is used to define the shower $dE/dx$ . This process is repeated for each plane.	6.29	EMShower
<i>Sliding Calorimetry</i>	$dE/dx$	Performs the same procedure as the <i>Standard Calorimetry</i> ; however, the $dx$ is calculated using the direction of the reconstructed initial track at several points along the track. Points where the direction is inline with the wire direction ( $< 15^\circ$ ) are removed due to poor hit reconstruction.	6.29	TRACS
<i>Bayesian Truncating</i>	$dE/dx$	The algorithm uses the $dE/dx$ , calculated via the above method. The closest hits to the start position (7) form an initial seed. The mean likelihood of the $dE/dx$ values of the seed for an electron prior (Figure 6.28a) and a photon prior (Figure 6.28b) is calculated. Hits are removed iteratively starting with the least likely hit if the mean likelihood is below a user defined threshold. After this, the algorithm recursively adds hits if the likelihood the hit came from the corresponding prior particle is above a set tolerance. Four hits are allowed to be skipped before this recursive procedure stops to allow for Landau fluctuations. Once the algorithm stops, the median of the distribution with the highest posterior probability from the electron prior or photon prior hits is used as the $dE/dx$ value.	6.29	TRACS
<i>Shower Linear Energy</i>	SE	Calculates the shower energy from the total charge deposited in the reconstructed shower using a linear scale factor.	6.31	EMShower
<i>Eigen Length</i>	SL & SOA	Calculates the shower length and width using the eigenvalues $E$ that result from performing a PCA analysis. Length and width are defined as $2E\sigma$ where $\sigma$ is the number of standard deviations, set to 2. The opening angle $\theta = \tan^{-1}(Width/Length)$ .	6.33	Pandora

<i>Percentile Length</i>	SL & SOA	The 3D spacepoints are ordered with respect to their projected distance from the shower start position on the shower axis. The length is then defined as the distance from the start position of the shower to the last spacepoint within a set percentile (90%) of the shower spacepoints. The width is then calculated likewise but along the secondary axis and the opening angle is defined as above.	6.32	TRACS
--------------------------	----------	---	------	-------

Table 6.2 The list of tools available in the TRACS software. The property that the tool calculates is given ( SP="Start Position", SD="Shower Direction", ITH="Initial Track Hits", IT="Initial Track", SE="Shower Energy", SL="Shower Length" and SOA="Shower Opening Angle"). A description of the tool is also given along with a reference to a figure of merit for the tool. In addition, the origin of the tool methods within shower reconstruction is given.





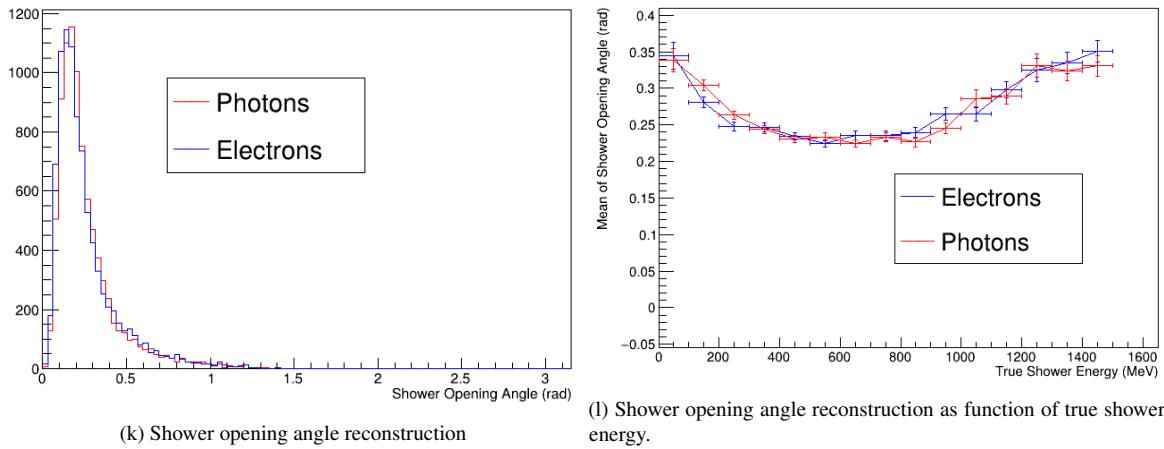


Fig. 6.18 The shower reconstructed metrics for the tuned TRACS tools used in the Analyses in Chapters 7 and 8. The errors on the histograms are Poissonian and the errors in  $y$  on the energy mean magnitude are the standard error on the mean value for events with  $x$  error width. The histograms are the number of entries.

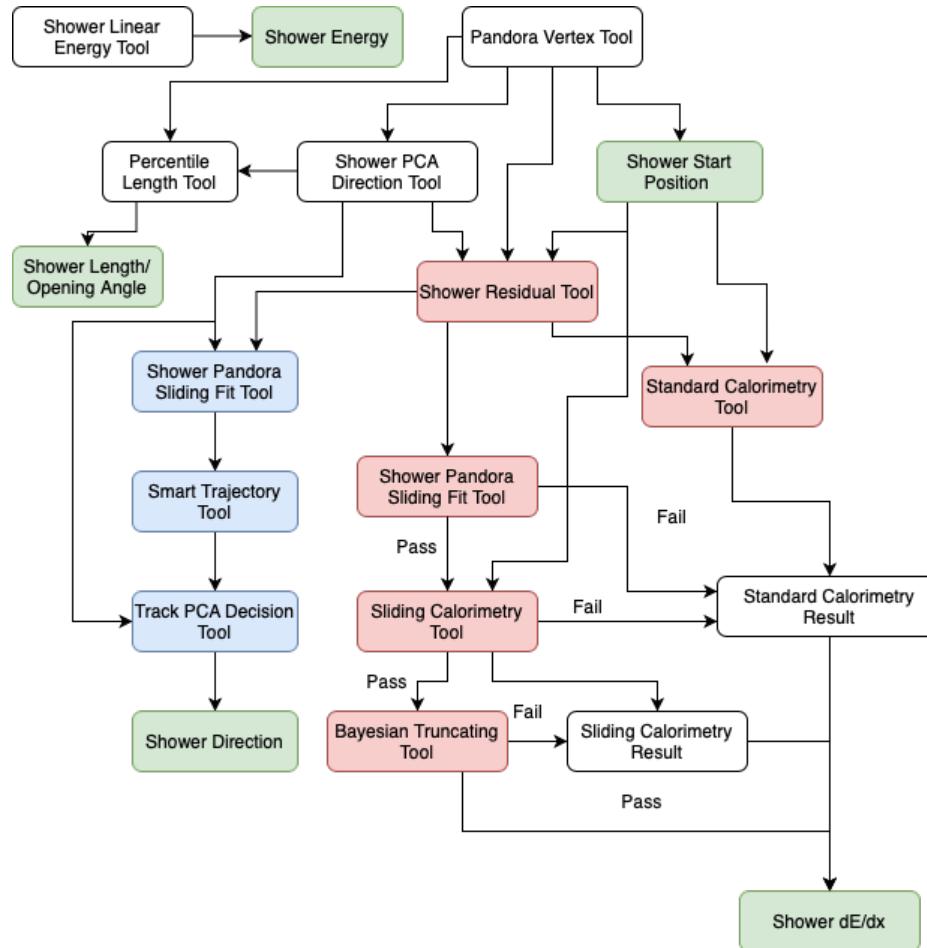


Fig. 6.19 The full TRACS reconstruction pathway. The green boxes correspond to calculated shower variables. The blue and red tools have been separately tuned using the Bayesian Optimisation.

### Start Position Reconstruction Evaluation

The two most successful shower start position tools are *TRACS Pandora Vertex* tool and the *PCA Propagation* tool. For the simulated samples used the *TRACS Pandora Vertex* tool reconstructs the start position of 70% of electrons and 75% of photons to within 5 cm of the true start position. The difference between the photons and the electrons for the *TRACS Pandora Vertex* tool is attributed to the incorrect clustering of hits near the vertex. This occurs predominantly in interactions where the charged pion is isochronous which causes the reconstruction difficulties in projection matching hits. Although in Figure 6.20a the EMShower approach looks to be more accurate the tail of the distribution extends further than the *TRACS Pandora Vertex* tool and less of the events (58%) are within 5 cm of the true start position.

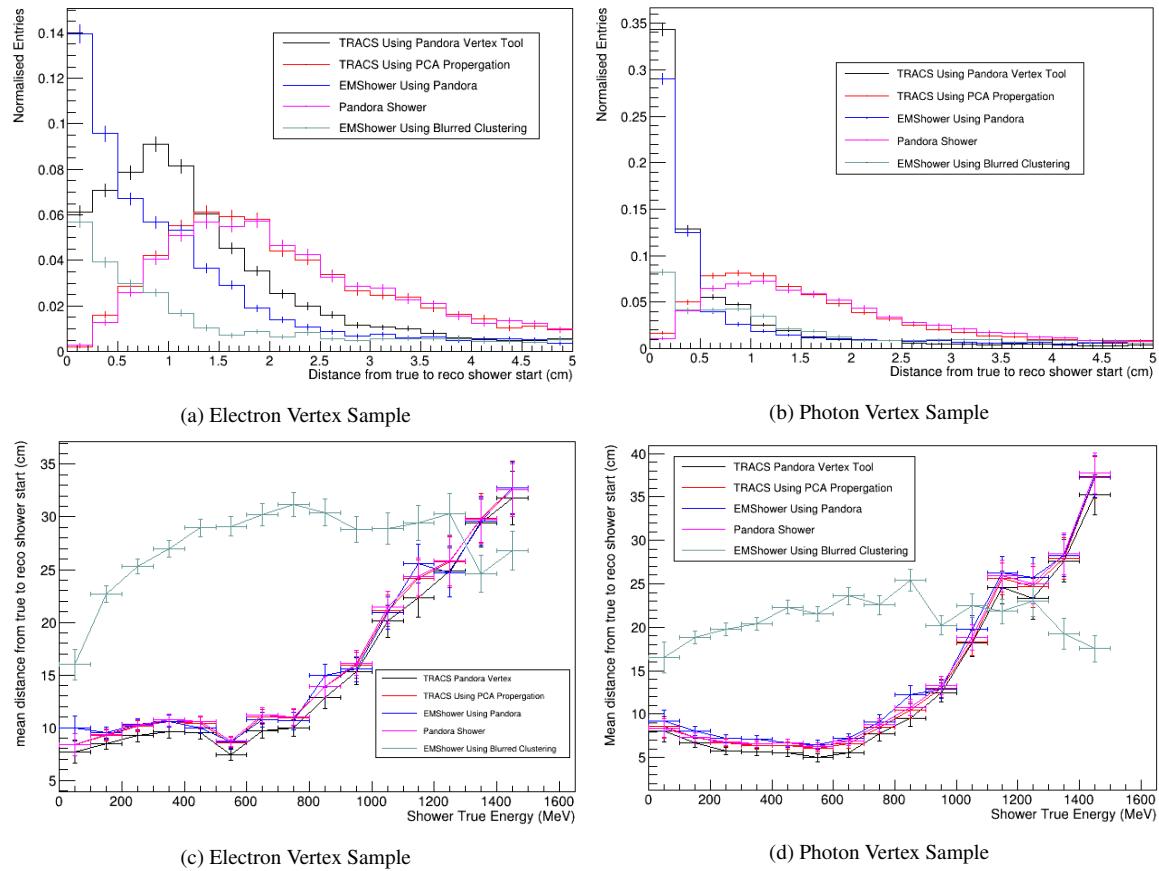


Fig. 6.20 The magnitude of the difference between the reconstructed and true start position of the shower 0-5 cm (a,b). The mean magnitude is also plotted as a function of the true shower energy (c,d). The histograms are normalised in order to consider only the number of events within the range. The errors on the histograms are Poissonian. The error in y on the graphs is the standard error on the beam for the entries with the x error.

Figures 6.20c and 6.20d which show the average difference between the reconstructed and true start position show the *TRACS Pandora Vertex* tool is the most successful at all energies using Pandora pattern recognition downstream produces the best performance. This is to be expected from the low completeness and high segmentation of the *BlurredCluster\_module* discussed in Section 6.1.5.

As the shower start position is the first property calculated, downstream tools are affected by to the performance of these tools, which rely heavily on the Pandora pattern recognition software. The increase in the distance between the reconstructed and true shower start position as a function of energy shown in the Figures is attributed to the mischaracterisation of the bulk of the shower in the pattern recognition. In this case, the reconstructed shower analysed is a small fragment of the shower and can be a large distance away from the true start position. This is cause of the increase at higher energies in the distance from the true and reconstructed start positions in Figures 6.20c and 6.20d. A Support Vector Machine (SVM) has been designed for MicroBooNE to aid the track-shower characterisation. This has been implemented in Chapters 7 and 8. However, the SVM is not implemented in the SBND reconstruction and so is not used in this analysis.

### Direction Reconstruction Evaluation

Once the start position has been calculated the orientation of the shower direction can be defined using a PCA axis of the Spacepoint distribution and the relative direction of the start position to the center of the shower. This is done by *Shower PCA Direction* tool. The performance of the method is shown in Figure 6.22 which shows that 71% of electron and 74% of photon events are reconstructed within  $10^\circ$  of the true direction using the PCA tool.

After the direction and start position are defined one can identify the initial track and the associated hits. From this information several tools can redefine the shower direction defining the initial track-stub direction. For brevity, only one (the *Smart Trajectory* tool) of these will be discussed. This is because there are over 100 different results when using different downstream tools and different configuration of the direction tool. All the different combinations were studied and the reader can request the results from the author if required.

The *Smart Trajectory* tool, is additionally described in Figure 6.21. As mentioned in Table 6.2, the algorithm iteratively progresses through the 3D trajectory points of the initial track stub (the blue triangles in Figures 6.24) until the track is straight. The straightness of the track is evaluated by considering two angles between the vectors of two or three trajectory points and a base point. Figure 6.22 shows the performance of the tool using the directional information and the previous trajectory point as the base point with angle cut of  $3.28^\circ$ . The results are similar to the other track-based methods and show the PCA

performs better on average. However, when looking at events where the angle between the reconstructed and true direction is less than  $11^\circ$ , shown in sub-Figures (c) and (d) in Figures 6.22c, the track-based tool performance is enhanced. Therefore the track direction tools are less accurate but are more precise.

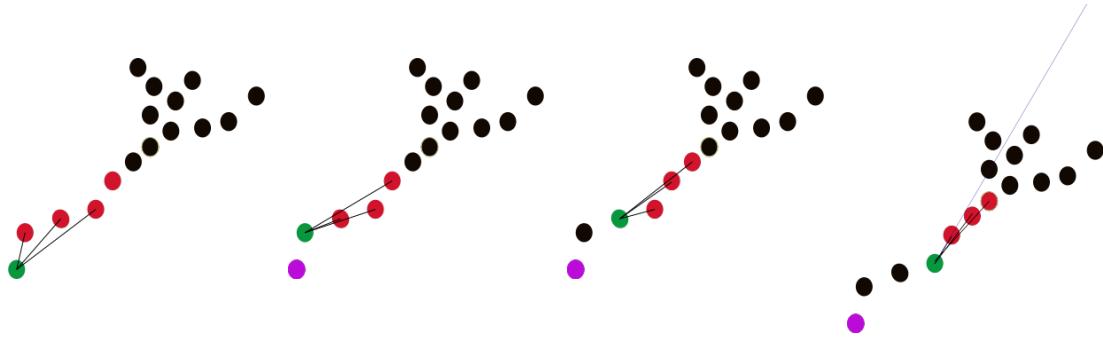


Fig. 6.21 The TRACS *Smart Trajectory* tool algorithm. The algorithm iteratively checks the angles (between the black lines) between the trajectory points (in red) and the base point (in green) until both angles are below the set threshold. The resultant direction is shown as the blue line. The dynamic sliding of the base point can be turned off and the shower start position or the initial track stub start position (pink) can be used instead. If the directional information is used, the last trajectory point is not required.

A *Track PCA Decision* tool was created so that the additional precision from *Smart Trajectory* tool could be coupled with overall performance of the *Shower PCA Direction*. Figures 6.22 shows the decision tool where the track direction is used if the angle between the track direction and the PCA direction is less  $21^\circ$ . As can be seen in Figures 6.22, the decision tool provides an effective compromise at improving the accuracy of reconstructed events within  $25^\circ$  of the true direction whilst maintaining the accuracy for the rest of the distribution.

A comparison of the *Shower PCA Direction* tool, the *Smart Trajectory* tool, the *Track PCA Decision* tool, the *PandoraShower\_module* and the *EMShower\_module* is shown in Figures 6.22. As shown in the Figures the PCA method used also by the *PandoraShower\_module* on average outperforms the track based methods: *Smart Trajectory* tool and *EMShower\_module*. The decision tool provides a compromise between accuracy and precision and so this was used in the reconstruction chain.

The comparison, Figures 6.22, also shows that the Pandora pattern recognition outperforms the blurred clustering approach. Is it to be expected as the track based method of the *EMShower\_module* relies on the correctness of the shower start position which is also outperformed by Pandora (shown in Figure 6.20).

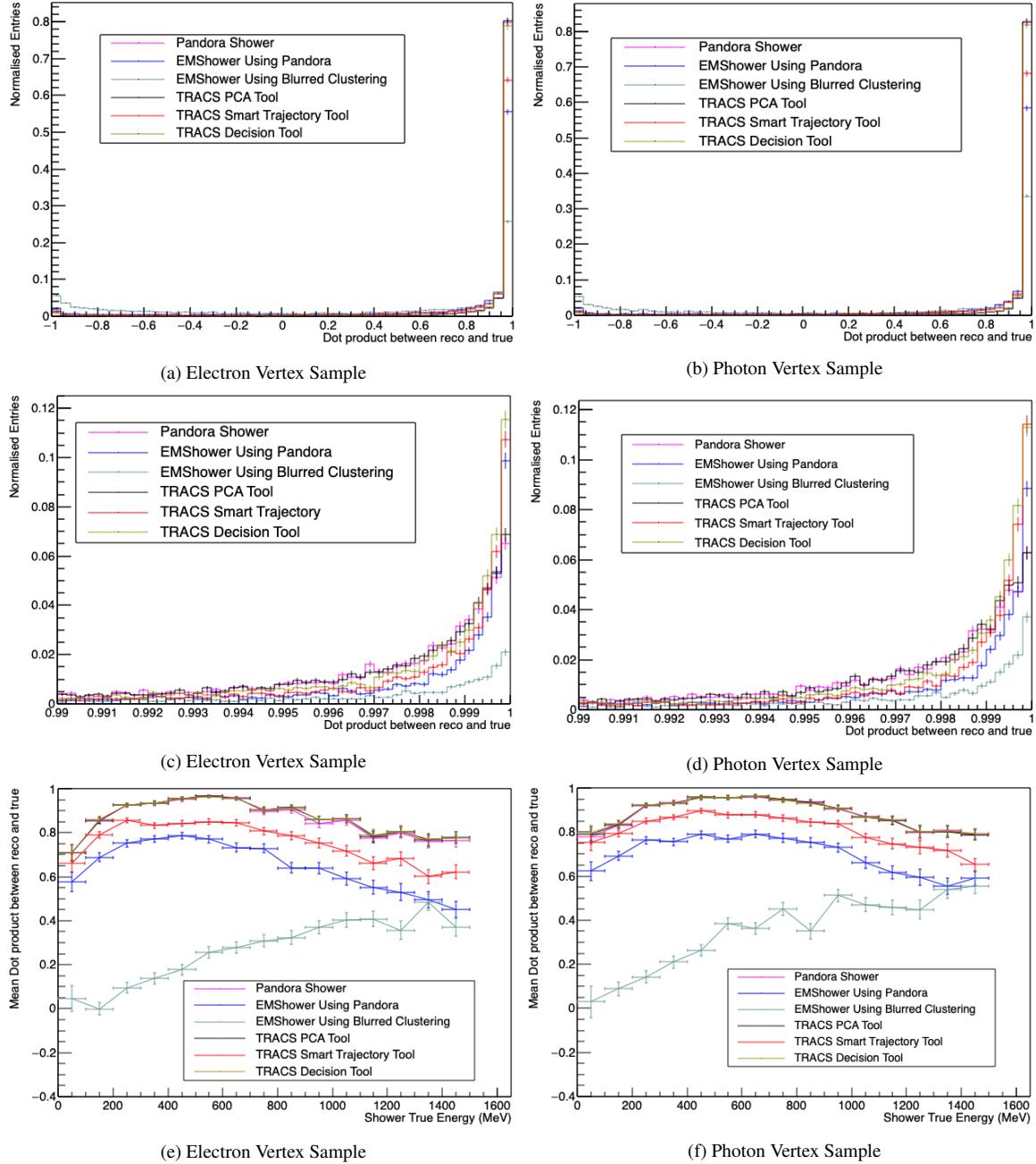


Fig. 6.22 A comparison of the direction calculations of all the existing shower modules. The errors on the histograms are Poissonian. The histograms are normalised in order to only consider the number of events within the range. The errors in  $y$  on the energy mean are the standard error on the mean value for events with true energy within the  $x$  error width.

In conclusion, the PCA analysis provides the best shower direction reconstruction on average. However, in a sub-sample of cases, when the initial track direction tools are successful the result is, on average, more precise than the PCA. Therefore, a decision tool

was created which successfully obtains the PCA average whilst providing a more precise measurement for the sub-sample.

### Initial Track Hit Finder Evaluation

Three algorithms for finding the initial hits of the shower track stub exist in the TRACS framework: the *Shower 3D Cylinder* tool, *Shower 2D Regression* tool and *Shower Residual* tool. The hits found by these algorithms are then used to make a 3D reconstructed track of the track stub and calculate the  $dE/dx$ . For all three, the hits associated to the shower are ordered with respect to their projected distance from the start position of the shower along the shower direction. The specifics of each algorithm are then briefly described in Table 6.2 along with Figure 6.23 which describes the algorithms graphically. An example showing the results of each algorithm on a particular event is show in Figure 6.24. In the figure the red points are the 3D spacepoints (which are associated to the 2D hits) that have been identified as part of the initial track.

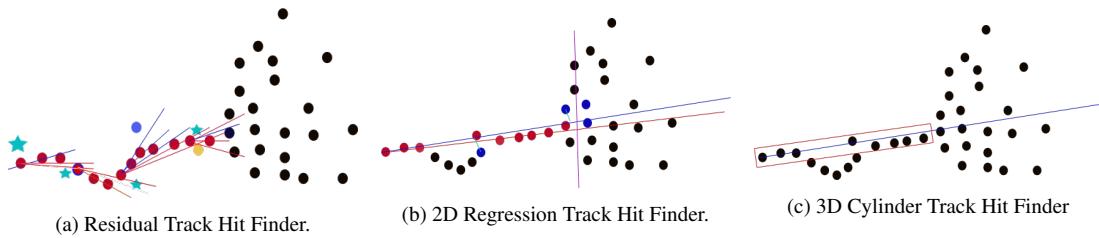


Fig. 6.23 TRACS track hit finder algorithms. For the *Shower Residual* (a) Blue lines indicate successful PCA fits and red lines indicate failures. Red dots with a blue outline indicate hits removed to force the initial seed to fit. Blue hits indicate points which failed the residual cuts, orange hits are removed due to pruning and cyan stars indicate new seed positions. Red points are identified as the initial track stub hits. For the *2D regression* (b) the blue line is the resultant fit of the red and blue hits. The blue points are removed by either the perpendicular tolerance (cyan lines) or the max number of hits cut (after the pink line). The red line is the next fit in the iteration. For the *3D cylinder* tool (c) a cylinder (red rectangle) is created from the start along the shower direction (blue line). All hits within the cylinder are used.

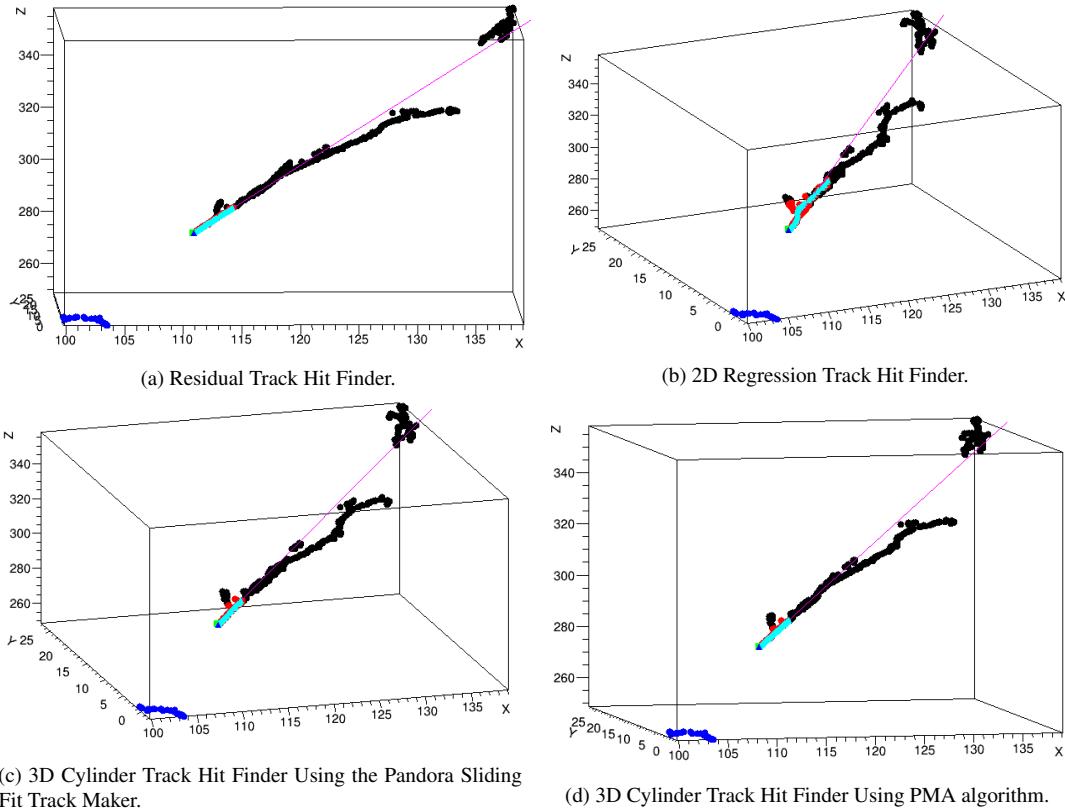


Fig. 6.24 TRACS reconstructed event display for the various track hit finder and track maker tools. The 3D spacepoints of the reconstructed shower are in black. The red 3D points correspond to the spacepoints which have been identified as being part of the initial track stub. The light blue triangles are the 3D trajectory points made by the track making tools. The dark blue triangle is the start position of the 3D initial track stub. The green square is the reconstructed shower start position and the pink line corresponds to the reconstructed shower direction. Other spacepoints in blue correspond to other reconstructed particles.

To identify which method produces the best results the correctness of the shower direction and the electron/photon separation is examined. For a particular reconstruction chain which uses the *Smart Trajectory* tool to define the shower direction (and hence requires the initial track hit information) the *Shower 3D Cylinder* has the most events reconstructed with  $10^\circ$  with fractional value of  $39.4\% \pm 0.7\%$  for the electron sample and  $55.2\% \pm 0.8\%$  for the photon sample. The *Shower Residual* tool has the next highest fraction at  $35.1\% \pm 0.6\%$  for the electron sample and  $50.1\% \pm 0.8\%$  for the photon sample. Lastly, *Shower 2D Regression* tool is the worst performing algorithm with  $30.3\% \pm 0.6\%$  for the electron sample and  $49.4\% \pm 0.8\%$  for the photon sample.

Figure 6.25 shows the performance differences between the hit finder tools when using  $dE/dx$  (calculation from the *Sliding Calorimetry Tool*) to separate electron and photon

showers. This was performed for all events where the shower start position was reconstructed within 1 cm of the true position. This is to ensure the hit finding algorithms are compared for events where they could identify the correct hits and hence the correct  $dE/dx$  could be calculated. As can be seen, there is currently a marginal difference between the methods. This similarity is mainly because only the first 2.4 cm (8 energy depositions) are used to calculate the  $dE/dx$  and, as can be seen in Figures 6.24, there is usually no difference in the hits chosen at that stage. For events reconstructed with 1 cm of the true start position, there is a maximum efficiency multiplied background rejection of  $0.66 \pm 0.04$  (efficiency  $70.5\% \pm 2.1\%$  background rejection  $94.0\% \pm 5.5\%$ ) for the *Shower 3D Cylinder* tool,  $0.64 \pm 0.04$  (efficiency  $69.7\% \pm 2.2\%$  background rejection  $92.4\% \pm 5.0\%$ ) for the *Shower 2D Regression* tool and  $0.65 \pm 0.04$  (efficiency  $70.2\% \pm 2.1\%$  background rejection  $92.35\% \pm 4.9\%$ ) for the *Shower Residual* tool. Although these differences are marginal and within statistical error, for  $\nu_e$  oscillation analysis a small increases in the background rejection is significant. This is because the muon neutrino beam background are very large comparison to the  $\nu_e$  oscillation signal for the current global best fit values, see Chapter 7 for further detail.

Although the *Shower 3D Cylinder* tool is the best performing tool the *Shower Residual* tool has been used in the reconstruction. This is because the *Shower 3D Cylinder* tool predetermines the initial track length (and width). As is discussed in Chapter 7 these can be used to separate electrons and photon showers so it is advantageous to allow an algorithm which determines the length of the track itself.

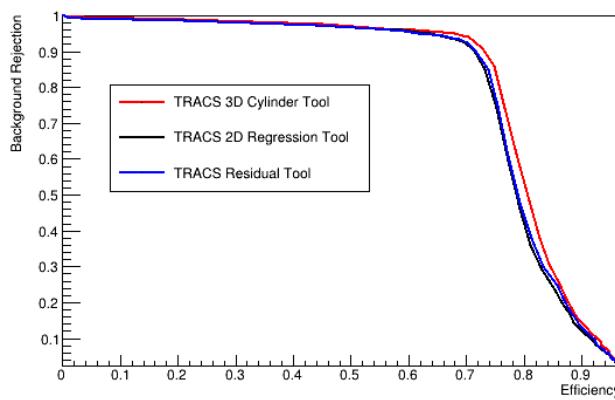


Fig. 6.25 ROC Curves for the three track hit finder algorithms for identifying vertex electrons and rejecting vertex photons using a  $dE/dx$  cut. The shower start position must be reconstructed within 1 cm of true start position.

## Initial Track Reconstruction Evaluation

A 3D track is made from the hits found by the initial track hit tools. Two tools exist which make a 3D track: the *Shower PMA Track Finder* tool and the *Shower Pandora Sliding Fit* tool. Like the track hit finder algorithms the direction and  $dE/dx$  is used to identify the best performing algorithm

Figures 6.26 show the performance of the track making algorithms using the *Shower Track Trajectory Point Direction* tool to identify the shower direction. As can be seen, the *Shower Pandora Sliding* is a better method for calculating the shower direction.

There is no significant difference between the algorithms when considering the  $dE/dx$  calculated from the *Sliding Calorimetry* tool. The efficiency  $\times$  background rejection for the *Shower Pandora Sliding Fit* tool is marginally higher at  $0.657 \pm 0.036$  (efficiency =  $73.7\% \pm 2.2\%$ , background rejection =  $91.12\% \pm 4.0\%$ ) than the *Shower PMA Track Finder* tool at  $0.656 \pm 0.036$  (efficiency =  $73.6\% \pm 2.2\%$ , background rejection  $89.1\% \pm 4.1\%$ ) for events with a start position correctly reconstructed within 1 cm. There is also an loss in reconstruction efficiency of 1.5% when using *Shower PMA Track Finder*.

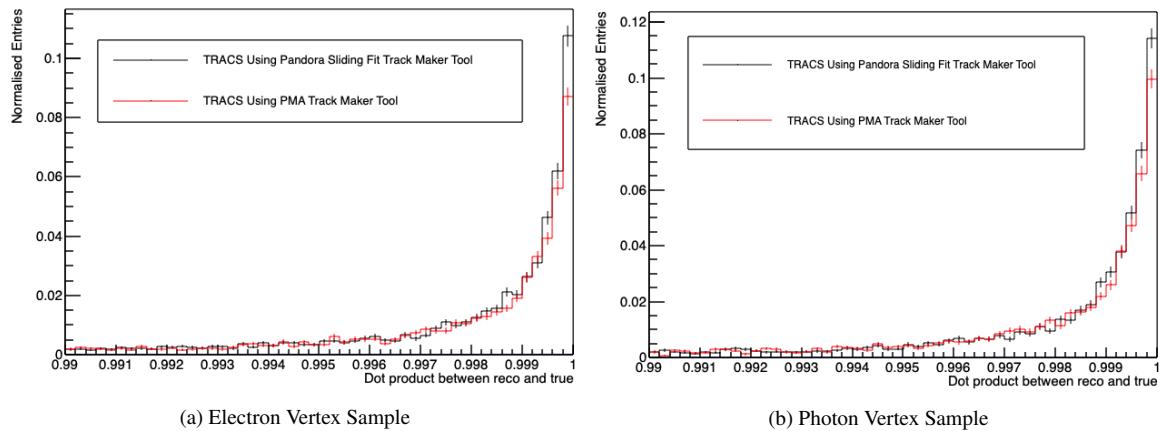


Fig. 6.26 The  $\cos(\theta)$  between the true and reconstructed directions for the various track makers in TRACS using the *Smart Trajectory* tool. The errors on the histograms are Poissonian. The histograms are normalised by the number of events within the range.

In conclusion, the *Shower Pandora Sliding Fit* outperforms the *Shower PMA Track Finder* in both the direction and  $dE/dx$  reconstruction as was thus used in the reconstruction chain.

## $dE/dx$ Reconstruction Evaluation

Using the collective information of the start position, shower direction, initial track and initial track hits there are two tools which calculate  $dE/dx$  of the initial track stub: the *Standard*

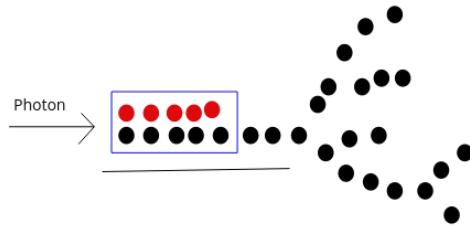


Fig. 6.27 Diagram of photon decaying into  $e^+e^-$  pair. One of the pair stops (red hits) within the initial track stub and therefore a subset of the hits in the initial track have a  $dE/dx$  corresponding to an electron landau (outside the blue box).

*Calorimetry* tool and the *Sliding Calorimetry* tool. The main difference between the two methods is how the  $dx$  is calculated. The *Standard Calorimetry* tool uses the global shower direction to define the  $dx$  for each hit whereas the *Standard Calorimetry* tool evaluates the direction of the initial track at each of the hits to provide independent measurements of the  $dx$ . The  $dE/dx$  of the shower is defined as median value of the  $dE/dx$  distribution of the initial track.

The *Bayesian Truncating* then takes the  $dE/dx$  distribution values, calculated by the calorimetry tools, and uses Bayesian statistic to identify if any of the  $dE/dx$  values should be removed. The principle is to identify a shift in the  $dE/dx$  along the initial track. In the case of a photon shower this could be due to a particle stopping or interacting, as shown in Figure 6.27. For an electron shower, a forward going delta ray could be created causing a shift. In both cases the topology of the initial track could remain unchanged and hence this change could be missed by the track hit finder algorithms and causes an incorrect measurement. The  $dE/dx$  distribution of the track is compared to an electron prior, which can be seen in Figure 6.28a, and a photon prior, which can be seen in Figure 6.28b (both of which were created using a truth based Monte Carlo), to identify which hits should be removed, if any. This decision is based on the likelihood of the hit occurring when assuming the prior distribution. The removal is done independently for both prior distributions and the final  $dE/dx$  distribution with the highest posterior probability is chosen to define the  $dE/dx$  of the shower.

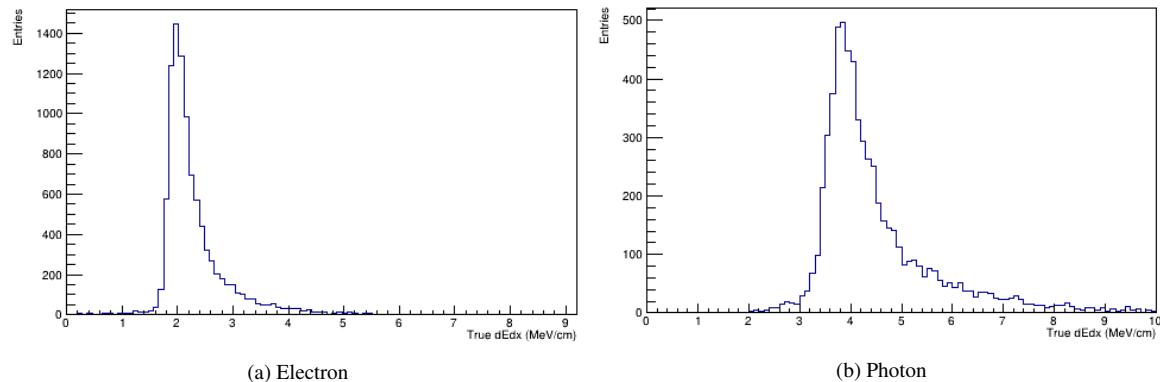


Fig. 6.28 The electron and photon priors used in the *Bayesian Truncating* tool.

The result of the analysis is shown in Figures 6.29 show the different shower calorimetry tools and the *Bayesian Truncating* tool used with the *Sliding Calorimetry*. The *Shower Residual* tool and the *Shower Pandora Sliding Fit* tool were used downstream. The extra precision on the track direction for each hit shifts improves the electron photon separation. For the *Sliding Calorimetry* the maximum efficiency  $\times$  background rejection for the electron photon separation is  $0.508 \pm 0.014$  (at 3.5 MeV/cm efficiency =  $70.0\% \pm 1.2\%$ , background rejection =  $72.27\% \pm 1.7\%$ ) compared to  $0.507 \pm 0.014$  (at 3.3 MeV/cm efficiency =  $70.1\% \pm 1.2\%$ , background rejection  $71.6\% \pm 1.6\%$ ) for the *Standard Calorimetry* tool. The truncating technique of the *Bayesian Truncating* tool forces more events into the peaks of the electron and photon distributions. The *Bayesian Truncating* also provides a higher electron photon separation with a maximum efficiency  $\times$  background rejection of  $0.513 \pm 0.014$  (at 3.1 MeV/cm, efficiency =  $73.5\% \pm 1.2$ , background rejection =  $69.7\% \pm 1.6\%$ ).

Comparing to existing shower modules, the TRACS framework improves on the electron photon separation as can be seen in Figure 6.29. Primarily, the difference between the *EMShower\_module* using the Pandora pattern recognition and TRACS is due to the increased accuracy when reconstructing the start position, which was discussed in the start position Section above. The resultant maximum efficiency multiplied by background rejection for the *EMShower\_module* using Pandora is  $0.477 \pm 0.013$  (at 3.5 MeV/cm, efficiency =  $66.8\% \pm 1.1\%$ , background rejection =  $71.4\% \pm 1.6\%$ ). Due to the downstream reconstruction, the *EMShower\_module* using blurred clustering is the least successful performing algorithm with a maximum efficiency multiplied by the background rejection of  $0.316 \pm 0.8$  (at 3.5 MeV/cm, efficiency =  $57.1\% \pm 1.0\%$ , background rejection =  $55.3\% \pm 1.1\%$ ).

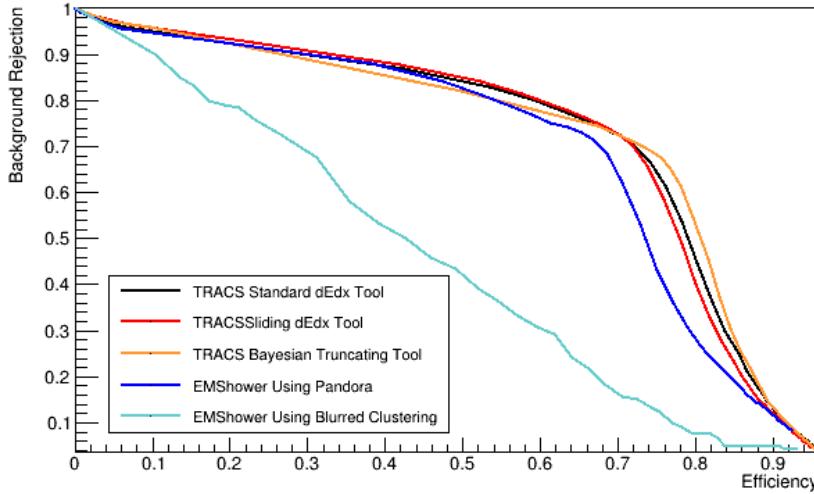


Fig. 6.29 The  $dE/dx$  reconstruction for the various TRACS algorithms and existing shower characterisation algorithms. The errors on the histograms are Poissonian. The histograms are normalised by the number of events reconstructed. The ROC curve corresponds to the efficiency of keeping vertex electrons (a) and removing vertex photons (b).

In conclusion, the TRACS framework using the *Bayesian Truncating* tool results in the best separation of electron and photon showers. The difference between the tools is marginal; however, as mentioned earlier small difference can remove as significant number of background events. Therefore, the *Bayesian Truncating* is used in the SBND reconstruction.

### Shower Energy Reconstruction Evaluation

There is a linear dependency between the total integrated charge and the total energy deposited by a shower. The linearity can be seen in Figure 6.30 and is equivalent to approximating the recombination models to a constant form. The *Shower Linear Energy* tool has been created to take the integrated charge for each hit in ADC, perform a lifetime correction to calculate the charge deposited and then integrate all the deposited charge from the shower. The tool then performs the linear conversion to obtain the energy of the shower in MeV.

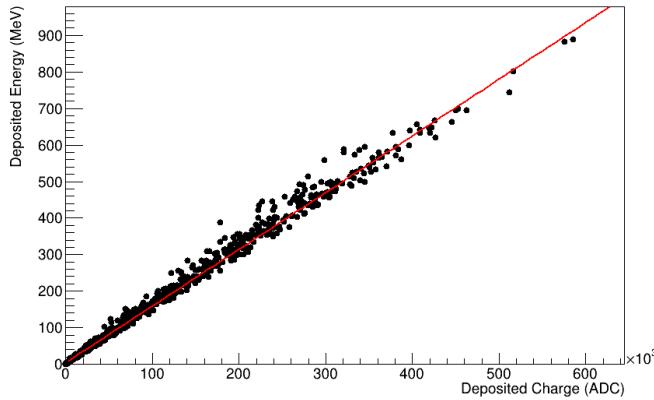


Fig. 6.30 Linear charge energy calibration from an isotropic sample of simulated muon.

The conversion factors are calculated using MC MIP muons where the true energy deposited in the detector is compared to the total charge in the reconstructed muon hits. As MIP muons deposit a similar energy per centimetre to an electron, the average recombination factor is comparable, and therefore muons can provide calibration constants for showers. This method can be used in the detector calibration during commissioning. Other standard candles, such as neutral pion showers, are also being considered for calibration.

Using the conversion factor results in the energy resolution shown in Figure 6.31. The bimodel nature of the distribution is due to incorrectly reconstructed events where hits have been missed in the clustering during the Pandora pattern recognition stage. This feature is also seen in the shower hit completeness metric discussed in Section 6.1.5. The calibration also only partially accounts for the losses in the hit finding as the muon hits account for more of the total energy collectively compared to electron shower hits, see Section 6.1.2. Combining the two effects results in 14.5% deficit from the most common reconstructed energy to the true energy.

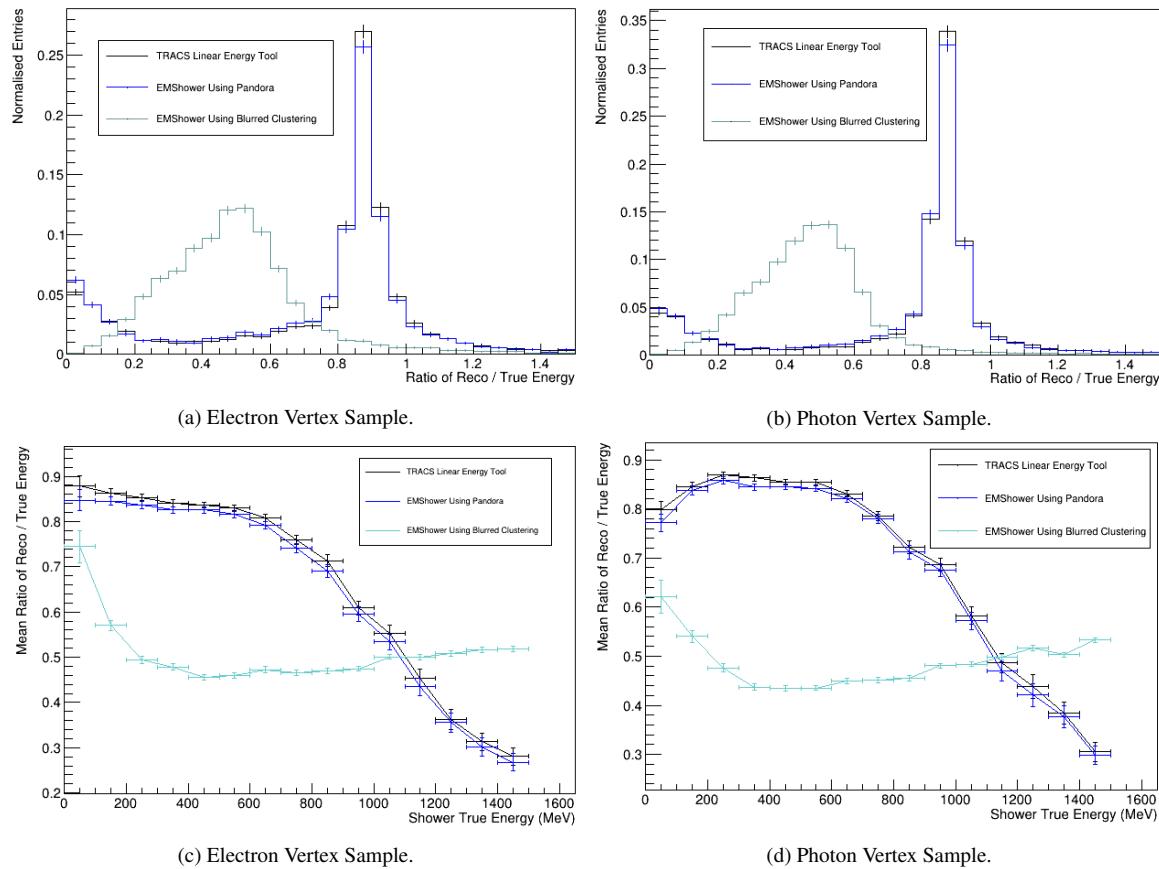


Fig. 6.31 The shower reconstructed energy for the module currently running in SBND. The errors on the histograms are Poissonian and the errors in  $y$  on the energy are the standard error on the mean value for events with true energy within in the  $x$  error width.

The *EMShower\_module* uses the same method for calculating the energy and therefore, as can be seen in Figure 6.31, provides similar results when using the Pandora pattern recognition. However, when using the Blurred Clustering algorithm, the clustering is less successful and the losses are apparent in the energy reconstruction. The *Pandora\_module* has no calorimetric reconstruction and so is not present in the Figures.

### Shower Length and Opening Angle Reconstruction Evaluation

There are two methods to calculate the shower length and opening angle: the *Eigen Length* tool and the *Percentile Length* tool. Figures 6.32 6.33 show the opening angle and shower length for vertex samples respectively. The peak at 0 cm in the shower length corresponds to showers with low hit completeness, which are discussed in Section 6.1.5.

As the shower length is ill-defined in truth the electron/photon separation using the length is used to compare the methods. The difference between photon and electron showers of

the same energy is due to the photon typically undergoing pair production to a  $e^+e^-$  pair and therefore the energy is split between two independent electron showers. The separation power is similar for both methods with an efficiency multiplied by background rejection of  $0.242 \pm 0.006$  (efficiency =  $51.6\% \pm 0.9\%$ , background rejection =  $46.9\% \pm 0.6\%$ ) for the *Percentile Length* tool and  $0.244 \pm 0.006$  (efficiency =  $49.0\% \pm 0.9\%$ , background rejection =  $49.0\% \pm 0.6\%$ ) for the *Eigen Length* tool. As can be seen due to large deviation in the length distribution the separation is poor. The use of the length and opening angle as a useful separation metric is discussed further in Chapter 7.

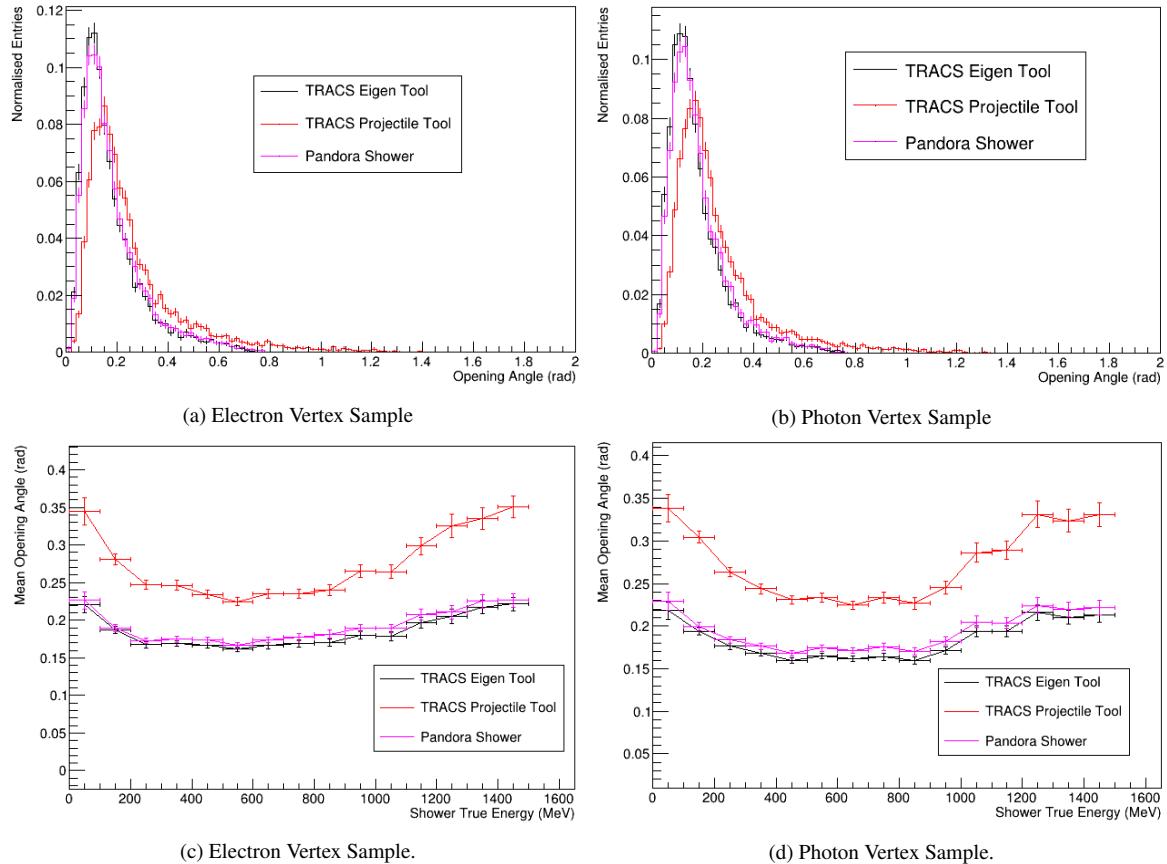


Fig. 6.32 The shower reconstructed opening angle for the TRACS tools and the *PandoraShower\_module*. The errors on the histograms are Poissonian and the errors in  $y$  are the standard error on the mean value for events with energy within  $x$  error width. The histograms are normalised using the number of reconstructed events.

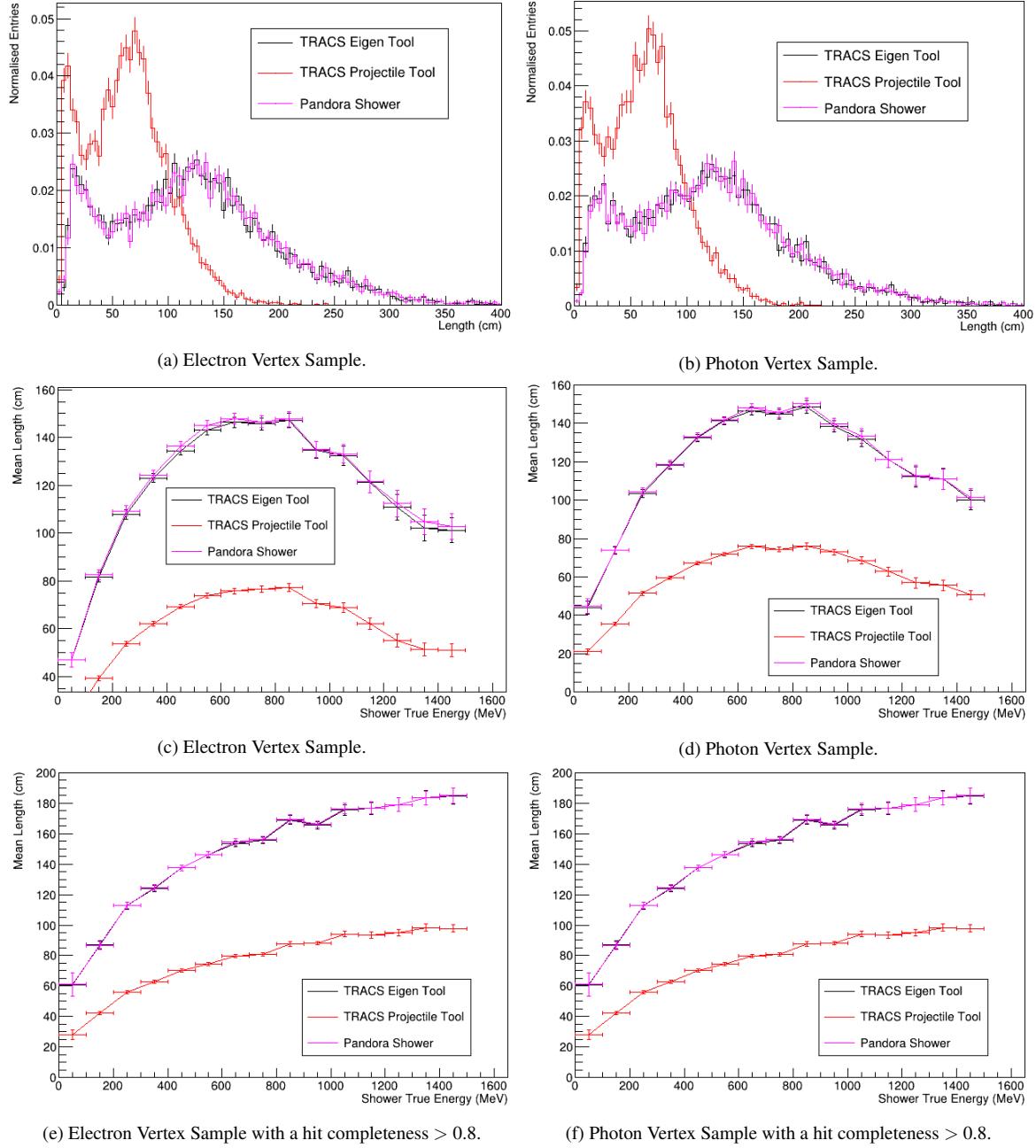


Fig. 6.33 The shower reconstructed shower length for the TRACS tools and the *PandoraShower\_module*. The errors on the histograms are Poissonian and the errors in  $y$  are the standard error on the mean value for events with energy within  $x$  error width. The histograms are normalised using the number of reconstructed events.

In conclusion, the percentile results in smaller length and a larger opening angle. For this thesis a brief study found that the separation when considering all events was comparable

for both methods but the *Eigen Length* tool was used in the reconstruction due to the higher operational power.

### Tuning TRACS

To construct a complete shower, several tools in TRACS are required. Depending on the tools chosen, several parameters must be tuned. For example, for  $dE/dx$  tools a cut off is given for the maximum distance a hit can be from the start position. This cut off is to maximise the purity of the initial track hits used. As the distance is decreased, the purity of the initial track hits improves, but the calculation is influenced more by Landau fluctuation, detector effects and poor reconstruction. Therefore, parameters such as this need to be fine-tuned to maximise the performance of the characterisation. The fine-tuning is undertaken using a Bayesian Optimisation approach [276, 277].

Bayesian Optimisation identifies the parameter set  $\mathbf{x}$  (e.g.  $dE/dx$  track length and others) that maximises a function  $f(\mathbf{x})$  (e.g. electron photon separation). The algorithm for the optimisation, found in Reference [278], can be seen in Algorithm 1. The algorithm is also depicted in Figure 6.34, where a few existing data points have been plotted. The Figure shows that preexisting data provides information about the parameter space. This information is used by an acquisition function, which suggests another point to probe. This point is where the maximum in the hyper-space is most likely. This location is probed and the information is used in the next iteration.

---

**Algorithm 1:** Basic pseudo-code for Bayesian optimization. From Reference [278].

---

Place a Gaussian process prior on  $f$ ;

Observe  $f$  at  $n_0$  points according to an initial space-filling experimental design. Set

$n = n_0$

**while**  $n < N$  **do**

Update the posterior probability distribution on  $f$  using all available data;

Let  $x_n$  be a maximizer of the acquisition function over  $x$ , where the acquisition function is computed using the current posterior distribution;

Observe  $y_n = f(x_n)$ . Increment  $n$

**end**

Return a solution: either the point evaluated with the largest  $f(x)$ , or the point with the largest posterior mean.

---

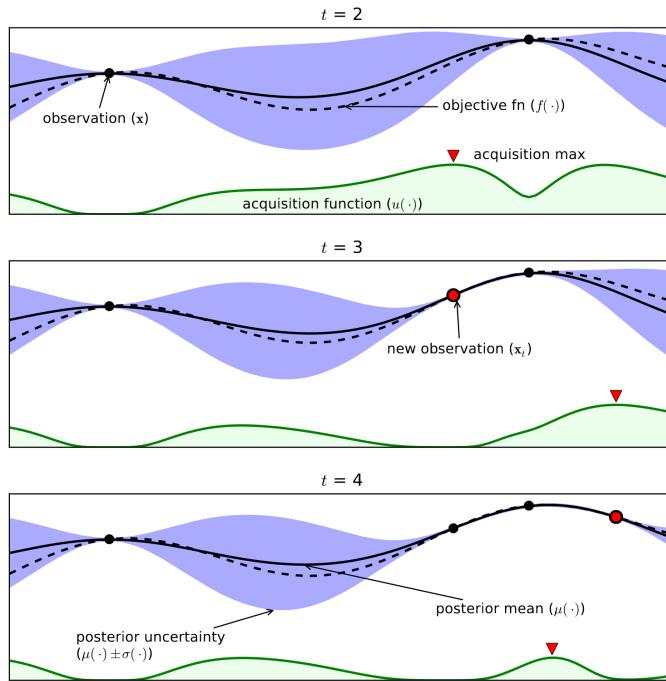


Fig. 6.34 The Bayesian Optimisation process. Where the points are added depends on the maximum of the acquisition function. Once the points are added the posterior probability changes with the new information. Figure from Reference [30].

In more detail, the first step of the algorithm is to assume that  $f$  is described by a multivariate Gaussian distribution. Then the Gaussian process is performed. The Gaussian process [279] in 1 dimension ( $\mathbf{x} = x$ ) states that the result  $f_i$  of each observed value  $x_i$  is derived from a Gaussian distribution which is correlated to the other possible  $x_i$  values through a kernel matrix  $K$  i.e. for two points  $x_1$  and  $x_2$

$$\begin{pmatrix} f_1 \\ f_2 \end{pmatrix} \sim N \left( \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{pmatrix} \right) \quad (6.9)$$

where  $\mu$  is the mean of the Gaussian. The kernel matrix usually takes the form

$$K(x_1, x_2) = \sigma^2 e^{\frac{-1}{2l^2}(x_1 - x_2)^2}, \quad (6.10)$$

where  $l$  determines the smoothness of the function and  $\sigma$  is the strength of the correlation. This form is known the Radial Basis Function (RBF) [279] kernel which ensures the smoothness of the functions.

Due to the marginalization properties [279], it can be shown that the expected value of (or posterior value),  $\mathbb{E}(f_2) = \mu_{2|1}$ , given  $f_1$  is

$$\mu_{2|1} = \mu_2 + K_{21}K_{11}^{-1}(f_2 - \mu_1) \quad (6.11)$$

and the covariance is

$$\Sigma_{2|1} = K_{22} - K_{21}K_{11}^{-1}K_{12}. \quad (6.12)$$

$\mu_{2|1}$  can be calculated at each value of  $x$  and gives the best estimate of the underlying function given previous data. The error on a point in  $x$  is given as  $\Sigma_{2|1}$ . Figure 6.34 shows the error is very small in regions where the data exists and large where it does not. Additional values of known  $x$  can be added by replacing  $x_1$  for a vector of the existing samples and  $x_2$  with the new sample. This, in turn, replaces  $f_1$  with the vector of functions and  $f_2$  with the new posterior that will be calculated. Once a few random samples are taken, the algorithm can move into the while loop and update the posterior iteratively every time a new data point is added [278].

The optimisation occurs as the acquisition function suggests areas of the space,  $x$ , to probe in order to find the maximum value. For the tuning of TRACS, the Upper Confidence Bound (UCB) [280] is used as the acquisition function, which has the form

$$UCB = \mu_{2|1} + \kappa\Sigma_{2|1}. \quad (6.13)$$

The maximum value of this acquisition function is the current best guess on the maximum of the function. Therefore, iteratively choosing the value of  $x$  which maximises the acquisition function results in finding the maximum of the function, see Figure 6.34. The value  $\kappa$  provides a weighting between exploration of the space where the variance is high and exploitation where it is known that the mean value is currently high.

Moving from 1 dimension to N dimension, these arguments still follow. Therefore for use in TRACS, several tuneable parameters are given to the Bayesian Optimisation. A parameter set is suggested (i.e. a  $x$  data point) and the result is calculated, e.g. electron gamma separation (returning the value of  $f$ ). Then a new posterior distribution is made from the result, and the UCB is evaluated. Finally, a new set of values for the parameters is given to try.

Attempts to tune TRACS by evaluating event displays have resulted in a less optimal performance compared to the Bayesian Optimisation tuning. One example of this can be seen in Figure 6.35, where the tuning of the *3D Cylinder* tool to maximise the  $dE/dx$  electron photon separation is performed. The red marker represents the result of the Bayesian Optimisations, the black line corresponds to the best result at the current iteration and the green line represents the experts best guesses based on the track hit length and width, discussed in the track hit finder Section. As can be seen in the Figure several iterations

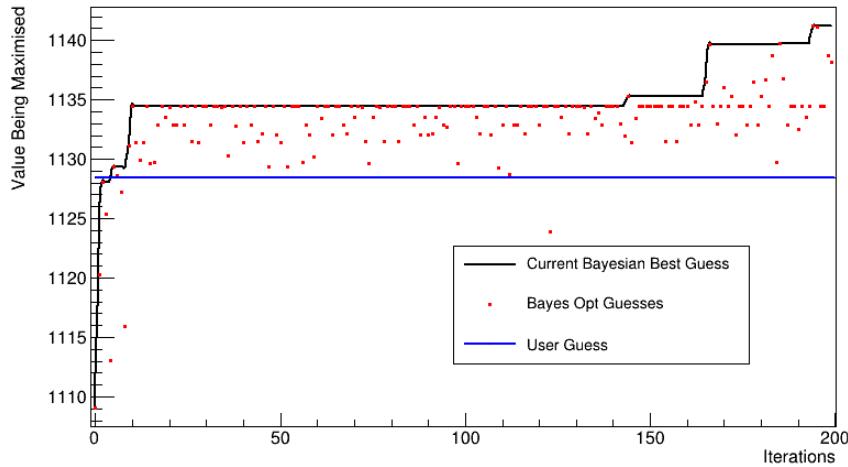


Fig. 6.35 The Bayesian Optimisation process over 200 iterations. The *3DCylinder* tool, see Section 6.2.2, is being tuned to maximise the  $dE/dx$  electron photon separation. The red markers represent the Bayesian Optimisations result at that iteration. The black line corresponds to the best result at the iteration. The green line represents the experts best guess based on the track hit length and width discussed in Section 6.2.2.

are required to find the best results. The tuning takes approximately one day to perform and  $\sim \mathcal{O}(50)$  iterations. Discrete parameters, such as the number of hits, are harder to tune. However, the package [277] used suggested mitigation techniques for discrete inputs, which have been successful.

The SBND TRACS reconstruction chain that has been tuning using the Bayesian Optimisation is shown in Figure 6.19. The resultant shower metric distributions for the simulated vertex electron and photons samples are shown in Figures 6.18. The relevant parameters (e.g. tolerances) set by the tools were tuned concurrently. The tools with a red border were tuned independently of the blue border to maximise the efficiency  $\times$  background rejection for electron photon separation, using  $dE/dx$ ,  $\times$  the fraction of reconstructed events. The tools with a blue border were tuned to maximise the average dot product between the reconstructed and true direction of the showers multiplied by the number of reconstructed events.

## 6.3 Concluding Remarks

This Chapter presents the current SBND reconstruction chain. The reconstruction starts by identifying charge depositions, known as hits, on the waveforms. The method used was discussed in Section 6.1.2 and the efficiency was evaluated in Section 6.1.2. By clustering the hits, 3D reconstructed objects are created using the Pandora pattern recognition, see

Section 6.1.3. From these objects, calorimetric information can be accessed, see Section 6.1.7. Using these 3D objects and the calorimetric information particle identification is possible, see Section 6.1.8. The use of the reconstruction in a selection procedure will be discussed in Chapter 7.

A new shower reconstruction framework TRACS, see Section 6.2.2, has been introduced which modularises the shower characteristic calculation. Doing so provides a flexible framework with an improved shower reconstruction that can be tuned for any liquid argon detector. A full breakdown of the TRACS tools used to calculate the shower properties has been presented along with an evaluation of the performance against existing shower reconstruction algorithms in Section 6.2.2. From this evaluation, a final shower characterisation workflow has been chosen for SBND which has been optimised using Bayesian Optimisation. This has maximised the reconstruction efficiency and the correctness of the shower properties. This characterisation is essential for performing the  $\nu_e$  selection and appearance analyses discussed in Chapters 7 and Chapter 8.

The next Chapter, Chapter 7, will use the reconstruction frameworks described above to identify  $\nu_e$  interactions in the detector and remove resonant  $\nu_\mu$  interactions which are the predominant background.

# Chapter 7

## Producing and Selecting Charge Current $\nu_e$ events in the SBN Programme

As discussed in Chapter 4, one of the main physics goals of the Short-Baseline Neutrino (SBN) programme is to confirm or refute anomalous electron neutrino excesses observed by the LSND [39] and MiniBooNE [143] experiments. To achieve this, the SBN programme will perform an electron neutrino appearance sensitivity measurement based on an electron neutrino selection. In the 2012 SBN proposal [17], truth information from simulated data was used to evaluate the potential sensitivity of the SBN programme. As shown in Chapter 4, the proposal predicted that the SBN programme would be sufficiently sensitive to achieve the goal described above.

Since the proposal analysis, the simulation of the SBN programme has become more robust and there have been changes to the physics model in the neutrino generator, GENIE [225]. Furthermore, some aspects of the SBN detectors have altered since the proposal analysis. Because of these changes, the truth-based proposal selection has been recreated for a new Monte Carlo (MC) sample, known as the "modern" sample.

Because of changes in the software, the truth-based selection had to be verified against the original proposal. To perform this, a "proposal" sample was created by applying weights to the "modern" sample. The production of the "modern" sample and the weightings applied to produce the "proposal" sample will be discussed in Section 7.1. The truth-based selection was then performed on both the "proposal" sample and the "modern" sample, discussed in Section 7.2. The differences resulting from each of these selections are discussed in Section 7.3.

The original proposal selection was then performed using reconstructed data on a simulated sample in the Short-Baseline Near Detector (SBND) in order to evaluate the relative success of the original selection on more realistic data. This is described in Section 7.4.3.

Additional metrics were developed to improve on the proposal selection and these are presented in Section 7.4.4. This selection was then adapted further using a Multivariate Analysis (MVA) [281] to improve the  $\nu_e$  efficiency. This is described in Section 7.4.5. Each of these improved selections will then be compared to the proposal-based selection in Section 7.5.

Following on from this, the potential sensitivity of the SBN programme was reevaluated using the verified "modern" sample and each of the selections discussed in this Chapter. This is discussed in Chapter 8.

## 7.1 Sample Production

### 7.1.1 Producing the Modern Samples

This Section discusses the production of the "modern" sample. This requires the generation of all the backgrounds and the oscillated  $\nu_e$  signal.

The first background generated was the  $\nu_\mu$  events that occur within the detector. The SBN programme expects to take  $6.6 \times 10^{20}$  Protons On Target (POT) of data for SBND and the same amount for ICARUS during the run-time.  $13.2 \times 10^{21}$  POT are expected for MicroBooNE, as this is already running. This amount corresponds to approximately 7 million  $\nu_\mu$  events in the SBND detector. However, a simulated sample of 7 million events is not feasible. Therefore, in order to minimise statistical fluctuation and still be within the computational means of the experiment, 1 million BNB events were created using the event generator GENIE [225] in the LArSoft framework [282]. These events were then scaled by the ratio of the expected POT during the run times of each experiment compared to the POT required to simulate the events. The BNB events created contain all neutrino flavours, in the correct proportions expected in the BNB beam. Of these 1 million BNB events, resonant Neutral Current (NC)  $\nu_\mu$ s with final state neutral pions are one of the most critical backgrounds in the analysis. They account for  $\sim 12\%$  of the  $\nu_\mu$  events generated. See Chapter 4, Section 4.7 for further details on the simulation.

The second background generated was the intrinsic  $\nu_e$  signal. This was necessary as only  $\sim 0.5\%$  ( $\sim 50$  K) of the BNB beam corresponds to the Charge Current (CC) intrinsic  $\nu_e$ . Therefore, a dedicated intrinsic  $\nu_e$  sample of 1 million events was also produced to increase the statistics for this particular background.

The first two backgrounds discussed were produced in all three detectors and corresponded to a total of 6 million events. The simulation generates events through a flux window of  $10 \text{ m} \times 10 \text{ m}$ , 10 m upstream of the front face of the detector. This ensures that the majority of the events occur within the detector or nearby.

A third background of a further 100 M BNB-like events was generated in each detector with a larger flux window of  $80\text{ m} \times 80\text{ m}$  at the face of the detector. This was carried out in this way to incorporate more interactions that occur outside of the detector. Events were only stored if a photon propagated into the active volume and had an energy greater than 10 MeV. These events are known as dirt events. A dirt event can mimic a CC interaction where no track-like (muon, proton, charged kaon, charged pion) final state particles are emitted from the vertex. This procedure resulted in  $\sim 100\text{ K}$  dirt events in the sample, corresponding to  $6.6 \times 10^{20}$  POT in SBND.

A final background of an additional 100 K of dirt events in each detector were overlaid with cosmic events generated by CORSIKA [227]. A cosmic event can mimic a CC event if a high energy cosmic photon interacts within the detector. During the three year run-time of the experiment, 211 seconds of beam spill data will be recorded. This sample corresponds to  $\sim 0.04\%$  of the events in the three year run. The spill time from the BNB is 1200 ns, therefore 0.12 seconds of in-time cosmics were generated in the sample.

To create the oscillated signal sample, 1 million BNB events were generated in each detector with the neutrino flavour swapped. An oscillation weight is then applied within the fitting frameworks, which is discussed in Chapter 8, Section 8.1. For the global best fit of the 3+1 sterile hypothesis ( $\sin^2(\theta_{\mu e}) = 0.003$  and  $\Delta m^2 = 1.2\text{ eV}^2$  [9]), at approximately the peak BNB energy (110 m baseline at 1 GeV), the oscillation probability at the SBND detector is  $\sim 4 \times 10^{-4}$ . The oscillation probability is given as a weight to the event, and therefore the sample has a weight corresponding to  $\mathcal{O}(100)$  events. This sample was also generated through a flux window of  $10\text{ m} \times 10\text{ m}$ , 10 m upstream of the front face of the detector.

Note that the simulation also includes Meson Exchange Current (MEC) interactions which correspond to neutrino interactions with pairs of hadrons, see Chapter 2 Section 3.6 for further details on the cross-section model.

The full POT for each sample can be found in Table 7.1. The generated events described here are then defined as the "modern" sample. The next Section will describe how the "modern" sample is altered to the "proposal" sample.

Sample	Experiment POT		
	SBND	MicroBooNE	ICARUS
$\nu_\mu$ BNB	$4.52 \times 10^{18}$	$9.83 \times 10^{20}$	$1.19 \times 10^{20}$
$\nu_e$ BNB Intrinsic	$5.29 \times 10^{21}$	$1.51 \times 10^{23}$	$1.78 \times 10^{22}$
$\nu_e$ BNB Oscillated	$4.49 \times 10^{19}$	$8.78 \times 10^{20}$	$1.09 \times 10^{20}$
Dirt	$6.18 \times 10^{20}$	$5.39 \times 10^{21}$	$2.82 \times 10^{21}$
Cosmic	$7.67 \times 10^{17}$	$7.68 \times 10^{17}$	$2.78 \times 10^{18}$

Table 7.1 The number of protons on target for each sample generated.

### 7.1.2 Producing the Proposal Samples

In order to verify and compare the "modern" sample generated in the previous Section, a "proposal" sample was created. This was achieved by applying weights to the "modern" sample to account for changes to the experiment and the physics model that have occurred since the proposal. From the two samples a comparison can then be made to identify how changes the understanding of neutrino interactions has affected the  $\nu_e$  oscillation analysis.

The first of these changes is that SBND has moved since the proposal era from 100 m along the test beam to 110 m. This difference in position results in a change to the flux. To account for this, a weight of 1.21 is applied to all SBND events. This scaling assumes the beam is uniformly emitted as a cone [95]. Other changes in the flux simulation during this era were not considered in the analysis.

The second change is that MEC events were not generated in the proposal era but are generated in the modern era. Therefore all MEC events were removed from the "modern" sample during the selection to produce the "proposal" sample. As the addition of the MEC events increases the overall cross-section of the neutrino interactions, no rescaling of the interaction rates needs to be considered.

The third change accounted for is that of the change to the physics model from GENIE version v2\_8 to v2\_12. There are several differences to the model which are documented in the GENIE version release notes [283]: small tweaks in the resonant models in version v2\_12; minor bug fixes in coherent pion model in version v2\_8\_2; and in v2\_10, there was a re-tune of the parameters used for producing the Deep Inelastic Scattering (DIS) cross-section spline. To account for these differences, a weight is applied to the events as a function of the neutrino energy for each interaction mode [284]. The GENIE release notes [283] do not give further information on the effects of these changes and additional changes can be found in the release notes.

To determine the weight applied to account for the change in physics model, the following work was carried out: 200 K of BNB-like events were generated using v2\_8 and v2\_12. The rate difference between the samples was calculated as a function of the true neutrino energy for the interaction modes. The resultant ratio was used as the GENIE weight. For each event the interaction and true neutrino energy was used to identify which weight to scale the event by in the final selection.

For the  $\nu_e$  selection, the modes that are important when considering the signal are the charged current events where the electron candidate is visible. Quasi-elastic CC events with zero final state charge pions (CC0Pi), are the most common signal event corresponding to approximately 62% (with MEC events removed) of the CC events. The version ratio is shown in Figure 7.1. There is a 6.3% decrease in the total event rate of CC0Pi events from the BNB in the modern era (v2\_12) by comparison to the proposal era (v2\_8). This loss is relatively constant as a function of the neutrino energy.

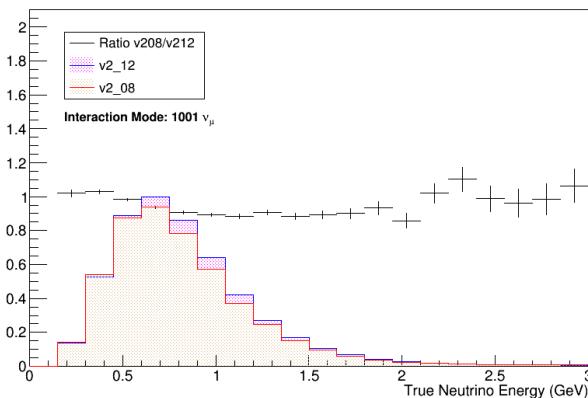


Fig. 7.1 The interaction rates of CC0Pi events in GENIE v2\_8 and v2\_12. The errors on the plot are Poissonian.

One of the main backgrounds for a  $\nu_e$  selection is photon showers arising from a final state neutral pion decay from neutral current interactions. The ratio for resonant interactions acting upon a proton and resulting in a neutral pion is shown Figure 7.2. This is one particular interaction that can result in neutral pion and others are discussed in Appendix C. In total there is a loss of 7.7% from the "modern" sample to the "proposal" sample for NC resonant events, with final state neutral pions.

Other interactions modes and differences in the final state particles are discussed further in Appendix C. A comparison between the weight "proposal" sample and the non-weighted "modern" sample is discussed in Section 7.3.

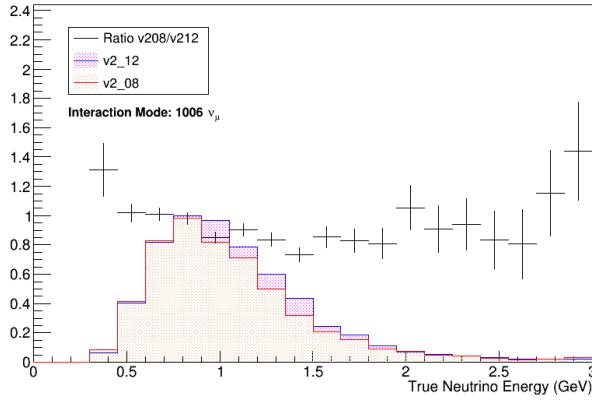


Fig. 7.2 The interaction rates of NC resonant interactions on protons producing a final state neutral pion for GENIE v2\_8 and v2\_12. The errors on the plot are Poissonian.

## 7.2 Truth-Based Proposal Like Selection

The following Section will discuss the application of the truth-based proposal like selection to both the "proposal" and the "modern" sample created in the previous Sections.

Firstly, Section 7.2.1 will discuss the application of the selection to the beam related signal and background events which occur in the active volume. The process is the same for both samples and the results for each sample is also discussed in this Section. Following this, the application of the selection to dirt and cosmic events, as well as the results for each sample, will be discussed in Sections 7.2.2 and 7.2.3 respectively. Note: where efficiencies are presented in these Sections, the proposal sample efficiency can be found within the text and modern sample efficiency is presented in brackets.

### 7.2.1 Beam Induced Active Volume Signal and Background Events

Firstly, for the non-dirt based samples, an active volume cut is applied to the vertex to remove all events which do not interact in the liquid argon.

Firstly, for the non-dirt based samples, an active volume cut is applied to the vertex to remove events outside the liquid argon. Showers arising from the vertex are identified and the reconstructed energy of the showers is found by smearing the summation of the ionisation depositions in the argon. A smearing corresponding to a resolution of  $15\% \sqrt{E(\text{GeV})}$  is applied [17]. Then, if more than one shower with a smeared energy above 100 MeV exists, the event is removed. The purpose of this cut is to remove neutral pion events where the pion decays into two photon showers. The branching ratio for this decay 98.823% [92] hence

the cut in theory removes almost all the neutral pion events. However, the photons showers can be incorrectly reconstruction or the decay can be very asymmetric which diminishes the background removal rate of the cut. The cut also reduces the selection efficiency by removing higher energy  $\nu_e$  CC resonant events. The effectiveness of this cut at removing background can be seen in Figure 7.3. The efficiencies and background rejection for both samples are presented in Table 7.3. As can be seen in the Figure the background rejection increases as function of neutrino energy. This is because resonant events become more dominate. Hence, the efficiency of the selection decreases as function of the neutrino energy.

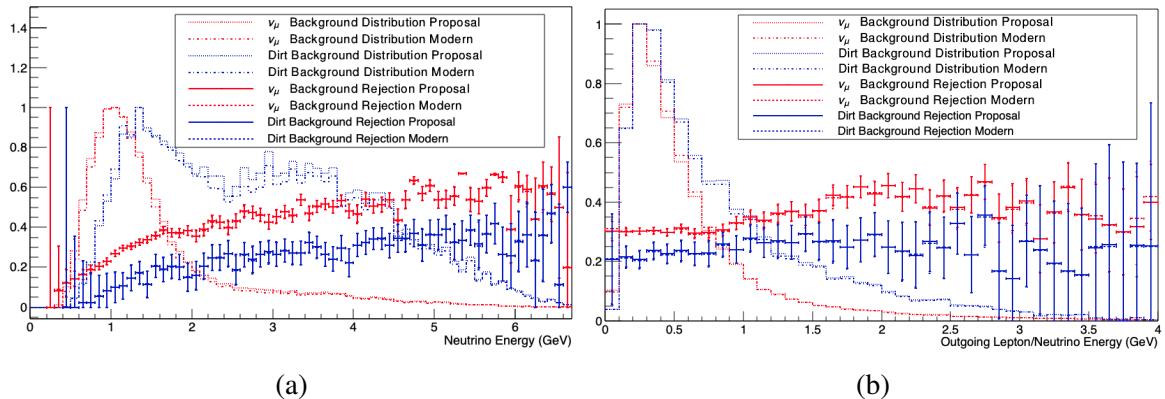


Fig. 7.3 Efficiency of removing events where there are two showers greater than 100 MeV for the  $\nu_\mu$  background events in the active volume (red) and the beam dirt background events (orange). This efficiency is in terms of true incoming neutrino energy (a) and the true outgoing charged lepton/neutrino energy (b). The initial distributions are shown for the proposal (solid line) and the modern (dashed line) samples. The efficiency is presented for both samples and the errors are calculated using the normal approximation to  $1\sigma$ .

If there is only one photon candidate in the event, a conversion gap cut is then applied. Photons can travel several centimetres from the vertex before they interact as depicted in Figure 7.4. If the vertex is deemed visible, due to the presence of at least 50 MeV of hadronic kinetic energy, and the photon starts to shower further than 3 cm from the vertex the event is removed. The kinetic energy of the tracks is smeared by 5% [17].

The efficiency of the conversion gap cut to removing backgrounds can be seen in Figure 7.5 and the total background rejection is shown in Table 7.3. The Figure shows that the background rejection increases as a function of neutrino energy initially. This is due to the hadronic energy increasing. When the hadronic energy becomes visible, increasing the energy does not improve the cut as the vertex is now visible.

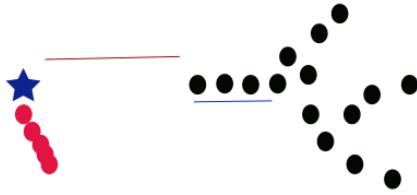


Fig. 7.4 Diagram of a  $\nu$  interaction. The vertex is visible due to hadronic activity (red dots). The black dots refer to a photon shower. Due to the hadronic activity, the conversion gap (red line) is visible. The initial track stub (blue line) is used to calculate the  $dE/dx$  of the shower.

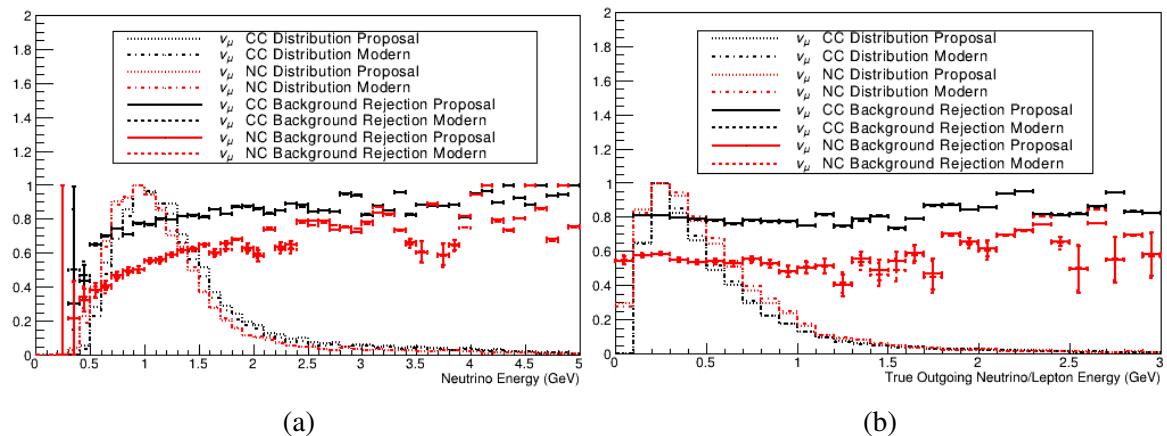


Fig. 7.5 The background rejection from removing events with a visible vertex and conversion gap greater than 3 cm for the  $\nu_\mu$  CC (black) and the  $\nu_\mu$  NC (red) background events in the active volume. This is displayed as a function of the true incoming neutrino energy (a) and the true outgoing charged lepton/neutrino energy (b). The distributions correspond to the events which pass the multiple shower cut in the proposal (solid line) sample and the modern (dashed line) sample. The errors are calculated using the normal approximation to  $1\sigma$ .

The remaining photon background events undergo a  $dE/dx$  cut. When a photon interacts, it usually undergoes pair production where an  $e^+e^-$  pair is produced. The particles travel in the same direction, as depicted in Figure 7.4, and each particle deposits energy that is representative of a Landau distribution that peaks at  $\sim 1.9$  MeV/cm, producing a convolved peak around  $\sim 3.8$  MeV/cm. Therefore, placing a cut on the  $dE/dx$ , for example as in Figure 7.26b, can remove photon showers. Some photons however, undergo Compton scattering where one electron is produced. Such events are indistinguishable from the CC signal.

A weight of 0.06 is applied to the photon background to account for a 94% background rejection using the  $dE/dx$  cut. The weight arises from studies performed in ArgoNeuT [246], where a  $e\gamma$  separation analysis was performed. The study also had an electron selection efficiency of 80%. However, a weight of 0.8 is not applied to the signal resulting in a signal

efficiency of 100%. Furthermore, it should be noted that the analysis was performed in the NuMi beam [96] which has a higher energy than the BNB, therefore pair production is more dominant. A truth-based study of the  $dE/dx$  separational power has since been performed for BNB events to identify the correct limit of the cut. This is discussed in Section 7.4.3.

If a misidentified photon originates from a resonant  $\nu_\mu$  CC interaction, then a further cut is applied which identifies the muon lepton. As muons are minimum ionising, they produce long tracks in the detector. Therefore, events where the muon travels greater than 1 m are assumed to be  $\nu_\mu$  CC interactions and the event is removed. This cut does not effect the signal. Figures 7.6 shows the efficiency of removing the background as a function of neutrino and lepton energy, after the previous cuts. The average background rejection is shown in Table 7.3. The background rejection increases as function of neutrino energy and the outgoing charged lepton energy. This is because higher energy muons travel further.

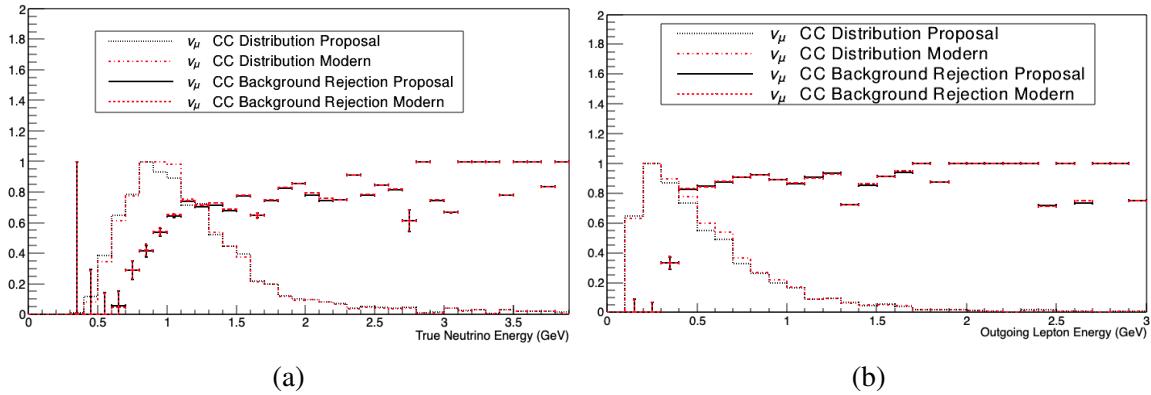


Fig. 7.6 The resultant background rejection due to removing events with a muon with a track length greater than 1 m for the  $\nu_\mu$  CC background events in the active volume (black). This is shown in terms of the true incoming neutrino energy (a) and the true outgoing charged lepton energy (b). The distributions of the events after the conversion gap cut and the two shower cut are shown for the proposal (solid line) and the modern (dashed line) samples. The errors are calculated using the normal approximation to  $1\sigma$ .

Events where the shower has an energy less than 200 MeV are removed from the analysis. Figure 7.7 shows the result of applying this cut on the signal, the NC background and the CC background. Table 7.3 describes the average efficiency and background rejection. As expected, the Figures shows the cut reduces the efficiency at lower lepton energies. As the lepton energy is correlated with the neutrino energy the efficiency decreases as the neutrino energy decreases. The Figure also shows, as expected, the background rejection decreases as a function of the neutrino energy as lower neutrino energies equate to lower energy final state particles. However, the background rejection is independent of the muon/outgoing neutrino energy as this particle is not the source of the photon background.

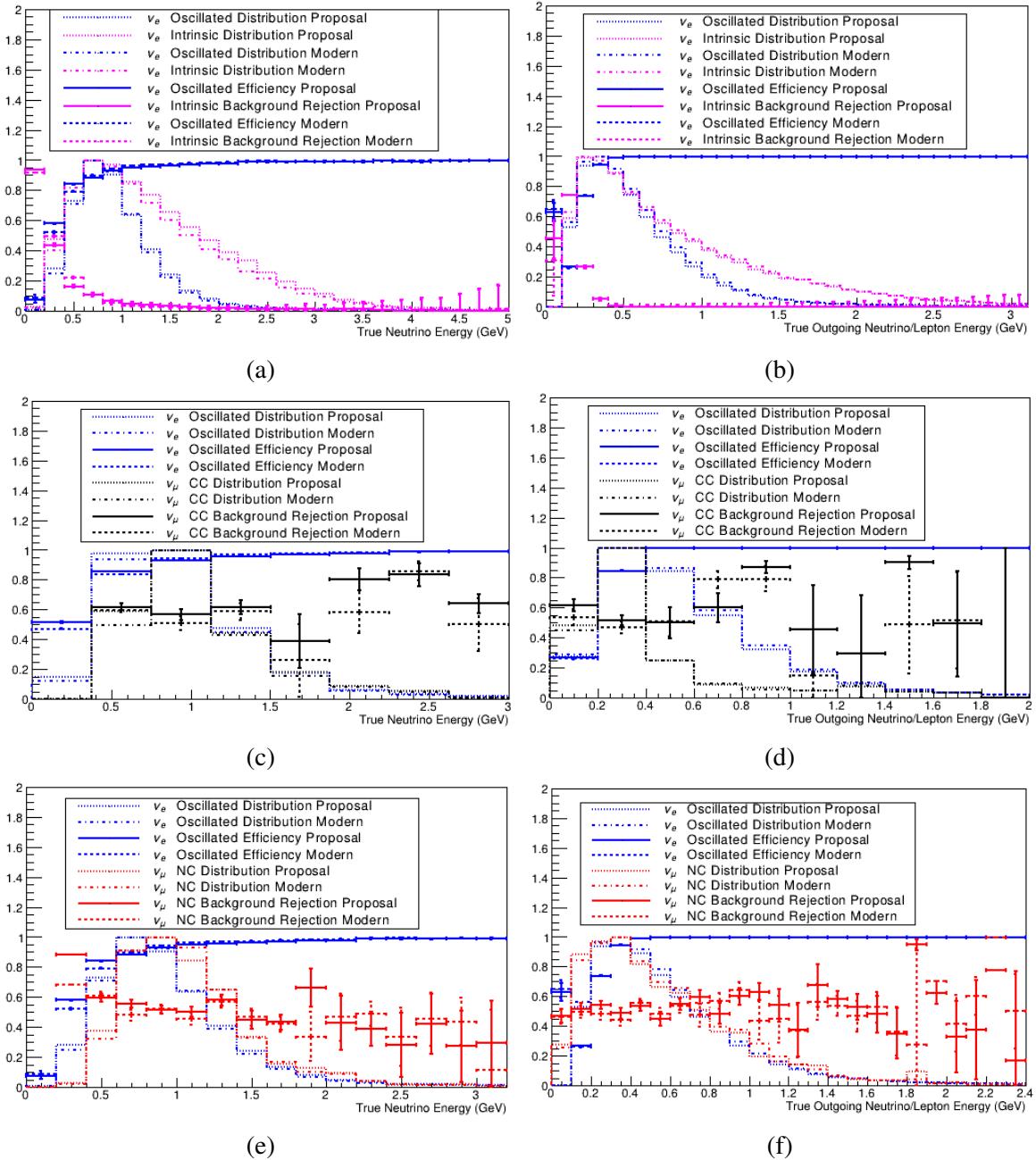


Fig. 7.7 The background rejection due to removing events where the primary shower has less than 200 MeV of energy for the  $\nu_e$  intrinsic CC background events (a,b) (pink),  $\nu_\mu$  CC background events (c,d) (black) and  $\nu_\mu$  NC background events (e,f) (red) in the active volume. The oscillated  $\nu_e$  efficiency (blue) is drawn on each graph for reference. The background rejection is presented in terms of the true neutrino energy (left-hand side) and the true outgoing neutrino/charged lepton energy (right-hand side). The distributions of events after the conversion gap cut, the muon length cut and the two shower cut are shown for the proposal (solid line) and the modern (dashed line) samples. The background rejection is presented for both distributions and the errors are calculated using the normal approximation to  $1\sigma$ .

Finally, a fiducial volume cut is applied. The fiducial volume cut is tabulated in Table 7.2 and is mainly in effect to remove the dirt and cosmic backgrounds and ensure events are contained.

Face	Distance Removed (cm)
Sides	15
Top/Bottom	15
Upstream	30
Downstream	50
Cathode Removal	1.5

Table 7.2 Fiducial volume cut values showing the distance from the active volume faces to the fiducial volume. The cut is applied to all three detectors in the SBN programme. A 1.5 cm cut is applied either side of cathode so that interactions in the cathode are not considered.

The total efficiency and background rejection of the full selection is presented in Figures 7.13 for the main backgrounds of the analysis. The overall efficiency and background rejection is presented in Table 7.3. Note that a weight of 0.8 is applied to all selected events to account for an 80% reconstruction efficiency. This 80% is not accounted for in the Figures or the total efficiencies.

The neutrino energy is reconstructed by summing the energy depositions of the particles in the event within the active volume. This includes particles associated with the vertex and the lepton candidate. The final selection distributions are discussed in Section 7.3.

Cut	Sample									
	v <sub>μ</sub> CC		v <sub>μ</sub> NC		Intrinsic v <sub>e</sub>		Oscillated v <sub>e</sub>		Dirt	
	Prop	Mod	Prop	Mod	Prop	Mod	Prop	Mod	Prop	Mod
No Shower	9.7 <sup>+0.1%</sup> <sub>-0.4%</sub>	7.6 <sup>+0.0%</sup> <sub>-0.0%</sub>	16.0 <sup>+0.1%</sup> <sub>-0.1%</sub>	15.2 <sup>+0.1%</sup> <sub>-0.1%</sub>	91.8 <sup>+0.1%</sup> <sub>-0.1%</sub>	91.3 <sup>+0.1%</sup> <sub>-0.1%</sub>	90.7 <sup>+0.1%</sup> <sub>-0.1%</sub>	90.5 <sup>+0.1%</sup> <sub>-0.1%</sub>	N/A	N/A
2 Shower	67.7 <sup>+0.4%</sup> <sub>-0.4%</sub>	68.1 <sup>+0.4%</sup> <sub>-0.4%</sub>	69.7 <sup>+0.2%</sup> <sub>-0.2%</sub>	69.6 <sup>+0.5%</sup> <sub>-0.5%</sub>	85.4 <sup>+0.1%</sup> <sub>-0.1%</sub>	88.7 <sup>+0.1%</sup> <sub>-0.1%</sub>	91.3 <sup>+0.1%</sup> <sub>-0.1%</sub>	93.1 <sup>+0.0%</sup> <sub>-0.0%</sub>	76.1 <sup>+0.1%</sup> <sub>-0.1%</sub>	76.4 <sup>+0.1%</sup> <sub>-0.1%</sub>
Fiducial Volume	67.6 <sup>+0.3%</sup> <sub>-0.3%</sub>	67.9 <sup>+0.3%</sup> <sub>-0.3%</sub>	68.2 <sup>+0.4%</sup> <sub>-0.4%</sub>	68.2 <sup>+0.4%</sup> <sub>-0.4%</sub>	68.2 <sup>+0.1%</sup> <sub>-0.1%</sub>	68.2 <sup>+0.1%</sup> <sub>-0.1%</sub>	70.0 <sup>+0.1%</sup> <sub>-0.1%</sub>	70.0 <sup>+0.1%</sup> <sub>-0.1%</sub>	43.3 <sup>+0.6%</sup> <sub>-0.6%</sub>	44.4 <sup>+0.6%</sup> <sub>-0.6%</sub>
Conversion Gap	21.0 <sup>+1.5%</sup> <sub>-1.5%</sub>	21.3 <sup>+1.5%</sup> <sub>-1.5%</sub>	45.6 <sup>+1.3%</sup> <sub>-1.3%</sub>	45.5 <sup>+1.3%</sup> <sub>-1.3%</sub>	99.1 <sup>+0.0%</sup> <sub>-0.0%</sub>	99.3 <sup>+0.0%</sup> <sub>-0.0%</sub>	99.1 <sup>+0.0%</sup> <sub>-0.0%</sub>	99.3 <sup>+0.0%</sup> <sub>-0.0%</sub>	N/A	N/A
Muon Track Length	46.9 <sup>+4.7%</sup> <sub>-3.0%</sub>	45.8 <sup>+4.7%</sup> <sub>-3.0%</sub>	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
dE/dx	11.4 <sup>+16.5%</sup> <sub>-11.4%</sub>	11.5 <sup>+16.4%</sup> <sub>-11.5%</sub>	7.69 <sup>+4.5%</sup> <sub>-4.5%</sub>	7.60 <sup>+4.5%</sup> <sub>-4.5%</sub>	99.8 <sup>+0.0%</sup> <sub>-0.0%</sub>	99.9 <sup>+0.0%</sup> <sub>-0.0%</sub>	99.8 <sup>+0.0%</sup> <sub>-0.0%</sub>	99.8 <sup>+0.0%</sup> <sub>-0.0%</sub>	6 <sup>+3.8%</sup> <sub>-3.8%</sub>	6 <sup>+3.7%</sup> <sub>-3.7%</sub>
200 MeV Energy	49.5 <sup>+3.1%</sup> <sub>-3.1%</sub>	50.1 <sup>+3.0%</sup> <sub>-3.0%</sub>	49.6 <sup>+1.5%</sup> <sub>-1.5%</sub>	51.3 <sup>+1.4%</sup> <sub>-1.4%</sub>	91.8 <sup>+0.1%</sup> <sub>-0.1%</sub>	91.6 <sup>+0.1%</sup> <sub>-0.1%</sub>	89.4 <sup>+0.1%</sup> <sub>-0.1%</sub>	89.5 <sup>+0.1%</sup> <sub>-0.1%</sub>	42.4 <sup>+0.8%</sup> <sub>-0.8%</sub>	39.9 <sup>+0.8%</sup> <sub>-0.8%</sub>
Total	0.022 <sup>+0.002%</sup> <sub>-0.002%</sub>	0.017 <sup>+0.001%</sup> <sub>-0.001%</sub>	0.122 <sup>+0.005%</sup> <sub>-0.005%</sub>	0.121 <sup>+0.004%</sup> <sub>-0.004%</sub>	48.1 <sup>+0.1%</sup> <sub>-0.1%</sub>	49.6 <sup>+0.9%</sup> <sub>-0.9%</sub>	50.8 <sup>+0.1%</sup> <sub>-0.1%</sub>	52.0 <sup>+0.1%</sup> <sub>-0.1%</sub>	0.015 <sup>+0.000%</sup> <sub>-0.000%</sub>	0.013 <sup>+0.000%</sup> <sub>-0.000%</sub>

Table 7.3 Selection cuts in the truth-based analysis and the percentage of events which pass the cuts for the main beam background modes and the oscillated v<sub>e</sub> signal in SBND. The proposal sample (Prop) and the modern sample (Mod) efficiencies are presented. The cuts are placed sequentially with the top cut of the Table being applied first. The efficiency of the next cut is then based on the remaining events which passed the previous cut. Events where there is no shower above 100 MeV do not pass the initial cut. Errors are calculated using the normal approximation to 1σ.

### 7.2.2 Dirt Events

Dirt events undergo a similar selection procedure to the Time Projection Chamber (TPC) events and any differences are due to the vertex being outside the detector volume. Therefore, the conversion gap and muon track length cuts are not undertaken. The distribution of dirt event photons within the active volume can be seen in Figure 7.8 which shows that the majority of events are at the face of the detector and near the detector walls. The fiducial volume cut therefore removes  $\sim 26.7\%$  more of the dirt events than the TPC events, removing  $56.7 \pm 0.6\%$  ( $56.6 \pm 0.6\%$ ) of the photons above 100 MeV. The two shower cut, shown in Figure 7.3, removes  $23.9 \pm 0.1\%$  ( $23.6 \pm 0.1\%$ ) of events. This is lower than the TPC background events as the probability that the second shower is in the TPC is smaller. The distribution has the same form as the TPC background. The events that remain after the fiducial volume cut and the two shower cut undergo the 200 MeV shower energy cut. In total, the cut removes an additional  $57.6 \pm 0.8\%$  ( $60.1 \pm 0.8\%$ ) of the events. Including the 94% background removal, due to the  $dE/dx$  cut, the total background rejection is  $99.99 \pm 0.00\%$  ( $99.99 \pm 0.00\%$ ). This total background removal is without the additional 20% removal due to assumed reconstruction inefficiencies. The results here are summarised in Table 7.3.

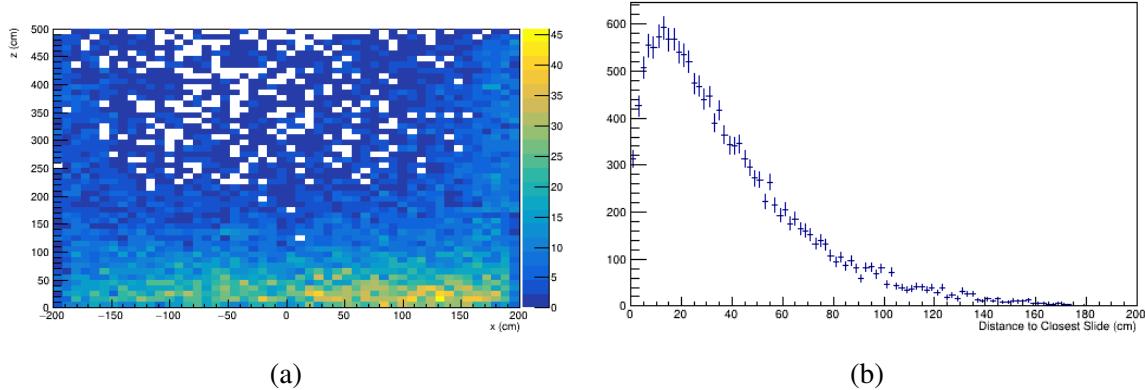


Fig. 7.8 a) The distribution of photons of energy  $>100$  MeV from dirt events within the active volume in terms of the drift direction (x) and the beam direction (z). b) The distance to the closest wall of the active volume bounding box for photons of energy  $>100$  MeV that enter the active volume from dirt events.

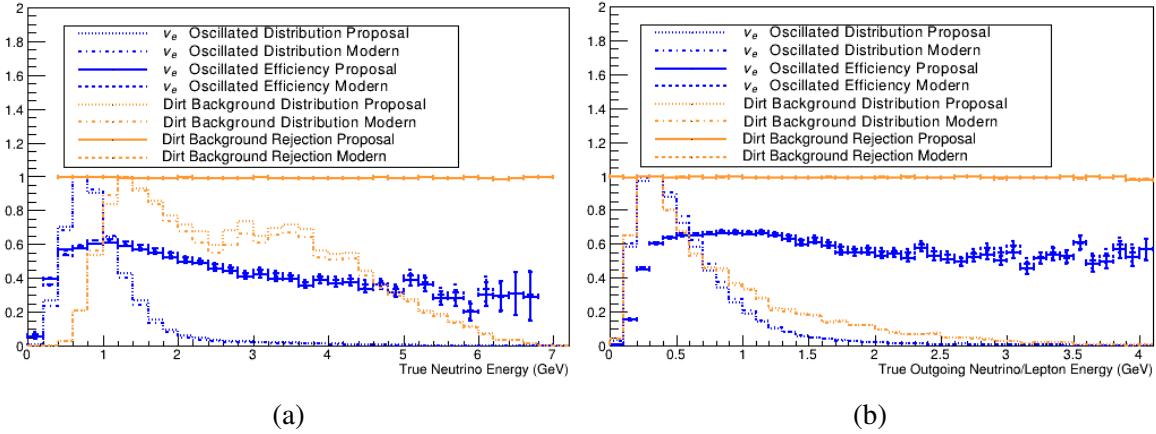


Fig. 7.9 The total background rejection in the truth-based analysis for removing dirt events with respect to the neutrino energy (a) and outgoing charged lepton/neutrino energy (b). The distributions of events after the two shower cut are shown for the proposal (solid line) and the modern (dashed line) samples. The oscillated  $v_e$  efficiency (blue) is drawn for reference. The background rejection is presented for both distributions and the errors are calculated using the normal approximation to  $1\sigma$ .

### 7.2.3 Cosmic Removal

For the cosmic events, a separate analysis is applied. Firstly, if a cosmic photon initially interacts outside the fiducial volume the event is removed. Figure 7.10 shows the photon decay point for cosmic showers above 100 MeV in the active volume. As can be seen in the Figure, the majority of the cosmic photons decay near the top of the detector. Thus,  $37.4^{+1.1}_{-1.1}\%$  of cosmic events are removed due to the fiducial volume cut. Cosmic events which occur outside the beam spill time window are also removed if there is no other activity in the TPC during the beam spill time. Doing so removes  $99.6^{+0.02}_{-0.02}\%$  of cosmic events; however, this assumes the PDS or CRT systems are able to time tag all cosmic events correctly. An additional weight of 0.05 is also applied to the cosmic events. This is to account for further reductions of cosmics within the beam spill time, which are achieved by using the PDS and CRT systems and taking advantage of the bucket structure of the beam. The beam properties are described in Chapter 4, Section 4.1.

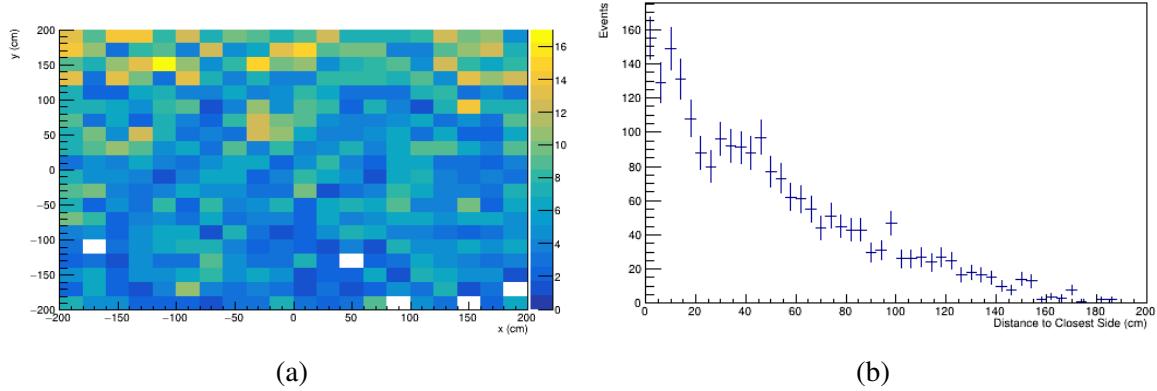


Fig. 7.10 a) The distribution of photons, with energy greater than 100 MeV, from cosmic events within the active volume in terms of the vertical direction ( $y$ ) and the beam direction ( $z$ ). b) The distance to the closest wall of the active volume bounding box for photons that enter the active volume from cosmic events.

Events are then weighted by 0.06 to mimic the 94% background removal from the  $dE/dx$  cut. Events below the reconstructed energy of 200 MeV are also removed. This results in the efficiency curve seen in Figure 7.11 and results in a  $99.4^{+0.02}_{-0.02}\%$  reduction in cosmic events.

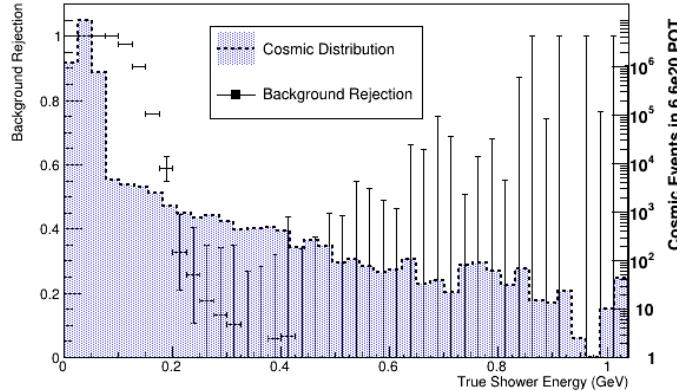


Fig. 7.11 The background rejection when removing cosmic events with an energy greater than 200 MeV. The errors are calculated using the normal approximation to  $1\sigma$ .

Finally, a topological cosmic cylinder cut is applied. Cosmic photons, which originate from cosmic muons that pass through the TPC, can be removed by a fiducial volume cut corresponding to a cylinder of radius 15 cm around the cosmic muon. This process is schematically shown in Figure 7.12 [17]. The average background rejection is  $16.0^{+1.5}_{-1.4}\%$ .

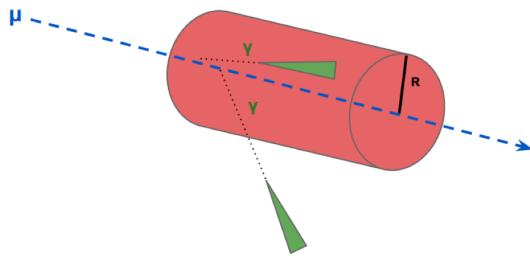


Fig. 7.12 A diagram of the cylinder cut from Reference [17].

The total background rejection of all the cuts is shown in Table 7.4.

Cut	Background Rejection
Fiducial Volume	$37.4^{+1.1\%}_{-1.1\%}$
Spill Trigger	$99.6^{+0.02\%}_{-0.02\%}$
$dE/dx$	94.0%
CRT exploitation	95.0%
200 MeV Energy Cut	$99.4^{+0.02\%}_{-0.02\%}$
Cosmic Cylinder Cut	$16.0^{+1.5\%}_{-1.4\%}$
Total	$99.9^{+0.0\%}_{-0.0\%}$

Table 7.4 The cosmic background rejection for each of the cosmic cuts in SBND. The errors on the figures are the  $1\sigma$  Clopper-Pearson intervals [32].

### 7.3 Comparisons Between the Truth-Based Analyses

This Section will compare the results for the "proposal" sample and the "modern" sample from the application of the selection discussed in the previous Sections.

The total oscillated  $\nu_e$  efficiency and intrinsic  $\nu_e$  background rejection in terms of the true incoming neutrino energy and the true outgoing charged lepton energy are shown in Figure 7.13. Due to the 200 MeV primary shower cut, the efficiency decreases at low energy. Also, at higher energies the  $\nu_e$  CC resonant interactions occur more frequently and are removed due to the two shower cut. This reduces the efficiency. The overall efficiency and background rejection are presented in Table 7.3.

A comparison of the "modern" sample to the "proposal" is shown as a function of the "reconstructed" neutrino energy in Figures 7.14, 7.15 and 7.16. The Figures show the resultant spectra from each detector (a weighting of  $\sin(\theta_{\mu e}) = 0.013$  and  $\Delta m_{41}^2 = 0.43 \text{ eV}^2$  using the 3+1 sterile model has been used on the oscillation sample [31]). The ratio plot

shows that the main difference is an overall increase in intrinsic and oscillated  $\nu_e$  CC events of  $\sim 20\%$ . This is due to the inclusion of MEC events. This inclusion results in a increase of 1.2% in selection efficiency in the "modern" sample compared to the "proposal" sample.

Other differences are present in the "modern" and "proposal" samples for the other backgrounds due to the GENIE weight. The main impact is a reduction of NC events from the "proposal" to the "modern" sample at low energies due the reduction in the resonant and DIS events weighting.

Comparing the actual proposal analysis, where the spectra are shown in Figures 7.20, to the new proposals shows an increase of NC events in the current analysis. This is concentrated in the lower energy range which can be seen in Figures 7.17, 7.18 and 7.19. One cause of the difference is the addition of rare events, such as electron scattering and Dalitz decays. These were not considered in proposal spectra but were considered in the sensitivity analysis. There is also a shift in the  $\nu_\mu$  CC event rate as to higher reconstructed neutrino energies. This is visible in the ratio plots in Figures 7.17, 7.18 and 7.19. This shift is due to the inclusion of the energy deposited by the muon in the reconstructed neutrino energy.

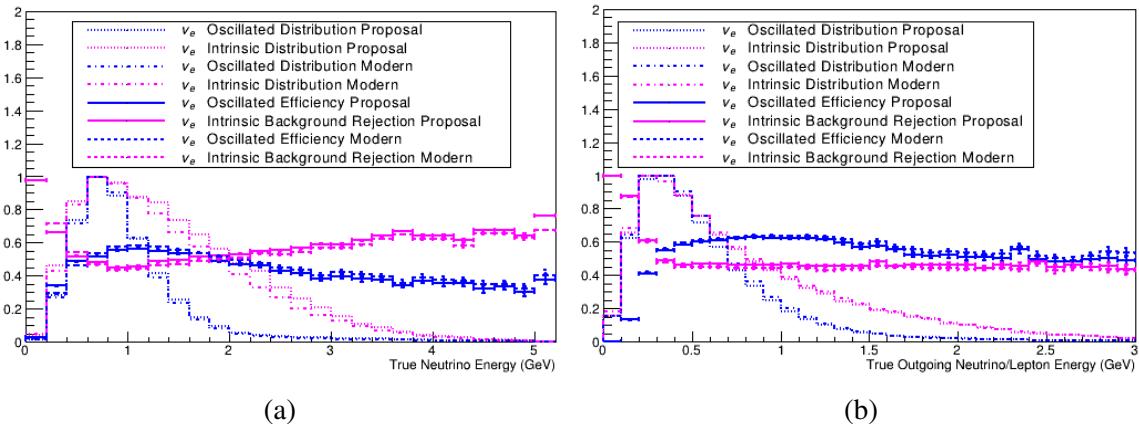


Fig. 7.13 The total background rejection of the analysis for the intrinsic  $\nu_e$  CC (a,b) (pink). The efficiencies are presented in terms of the true neutrino energy (left-hand side) and the true outgoing neutrino/charged lepton energy (right-hand side). The oscillated  $\nu_e$  efficiency (blue) is drawn on each graph for reference. The background rejection is presented for both the proposal and modern distributions and the errors are calculated using the normal approximation to  $1\sigma$ .

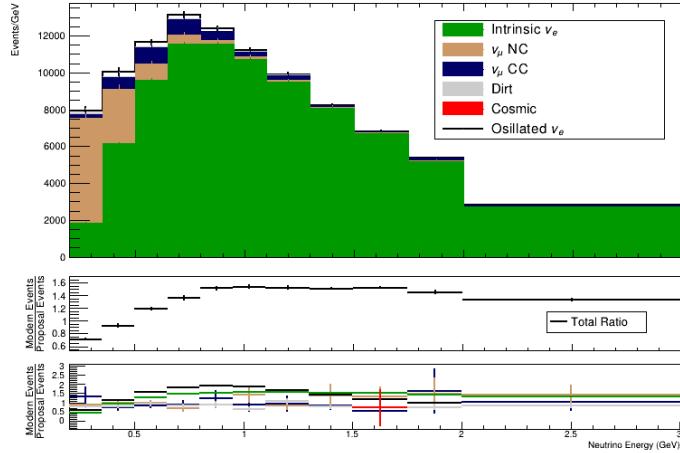


Fig. 7.14 The resultant spectra from the truth-based analysis in SBND using the "modern" sample". The ratio between the "modern" sample and "proposal" sample for the total event distribution and the different interaction types of the analysis is shown. Errors are derived from the Poissonian errors of the histogram. A weighting of  $\sin(\theta_{\mu e}) = 0.013$  and  $\Delta m_{41}^2 = 0.43 \text{ eV}^2$  has been used on the oscillation sample [31].

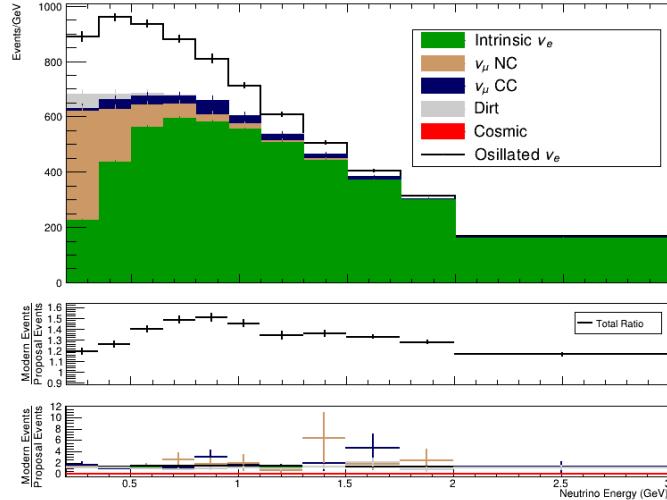


Fig. 7.15 The resultant spectra from the truth-based analysis in MicroBooNE using the "modern" sample. The ratio between the "modern" sample and "proposal" sample for the total event distribution and the different interaction types of the analysis is shown. Errors are derived from the Poissonian errors of the histogram. A weighting of  $\sin(\theta_{\mu e}) = 0.013$  and  $\Delta m_{41}^2 = 0.43 \text{ eV}^2$  has been used on the oscillation sample [31].

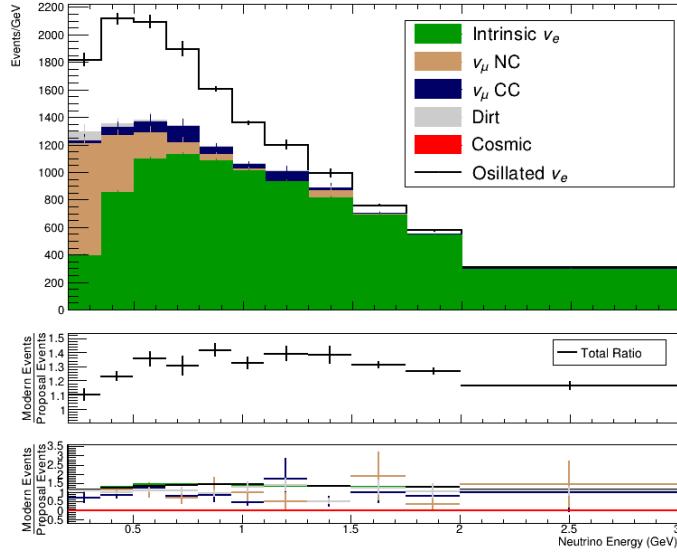


Fig. 7.16 The resultant spectra from the truth-based analysis in ICARUS using the "modern" sample. The ratio between the "modern" sample and "proposal" sample for the total event distribution and the different interaction types of the analysis is shown. Errors are derived from the Poissonian errors of the histogram. A weighting of  $\sin(\theta_{\mu e}) = 0.013$  and  $\Delta m_{41}^2 = 0.43 \text{ eV}^2$  has been used on the oscillation sample [31].

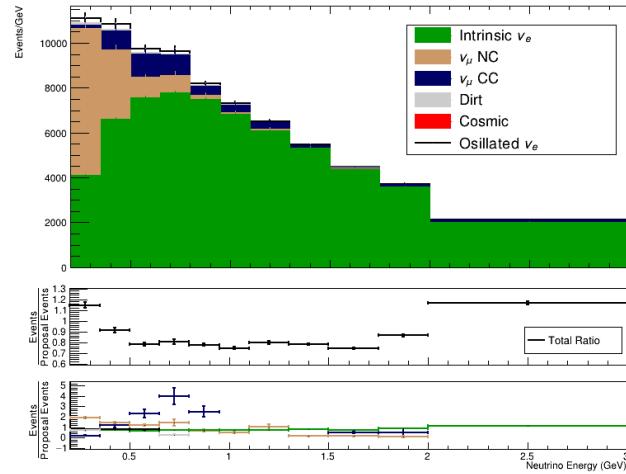


Fig. 7.17 The resultant spectra from the truth-based analysis in SBND using the "proposal" sample. The ratio between the total event distribution and the different interaction types of the analysis against the spectra from 2012 is shown. Errors are derived from the Poissonian errors of the histogram. A weighting of  $\sin(\theta_{\mu e}) = 0.013$  and  $\Delta m_{41}^2 = 0.43 \text{ eV}^2$  has been used on the oscillation sample [31].

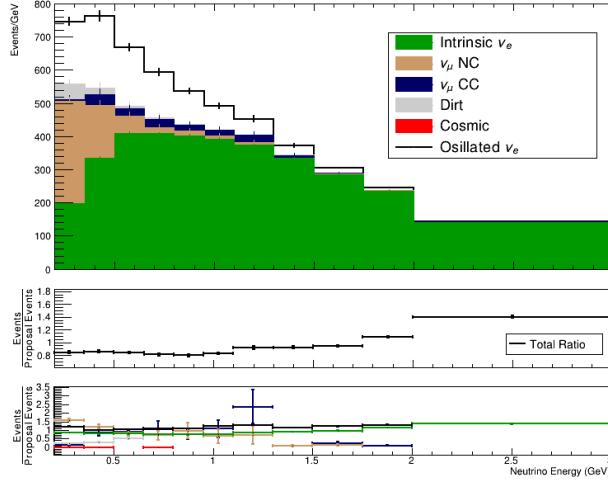


Fig. 7.18 The resultant spectra from the truth-based analysis in MicroBoone using the "proposal sample". The ratio between the total event distribution and the different interaction types of the analysis against the spectra from 2012 is shown. Errors are derived from the Poissonian errors of the histogram. A weighting of  $\sin(\theta_{\mu e}) = 0.013$  and  $\Delta m_{41}^2 = 0.43 \text{ eV}^2$  has been used on the oscillation sample [31].

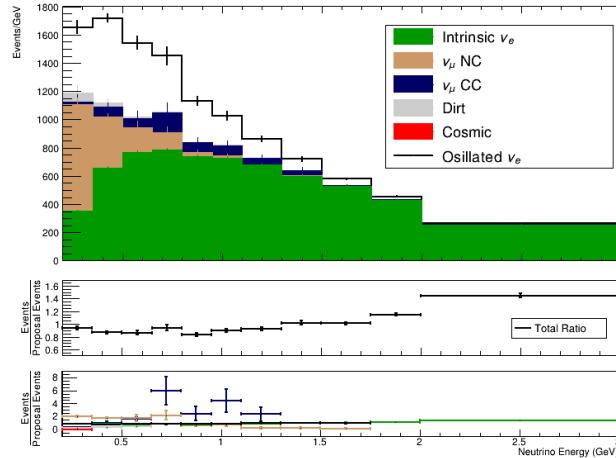


Fig. 7.19 The resultant spectra from the truth-based analysis in ICARUS using the "proposal" sample. The ratio between the total event distribution and the different interaction types of the analysis against the spectra from 2012 is shown. Errors are derived from the Poissonian errors of the histogram. A weighting of  $\sin(\theta_{\mu e}) = 0.013$  and  $\Delta m_{41}^2 = 0.43 \text{ eV}^2$  has been used on the oscillation sample [31].

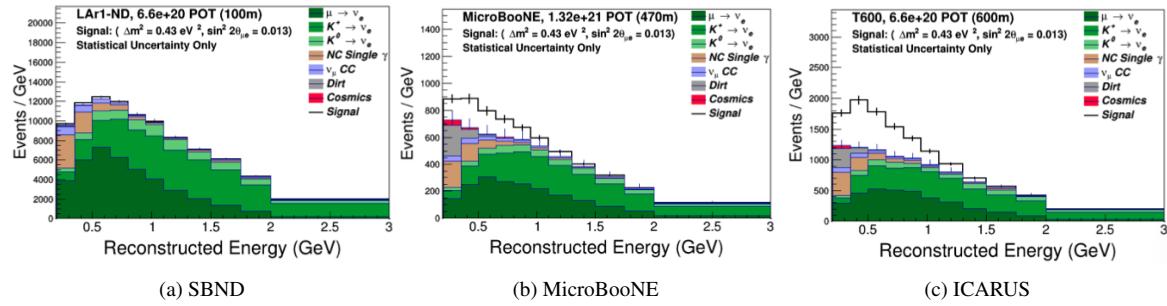


Fig. 7.20 SBN electron neutrino spectra after the  $\nu_e$  CC inclusive truth-based selection.

In the previous Section, a truth-based selection procedure was presented using the "proposal" and "modern" samples. Due to the selection being of the same form as the SBN 2012 proposal selection, the resultant spectra for the "proposal" can be compared to the spectra from the 2012 analysis and this has been undertaken in this Section. Figures 7.17, 7.18 and 7.19 show this comparison and show that differences not accounted by the application of the GENIE weighting described above. These unresolved differences in the spectra are due to the definitions of the backgrounds. For example neutrino-electron scattering not included in Figures 7.20 whilst they are included in Figures 7.17, 7.18 and 7.19. Unfortunately the MC from the proposal is no longer available and so the rare background rates cannot be verified. The result of the proposal spectra is used in the analysis in Chapter 8 where it is compared with the actual proposal sensitivity.

In addition to the proposal comparison, the "proposal" and "modern samples" have been compared in this Section. This was performed so that changes in selection samples could be identified due to the changes in physics and detector models. The main difference between the spectra from the "modern" sample and the "proposal" sample is the inclusion of MEC events which increases the signal size by approximately 20%. This has implications for the analysis and this is discussed in Chapter 8.

## 7.4 Evaluation of Reconstructed Selections in SBND

The truth-based analysis performed in Section 7.2.1 provides an estimate of the performances expected in the SBN programme. The next step is to perform selection procedures upon reconstructed events to provide more realistic simulated results. This process is described in this Section.

The events used to evaluate the selection are discussed in Section 7.4.1. The events underwent the reconstruction procedure in Chapter 6. Using these simulated events, two pre-selection cuts were performed to reduce the samples to manageable sizes. These cuts are

discussed in Section 7.4.2. Further cuts were then performed which mimicked the proposal and are discussed in Section 7.4.3, while the new metrics which are used to identify  $\nu_e$  CC events from the  $\nu_\mu$  background are discussed in Section 7.4.4. Finally, machine learning is applied to perform the selection and this is discussed in Section 7.4.5.

Note that the efficiencies, background rejection and purity of cuts are presented within the text for the oscillation parameters discussed in Section 7.4.1. The efficiencies and background rejection for a 1:1 normalisation are given in the brackets. Both samples have no cosmic overlay and so the reconstruction is not affected by the presence of cosmics. This must be addressed in future analyses.

### 7.4.1 Reconstruction Samples Used

To perform the selection on reconstructed events, two samples were created in SBND to mimic the oscillated  $\nu_e$  signal and the TPC  $\nu_\mu$  backgrounds. The additional samples were used to tune the selection independently from the samples described above. The final selections procedures below were then used on the larger samples described in Section 7.1.1 to perform the oscillation analysis. The signal events were made up of 43 K simulated  $\nu_e$  events within the active volume. These events have an energy spectrum corresponding to the energy of the BNB. The TPC background events correspond to 255 K  $\nu_\mu$  simulated events within the active volume. The number of events in the samples corresponded to  $3.37 \times 10^{18}$  POT for the oscillated sample and  $1.62 \times 10^{19}$  POT for the background  $\nu_\mu$  sample. These values account for 0.51% and 2.45% respectively of the total events within the SBN three year run.

The current best fit for  $\nu_e$  appearance places  $\sin^2(\theta_{\mu e}) \sim \mathcal{O}(10^{-3})$  in the 3+1 sterile model, see Chapter 2, Section 2.2.5. Oscillation weights were therefore chosen to normalise the signal and background. The parameters are chosen in order to mimic an expected signal in the far detectors for the current global best fit.  $\Delta m_{41}^2$  was set to 11.3 eV<sup>2</sup> resulting in a maximal oscillation probability at the SBND detector. The value of  $\sin^2(\theta_{\mu e})$  was then set to 0.003 which corresponds to the LSND best fit value [9]. In this case, every simulated event corresponds to  $\sim 10^3$  total events and hence the oscillated data corresponds to a data-set that is 510% larger than expected for the three year run.

The events were reconstructed in the manner described in Chapter 6. However, a Support Vector Machine (SVM) was implemented to improve the track-shower characterisation, the result of which is discussed in Appendix D.

The resultant reconstructed neutrino energy spectrum is discussed in Appendix F.

### 7.4.2 Pre-selection Cuts

The first procedure of the selection was to reduce the samples to a more manageable size with two selection cuts: removing events with no reconstructed PFP neutrino in the active volume and removing events with no reconstructed shower in the active volume. The remaining events are then used to evaluate further selection cuts used in the analysis. This includes the proposal cuts, discussed in Section 7.4.3, as well as the new cuts defined in Section 7.4.4 and the machine learning method selections in Section 7.4.5. Both cuts are discussed in more detail in Appendix E.

The first cut removed events where a reconstructed PFP neutrino did not exist in the drift window. This predominantly removed events where only elastic NC interactions occurred in the active volume during the 3 ms drift time. The events remaining events are Charge Current (CC) interactions and NC interaction with final state charged particles. The total CC signal efficiency is  $98.8^{+0.1}_{-0.1}$  (stat)% and the  $\nu_\mu$  background rejection is  $13.9^{+0.1}_{-0.1}$  (stat)% and is shown as function of the neutrino energy in Figure 7.21.

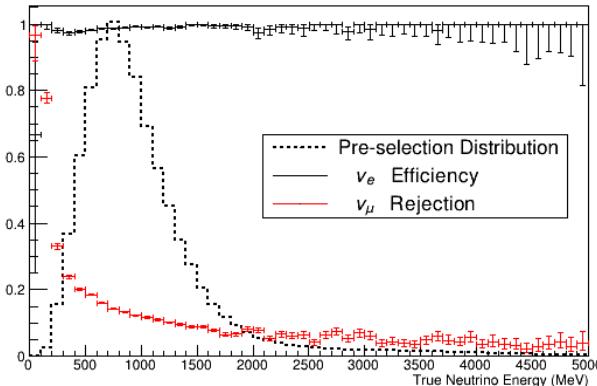


Fig. 7.21 The selection efficiency of CC  $\nu_e$  interactions and the background rejection of  $\nu_\mu$  interactions in the active volume for removing events where a PFP reconstructed neutrino does not exist in the event drift window. The underlying neutrino distribution is also plotted. The errors on the figures are the  $1\sigma$  Clopper-Pearson intervals [32].

The second pre-selection cut requires that the event contains at least one reconstructed shower in the drift window. This is a straight forward requirement of a  $\nu_e$  CC selection as the identifiable part of the interaction is the CC showering electron. Also performing the cut significantly removes Quasi-Elastic (QE)  $\nu_\mu$  interactions which are the most common beam event in the detector. Mostly, Resonant, Deep Inelastic Scattering (DIS) and other  $\nu_\mu$  interactions which produce a photon, either directly or from a neutral pion decay, remain as the main  $\nu_\mu$  backgrounds. Figure 7.22 shows the final  $\nu_e$  CC signal efficiency and the  $\nu_\mu$

background rejection. The resulting efficiency is  $83.5^{+0.2}_{-0.2}(\text{stat})\%$  and background rejection is  $79.2^{+0.2}_{-0.2}(\text{stat})\%$ .

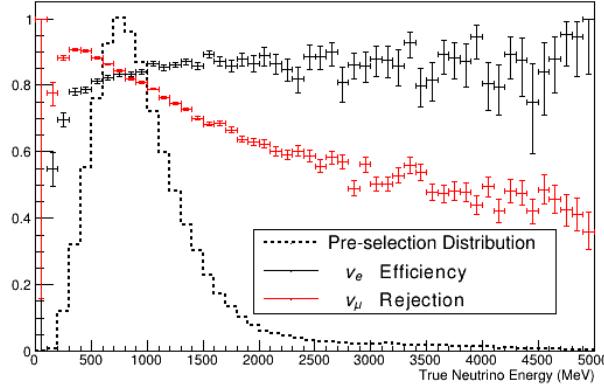


Fig. 7.22 The selection efficiency of CC  $\nu_e$  interactions and the background rejection of  $\nu_\mu$  interactions in the active volume for removing events where a reconstructed shower does not exist in the event drift window. The underlying neutrino distribution is also plotted. The errors on the figures are the  $1\sigma$  Clopper-Pearson intervals [32]. The results correspond to the sample created after performing the PFP cut.

After performing the two cuts, the oscillated sample is reduced to  $\sim 33$  K events and the  $\nu_\mu$  sample to  $\sim 45$  K. These reduced samples are then used to evaluate separate cuts defined in the upcoming Sections.

### 7.4.3 Evaluating the Proposal Selection Cuts on Reconstructed Events

After the samples are reduced by the two previous cuts, the proposal selection can be performed. In the following Figures, the signal is defined as CC  $\nu_e$  events and the background is defined as the  $\nu_\mu$  events.

#### Number of Showers Cut

Firstly, events are accepted if they have only one shower in the event. Due to the segmentation of showers discussed in Chapter 7, Section 6.1.5, one true shower can be mistaken for several showers, reducing the selection efficiency. Requiring one shower in the active volume equates to  $71.8 \pm 0.9\%$  efficiency and  $41.0 \pm 0.4\%$  background rejection. It should also be noted that multiple events can occur within the event drift window and therefore events exist with no reconstructed shower (2.9% signal and 6.5% background).

To remove small segments of large showers, a reconstructed energy cut was applied such that showers below the cut are ignored in the analysis. Such a cut also mimics the 100

MeV visibility threshold in the truth analysis. The results of selecting one shower with an energy cut are shown in Figure 7.23a. The best cut to maximise the efficiency  $\times$  background rejection is a value of 190 MeV, with an efficiency of  $64.2 \pm 0.4\%$  and a background rejection of  $78.6 \pm 0.2\%$ . The increase in background rejection is due to the topology of the energy distribution of the photons, which is shown in Figure 7.28. Most of the background is rejected because the reconstructed energy of the photon is below the threshold. This is proven in Figure 7.23b, where it is required that events are only removed when there are two showers above the energy threshold. This distribution shows that requiring no energy threshold results in the best efficiency ( $74.4 \pm 0.9\%$ )  $\times$  background rejection ( $34.4 \pm 0.3\%$ ).

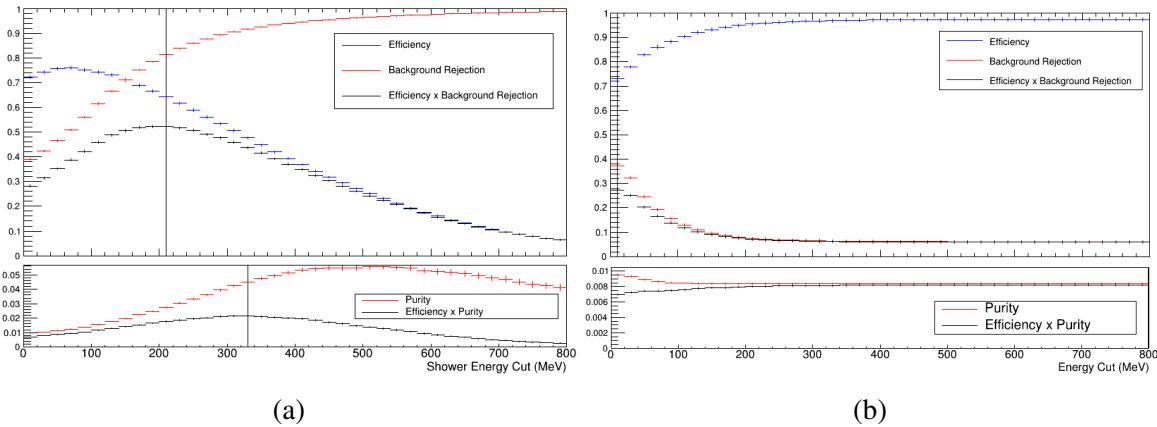


Fig. 7.23 The efficiency (blue), background rejection (red), efficiency  $\times$  background rejection (black), purity (red) and efficiency  $\times$  purity (black) for selecting events requiring a) one shower above the energy threshold and b) no more than one shower above threshold. The error on the curves are the cumulative statistical error.

### Conversion Gap Cut

The conversion gap is defined as the distance from the reconstructed start position of the most energetic shower and the reconstructed neutrino vertex position. Approximately 70% of showers are reconstructed within 5 cm of the true start position of the shower, see Chapter 6, Section 6.2.2. As demonstrated in Figure 7.24, defining the vertex as properly reconstructed if it lies within 5 cm of the true start position results in an 80% efficiency at  $\sim 200$  MeV hadronic energy. This compares to a 100% reconstruction at 50 MeV suggested by the proposal.

As the majority of events are reconstructed correctly, the conversion gap cut is one of the most successful cuts. As can be seen in Figure 7.25, the best cut to maximise the efficiency  $\times$  background rejection is 2.1 cm with an efficiency of  $71.2 \pm 1.9$  (stat) % and a background rejection of  $61.1 \pm 0.7$  (stat) %. See Table 7.6 for the full breakdown.

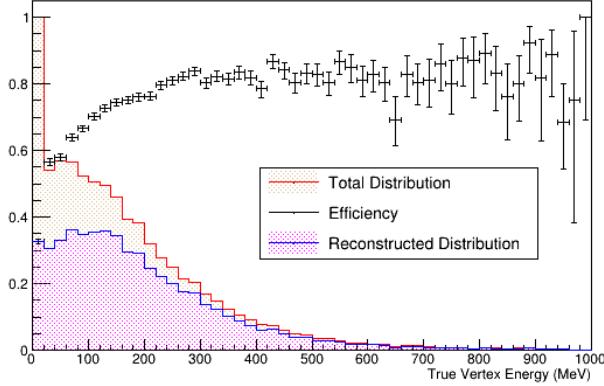


Fig. 7.24 The efficiency for reconstructing the neutrino vertex within 5 cm of the true vertex as a function of the true hadronic energy deposited in the TPC. The analysis is performed on the  $\nu_\mu$  events after the PFP neutrino and no shower cuts are applied. The underlying true vertex distribution is also plotted, along with the correctly reconstructed distribution. The errors on the figures are the  $1\sigma$  Clopper-Pearson intervals [32].

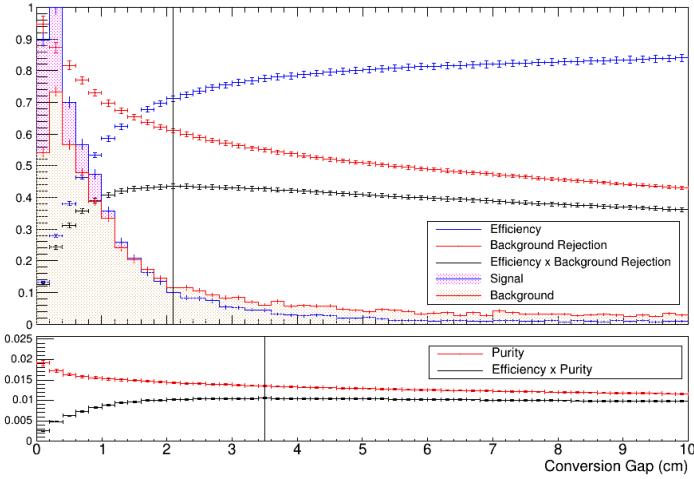


Fig. 7.25 The conversion gap distance for signal (blue hatched) and background (red hatched) events. The efficiency (blue), background rejection (red), efficiency  $\times$  background rejection (black), purity (red) and efficiency  $\times$  purity curves are also plotted. The error on the curves are the cumulative statistical error.

### Shower Track Stub $dE/dx$ Cut

The  $dE/dx$  of the most energetic shower is shown in Figure 7.26a. The performance for the neutrino samples is worse than the vertex samples, discussed in Chapter 6, Section 6.2.2, due to the changes in the photon shower energy distribution. This distribution is shown in Figure

7.28. In the vertex case, the photons are simulated with the same energy distribution as the electrons. Therefore, the average energy of the photon showers is higher, which causes pair production to be more dominant. This means more photons can be removed using  $dE/dx$ . Misidentification and inefficiencies in the reconstruction due to the more complex BNB topologies are also at fault.

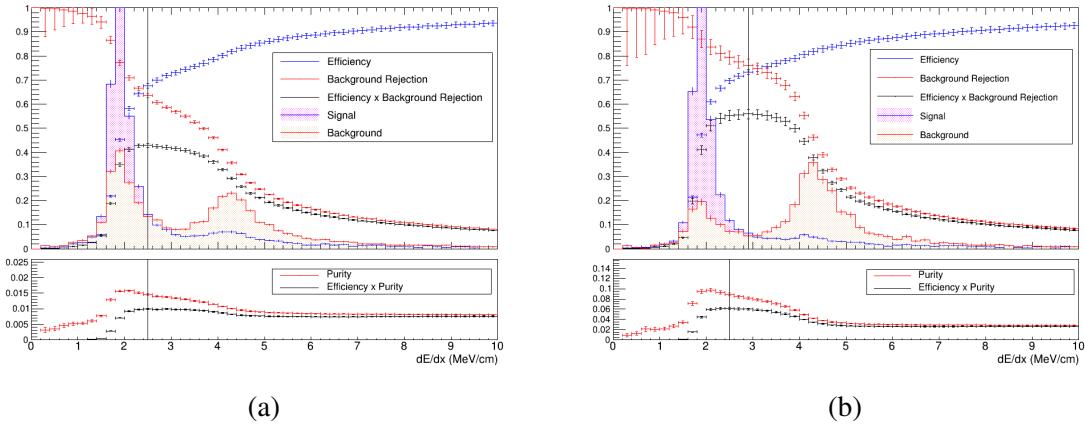


Fig. 7.26  $dE/dx$  for signal (blue hatched) and background (red hatched) for all events (a) and for events where the reconstructed shower energy is greater than 225 MeV (b). The efficiency (blue), background rejection (red), efficiency  $\times$  background rejection (black), purity (red) and efficiency  $\times$  purity (black) curves are also plotted. The error on the curves are the cumulative statistical error.

By using Figure 7.26a, it can be seen that the current best cut on the reconstructed  $dE/dx$  of the most energetic shower is at 2.5 MeV/cm with  $67.6 \pm 0.9$  (stat) % efficiency and a background rejection of  $63.5 \pm 0.8$  (stat) %. See Table 7.6 for the full breakdown. However, as shown in Figure 7.26b, by applying a minimum 225 MeV reconstructed energy cut to remove poorly reconstructed showers the  $dE/dx$  separation significantly improves. The resultant best cut is at 2.9 MeV/cm with an efficiency of  $73.4 \pm 1.1$  (stat) % and a background rejection of  $76.2 \pm 2.5$  (stat) %. It should be noted that the limit of these values are 95.4% for the efficiency and 83.3% for the background rejection based on a truth based analysis discussed in Appendix G.

### Longest Track Length Cut

The final topological cut in the proposal selection is the maximum track length cut. As can be seen in Figure 7.27, there is separation in the signal and background when considering the

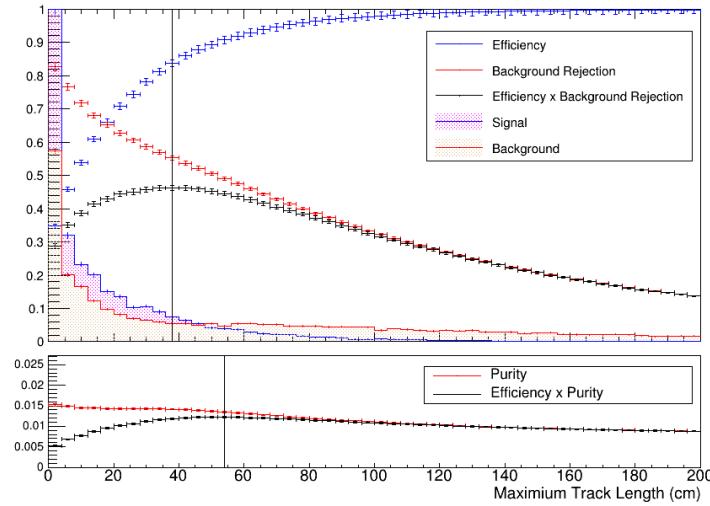


Fig. 7.27 Length of the longest track for signal (blue hatched) and background (red hatched) events. The efficiency (blue), background rejection (red), efficiency  $\times$  background rejection (black), purity (red) and efficiency  $\times$  purity (black) curves are also plotted. The error of the curves is the cumulative statistical error.

length of the longest track in the event. This is primarily due to CC muon producing long tracks in the detector. Resonant events smear the separation due to the creation of charged pions which travel a similar length to the muon. Additional smearing is due to the presence of protons, which will be discussed in Section 7.4.4. The resultant best cut-off to maximise the efficiency  $\times$  background rejection is a maximum track length of 38 cm with an efficiency of  $83.8 \pm 1.0$  (stat)% and a background rejection of  $55.3 \pm 0.6$  (stat)%.

### Shower Energy Cut

In the proposal selection, events where the shower is below 200 MeV are removed. In reconstruction the optimal efficiency  $\times$  background rejection is at 225 MeV with a  $75.1 \pm 1.0$  (stat)% efficiency and a  $77.1 \pm 1.1$  (stat)% background rejection, which is shown in Figure 7.28. As discussed previously in the Section, performing this cut increases the robustness of the reconstruction and therefore increases the separation power of other metrics. Figure E.4b demonstrates that  $\sim$ 225 MeV is also the point where the efficiency on removing events with no showers in the active volume is not affected, indicating that the reconstruction efficiency is limited below this value.

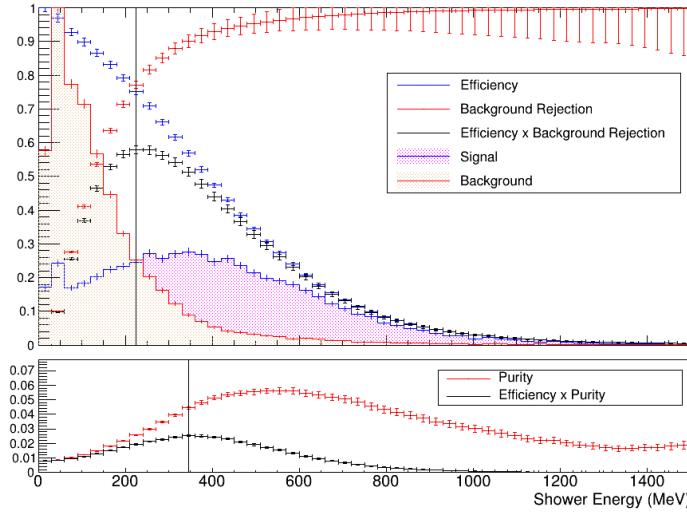


Fig. 7.28 The energy of the largest shower for signal (blue hatched) and background (red hatched) events. The efficiency (blue), background rejection (red), efficiency  $\times$  background rejection (black), purity (red) and efficiency  $\times$  purity (black) curves are also plotted. The errors on the curves are the cumulative statistical error.

### The Full Proposal Selection

Using the values which maximise the efficiency  $\times$  background rejection for each cut, seen in the Table 7.6, the total selection efficiency is  $21.44 \pm 0.30\%$  ( $19.84 \pm 0.26\%$ ) and the total background rejection is  $99.56 \pm 0.04\%$  ( $99.56 \pm 0.00\%$ ). This results in a purity of  $28.47 \pm 1.68\%$ , an efficiency  $\times$  background rejection of  $0.213 \pm 0.003$  ( $0.198 \pm 0.003$ ) and an efficiency  $\times$  purity of  $0.061 \pm 0.004$ . Choosing the parameters which maximise the efficiency  $\times$  purity in Table 7.6 results in an efficiency of  $19.36 \pm 0.29\%$  ( $17.27 \pm 0.25\%$ ), a background rejection of  $99.84 \pm 0.00\%$  ( $99.84 \pm 0.00\%$ ) and a purity of  $49.13 \pm 3.37\%$ . This results in an efficiency  $\times$  background rejection of  $0.193 \pm 0.003$  ( $0.172 \pm 0.003$ ) and an efficiency  $\times$  purity of  $0.095 \pm 0.006$ .

The efficiency is depicted as a function of the true neutrino energy in Figure 7.29a. Other interaction variables are discussed in Appendix H. These Figure show the efficiency curve when maximising the efficiency  $\times$  background rejection. It can be seen that the efficiency increases as a function of the outgoing charged lepton/neutrino energy. This is firstly due to the lepton becoming more visible in the CC signal events. Secondly, more of the energy is taken out of the event by the neutrino in the background NC cases.

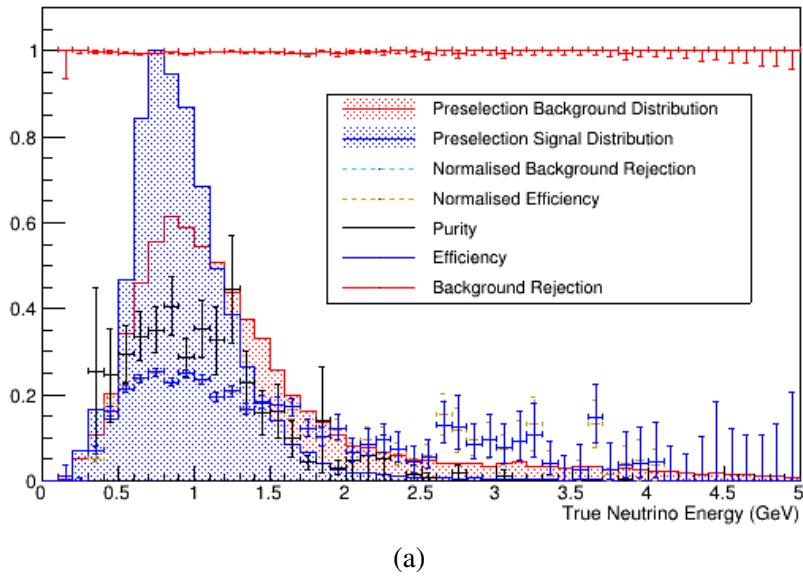


Fig. 7.29 The efficiency of selecting CC  $\nu_e$  events (blue), the background rejection for removing  $\nu_\mu$  events (red) and the purity (black) of the proposal-like selection as a function of the true neutrino energy. The underlying distributions before selection for the signal (red) and background (blue) are also plotted, along with the efficiency (dashed orange) and background rejection (dashed light blue) for the POT and oscillation normalised events. The efficiencies are calculated from the remaining events after the pre-selection cuts and the full selection. For the unnormalised events the errors are the  $1\sigma$  Clopper-Pearson intervals [32], whilst for the normalised distributions the normal approximation is used.

In order to improve the selection efficiency, additional metrics were used to separate the  $\nu_e$  CC signal from the  $\nu_\mu$  background. The next Section will discuss the metrics and performance changes.

#### 7.4.4 Including Additional Metrics in the $\nu_e$ Selection

This Section describes additional metrics to separate the  $\nu_e$  CC signal from the  $\nu_\mu$  background. The metrics can be split into two groups:  $\nu_e/\nu_\mu$  separation metrics and  $e/\gamma$  separation metrics. The  $\nu_e/\nu_\mu$  metrics can be used for a traditional cut-based selection and will be described first. Then the  $e/\gamma$  metrics, which have a smaller separational power, will be discussed. The  $e/\gamma$  metrics cannot be used effectively in a traditional cut-based selection and are instead used in a Boosted Decision Tree (BDT) selection described in Section 7.4.5.

### Shower Residual Analysis Cut

Firstly, an alternative method to count the number of showers in the interaction was created. Removing events with two showers reduces the  $\pi^0$  background. Due to the segmentation issue, discussed in Chapter 6, Section 6.1.5, a significant proportion of the signal is lost in the process. To prevent the loss of signal events, showers were only considered a secondary shower if they were above an energy threshold (10 MeV) and significantly far from the primary shower (the most energetic shower). The perpendicular distance from the secondary shower's start position and the primary shower's direction vector relative to the primary shower's start position was used to characterise how far away the secondary shower is. This distance is depicted in Figure 7.30 and is defined as the residual of the shower. The distributions of the residual for events can be seen in Figure 7.31, which shows discriminatory power. Figure 7.30 depicts secondary showers which are not considered in the analysis. The residual threshold analysis is discussed further in Appendix I.

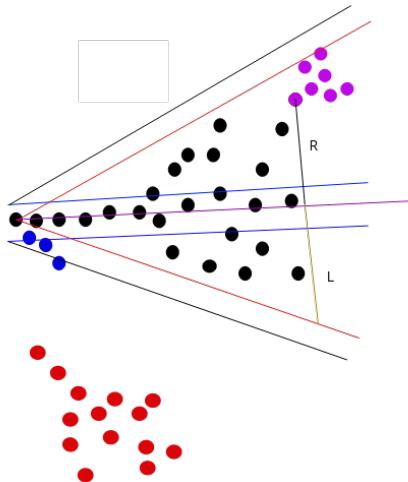


Fig. 7.30 The steps in the residual analysis. The distance from the start position of the secondary showers (blue, pink and red dots) is evaluated with respect to the primary shower (black dots). The residual of the secondary shower is calculated using the primary shower direction (pink line). Showers outside the black line (made from the 2 cm line (blue) and the cone outline (red)) are removed.

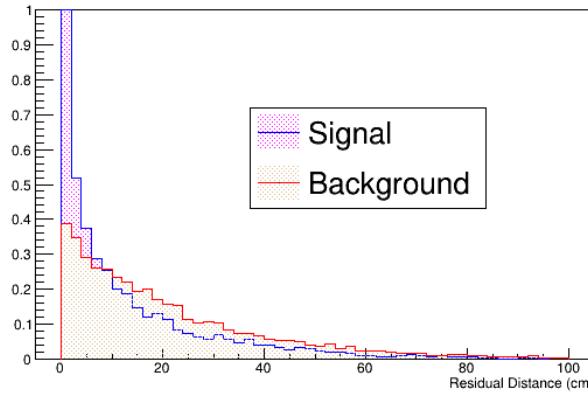


Fig. 7.31 The distribution of residual values of secondary showers.

### Reconstructed Neutrino PDGcode Cut

The Pandora pattern recognition provides a reconstructed neutrino PDGcode based on the largest primary daughter. If the daughter with the largest number of hits is labelled as shower-like, the neutrino PDGcode is reconstructed as a  $\nu_e$  (12). For a track-like particle, it is labelled as a  $\nu_\mu$  (14). Removing events with a PDGcode of 14 resulted in an efficiency of  $88.4 \pm 1.1$  (stat) % and a background rejection of  $58.7 \pm 0.7$  (stat) %.

### Maximum Track PIDA Cut

Track calorimetry information was also introduced to improve the selection using a PIDA analysis [271, 29], see Chapter 6, Section 6.1.8. The distribution of the PIDA score from the largest track in the signal and background is shown in Figure 7.32. The discriminatory power arises between pion/muon tracks and protons. Therefore, CC QE events where a proton track is released and reconstructed can be separated from the resonant NC and CC  $\nu_\mu$  background. The best efficiency  $\times$  background rejection results in an efficiency of  $48.6 \pm 0.7$ % and a background rejection of  $57.2 \pm 0.6$ % at a PIDA score of 10.75. The efficiency  $\times$  purity best cut is at 0.25 (minimum value in the search), resulting in an efficiency of  $100 \pm 1.2$ % and purity of  $0.1 \pm 0.0$ %.

The PIDA score was combined with the maximum track length to enhance the separation. As can be seen in the 2D distribution shown in Figures 7.33a and 7.33b, the signal has a concentration of maximum track lengths with a small track length and high PIDA score, whilst the background sample consists of longer tracks with a PIDA score of  $\sim 4$ . As discussed in Appendix I combining the two cuts provides an improvement of  $0.03 \pm 0.01$  in efficiency  $\times$  background rejection ( $3.7 \pm 1.4$ ,% loss in efficiency and  $5.7 \pm 1.0$ % increase

in background rejection) compared to the track length cut alone. Therefore, combining the track length cut with the PIDA score improves the selection.

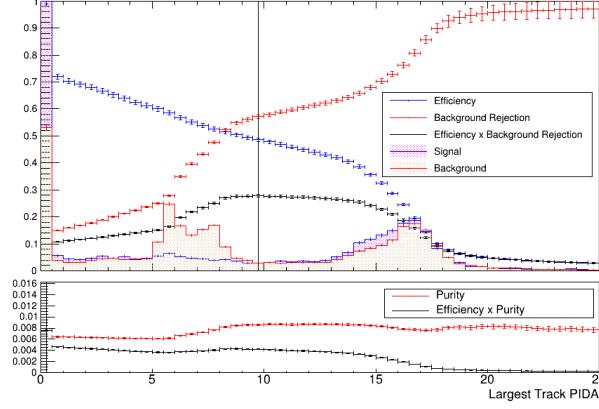


Fig. 7.32 The PIDA score for the largest shower for signal (blue hatched) and background (red hatched) events. The efficiency (blue), background rejection (red), efficiency  $\times$  background rejection (black), purity (red) and efficiency  $\times$  purity (black) curves are also plotted. The errors on the curves are the cumulative statistical error.

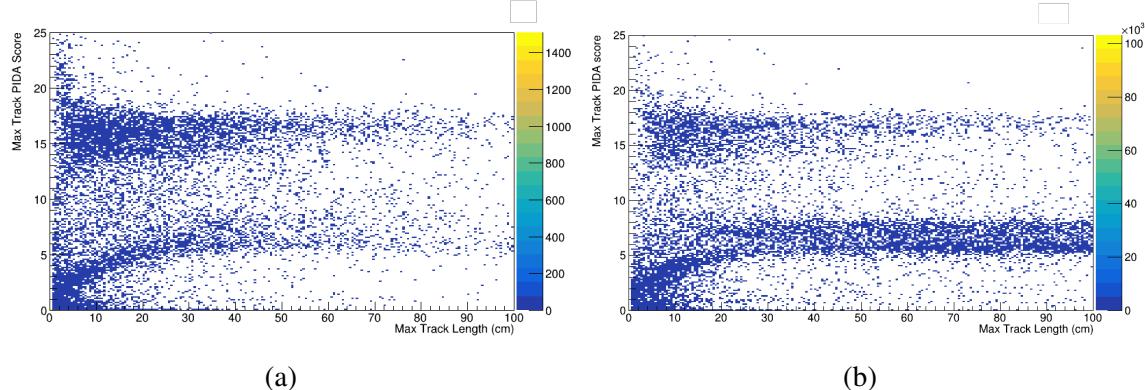


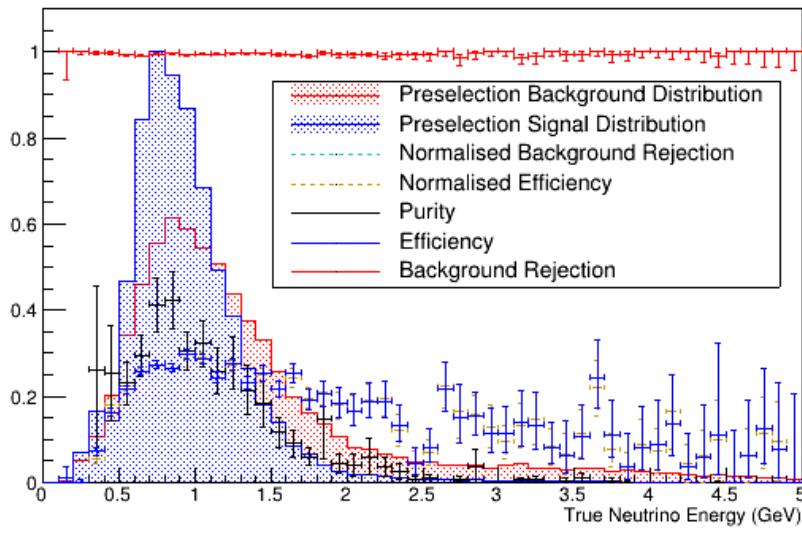
Fig. 7.33 The 2D signal (a) and background (b) distributions of the length and the PIDA score of the longest track.

### Total $v_e$ Cut-based Selection Using New Metrics

Replacing the one shower cut with the residual analysis, as well as including the neutrino PDGcode and including the PIDA 2D analysis, results in a total efficiency of  $25.23 \pm 0.31\%$  ( $23.55 \pm 0.28\%$ ) and a background rejection of  $99.39 \pm 0.00\%$  ( $99.38 \pm 0.00\%$ ). This is an improvement of  $3.81 \pm 0.43\%$  ( $3.71 \pm 0.38\%$ ) in efficiency and an decrease of  $0.17 \pm 0.00\%$  ( $0.17 \pm 0.00\%$ ) in background rejection. The purity ( $25.17 \pm 1.32\%$ ) decreased by  $3.30 \pm$

2.14%. The efficiency  $\times$  background rejection is  $0.251 \pm 0.003$  ( $0.234 \pm 0.003$ ), which is an improvement of  $0.037 \pm 0.004$ . The efficiency  $\times$  purity is  $0.064 \pm 0.003$ : an decrease of  $0.003 \pm 0.005$ . Comparing the efficiency and background rejection as a function of the neutrino energy, which are shown in Figures 7.34, whilst maximising with respect to the efficiency  $\times$  background rejection of each cut, shows no additional biases have been introduced into the selection as the structure of the curve is similar to the proposal selection.

Using the cuts which maximise the efficiency  $\times$  purity for each cut in the analysis, which are shown Table 7.6, results in a total efficiency of  $24.19 \pm 0.31\%$  ( $19.9 \pm 0.26\%$ ), an improvement of  $4.82 \pm 0.42\%$  ( $2.67 \pm 0.36\%$ ), and a background rejection of  $99.69 \pm 0.00\%$  ( $99.74 \pm 0.03\%$ ), a decrease of  $0.01 \pm 0.00\%$  compared to the proposal selection. The purity of  $39.16 \pm 2.35\%$  has decreased by  $9.97 \pm 4.11\%$  while the the efficiency  $\times$  background rejection of  $0.24 \pm 0.00$  ( $0.2 \pm 0.00$ ) has an improvement of  $0.047 \pm 0.004$  ( $0.026 \pm 0.036$ ). The efficiency  $\times$  purity is  $0.098 \pm 0.006$  is an decrease of  $0.0004 \pm 0.008$ .



(a)

Fig. 7.34 The selection efficiency of CC  $\nu_e$  events (blue), the background rejection for removing  $\nu_\mu$  events (red) and the purity (black) of the updated selection as a function of the true neutrino energy. The underlying distributions before selection for the signal (red) and background (blue) are also plotted, along with the efficiency (dashed orange) and background rejection (dashed light blue) for the POT and oscillation normalised events. The efficiencies are calculated from the remaining events after the pre-selection cuts and the full selection. For the unnormalised events, the errors are the  $1\sigma$  Clopper-Pearson intervals [32] whilst for the normalised distributions the normal approximation is used.

### Additional Metrics With Small Separational Power

The following metrics were also considered when trying to separate  $\nu_e$  CC interactions from the  $\nu_\mu$  backgrounds: Shower length and opening angle; the length and width of the initial track stub of the shower; The relative shower energy density gradient and form; the number of reconstructed neutrinos.

The above selection metrics are discussed in detail in Appendix I. The average selection efficiency and background rejection can be found in Table 7.6 and the distribution for the  $\nu_e$  CC signal events and  $\nu_{\mu}n$  background events can be found in Figures 7.38.

The first three metrics can separate the  $\nu_e$  events from the  $\nu_{\mu}n$  as they give an indication as to if a shower originated from an electron (i.e. a  $\nu_e$  CC event) or a photon (a  $\nu_{\mu}n$  event). For the length and opening angle this is because typically a photon shower is smaller than electron shower of the same energy. This is because the photon creates essentially two electron showers when pair production occurs. The  $e^+e^-$  pair then share the energy of the photon. Also, because photon showers have two particles in the initial track of the shower delta-rays and larger interactions are more common. This makes the initial track stub of the photon shower shorter and wider on average. In addition, as there is essentially two electron showers with a smaller energy on top of one another the energy profile of the photon shower is different to an electron shower. As shown in Appendix I this profile is well modelled by the function  $ax^b$  where a (the Relative Energy Density Gradient (REDG) and b (Relative Energy Density Power (REDP) are fit parameters. These fit parameters are used to separate the electron and photon showers.

The number of reconstructed neutrinos is used as well due to inefficiencies in the reconstruction. It is possible that a neutrino interaction with two showers is split into two separate interactions by the reconstruction. Thus the cut is applied to remove neutral pion interactions which have been incorrectly reconstructed.

This Section has presented new methods to the SBND selection for separating the  $\nu_e$  signal from the  $\nu_\mu$  background. Some cuts have been used to improve the traditional cut based selection. However, other cuts provide little separational power. Therefore, a BDT approach is introduced in the next Section so that these metrics can be included effectively.

#### 7.4.5 Performing a Multivariate Analysis to improve the selections

To use the weak classifiers described in the Section 7.4.4, a Boosted Decision Tree (BDT) was implemented. Firstly, the BDT process will be described then the performance of selecting  $\nu_e$  CC using the BDT mechanism will be evaluated.

## BDT Methodology

The gradient boosting algorithm within the TMVA [281] root package was employed to make the BDT. The premise behind the gradient boosting algorithm is to perform several cuts using decision trees. The decision tree splits the signal and background using multiple optimised cuts which depend on the parameters of the event. In the case of gradient boosting, several trees are combined in order to reclassify the misidentified events from previous tree iterations. Therefore, the algorithm requires training data in order to define the structure of the trees. Once the structure is defined, the BDT can be used on actual data. This procedure requires caution as the BDT can over-classify the training data, which is known as over-training, resulting in poor performance on actual data.

More specifically, the gradient boosting algorithm [285] for a training set  $x_i$  (containing  $p$  signal and  $n$  background events), which have a classification score  $y_i$ , is shown for  $M$  trees in Algorithm 2. To maximise the classification a function  $F$  is used which commonly corresponds to the  $\log(\text{odds}) = \log\left(\frac{P}{1-P}\right)$ , where  $P$  is the probability of the data set being classified as a signal event [286]. A loss function  $L$  is then used to define the ability to classify events using the decision process. The loss function is typically set as the negative binomial log-likelihood:  $\log(1 + e^{-2yF})$ , where  $y$  are the observed values of the sample and  $F$  is the prediction [286]. Therefore, minimising the loss function results in the predictor which maximises the likelihood of extracting the signal [285, 286]. The algorithm initialises the initial model such that the loss function,  $L$ , is minimised with respect to a  $F = \gamma$  value.

The residuals of the training data ( $F - y$ ) are then calculated and used to create a decision tree. Using the residuals to make the tree ensures that more weight is given to incorrectly matched data in the next tree. An example tree from the trained BDT, where the signal and background are normalised 1:1, with the input parameters in Table 5.1, is shown in Figure 7.35. Each stage of the decision tree is chosen to maximise the classification greedily. To choose which variables maximise the classification, the entropy loss (or information gain) at each node is maximised. The information gain is defined as the total entropy of the training set minus the remainder of the entropy after the node. Entropy at a two point classification node is defined as  $H = q \log_2 q + (1 - q) \log_2 (1 - q)$ , where  $q = \frac{p}{p+n}$  is the probability of the data point being classified [287]. The information gain for a two point classification node is then defined as

$$G = H\left(\frac{p}{p+n}\right) - \sum_{k=1}^d \frac{p_k + n_k}{p+n} H\left(\frac{p_k}{p_k + n_k}\right), \quad (7.1)$$

where the sum is over the distinct values of  $d$  for the variable in question [287]. For continuous variables, the value which splits the data resulting in the highest information gain is used.

---

**Algorithm 2:** pseudo-code for the gradient boost algorithm. From Reference [286]

---

1. Initialise model with constant value predictor  $F_0(\mathbf{x})$  which minimises the loss function ( $L$ ) for elements in the training set:  $F_0(x) = \arg \min_{\gamma} \sum_{i=1}^n L(y_i, \gamma)$ .
  2. **for**  $m=1$  to  $M$  **do**
    1. Compute the so-called pseudo-residuals for the elements  $i$  in the training set  $x$  by evaluating the differential of the loss function by the predictor  $F$  evaluated at the previous predictor:  $F_{m-1}$ :  $r_{im} = - \left[ \frac{\partial L(y_i, F(x_i))}{\partial F(x_i)} \right]_{F(x)=F_{m-1}(x)}$ .
    2. Fit a decision tree  $h_m$  to training data such that the tree maximises the information gain (or entropy loss) within a greedy search using the pseudo-residuals.
    3. Compute the multipliers for each leaf,  $\gamma_m$ , which minimise the loss function:  

$$\gamma_m = \arg \min_{\gamma} \sum_{i=1}^n L(y_i, F_{m-1}(x_i) + \gamma h_m(x_i))$$
    4. Recalculate the prediction using the old prediction and the new decision tree, multiplied by  $\gamma_m$  and a learning rate  $v$ :  $F_m(\mathbf{x}) = F_{m-1}(\mathbf{x}) + v \gamma_m h_m(\mathbf{x})$
  - end**
- 

Once the decision tree is calculated, the multipliers  $F_m = \gamma_m = \log(odds_m)$  for each leaf which minimise the loss function are calculated. The loss function is evaluated with the old prediction  $F_{i-m}$  and the new decision tree  $h_m$ . Therefore, the algorithm is analogous to gradient descent which increments along the fit parameter space to find the minimum in the loss function, i.e. the best fit. The final prediction is the initial guess plus the sum of the decision trees  $F_M = F_0 + \sum_m v \gamma_m h_m(\mathbf{x})$ . A learning rate,  $v$ , is applied such that a decision tree does not solve the training set exactly and therefore become overtrained.

Bagging [288] is also used to improve the BDT response and prevent overtraining. Bagging is the act of only allowing a decision tree to be trained on a random subset of the training data each time. This reduces the variance of the BDT result. Cross-validation [289] is also performed to enhance the robustness of the algorithm. To train a signal BDT a training set is required. A test data set is also required to evaluate the BDT response to new data. In cross-validation, the datasets are combined and a random split of the data is used to train several BDTs. A weighted average of the classification score of new data from each BDT is used as the new response. This method reduces overtraining by averaging out the responses and improves the BDT response to rare channels. Ten BDTs are used in this analysis.

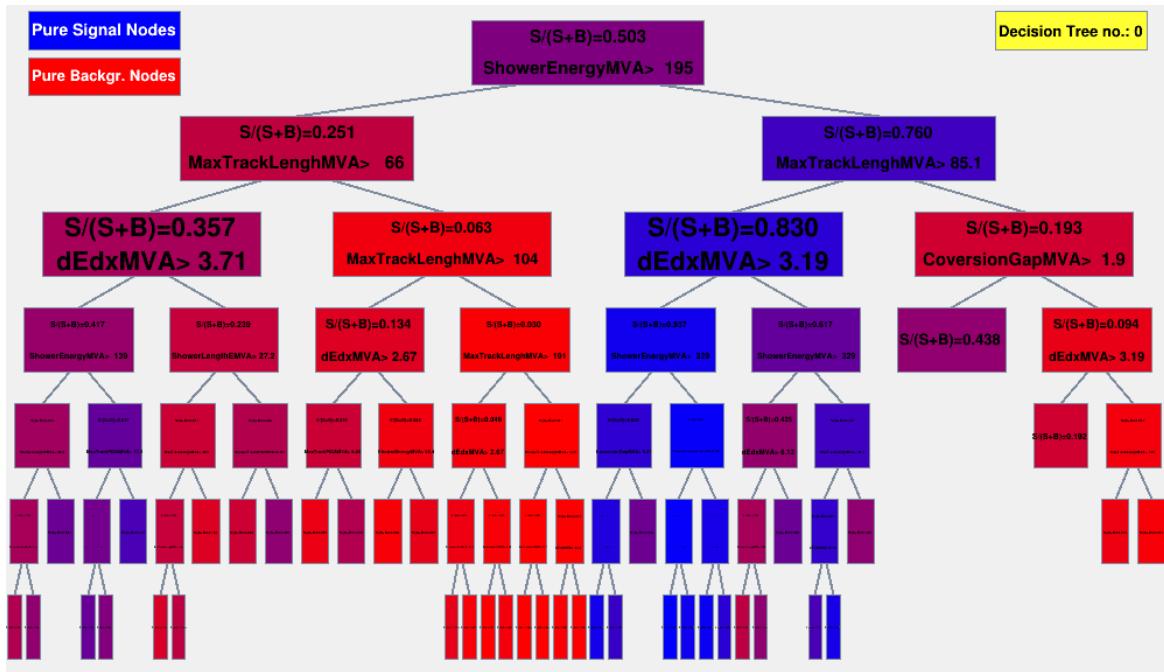


Fig. 7.35 An example of the first decision tree from the BDTG algorithm from the root TMVA package. The hyperparameters of the BDT are shown in Table 7.5 and the input parameters are shown in Table 7.6. The signal and background, Figure 7.38, are scaled to 1:1 before the BDT analyses the data.

The BDT class in TMVA allows for the hyperparameters: the number of decision trees, the bagging fraction, the minimum size of a node, the tree depth and a shrinkage parameter which defines the learning rate to be set by the user. These parameters were optimised using the Bayesian Optimisation [277] described in Chapter 6, Section 6.2.2. The parameters were tuned in order to maximise the area under the efficiency  $\times$  background rejection ROC curve of the full dataset. The RMS of the ten BDTs was evaluated to ensure the stability of the results and, if the RMS was greater than 0.01, the result was discarded. This was done by giving a ROC area result to the Bayesian Optimisation of 0, allowing the optimisation process to recognise the parameters as a poor choice.

To prevent overtraining, during optimisation, the BDT response for the training set was compared to the BDT response of a validation dataset using the Kolmogorov–Smirnov (KS) score (with option "X") in ROOT TH1 class [290]. The whole dataset after the two pre-selection cuts was split in a 25 K signal and 35 K background training sets, as well as a 5.2 K signal and 9.6 K background validation sets. The KS calculation takes random samples of the training data which are the same size as the validation data. The maximum cumulative difference between the full training data and the subsample of the training data is then evaluated and is compared to the maximum cumulative difference of the full training

data and the validation data. If the difference is larger for the subsample, then the validation data is deemed indistinguishable from the training data. The process is repeated over 1000 random subsamples. The fraction of subsamples which are deemed indistinguishable from the training data is then evaluated and is expected to be 50% when the training data and the validation data are within statistical error. To ensure no overtraining, BDT tunes where the KS score was less than 0.1 were labelled as overtrained, discarded and treated similarly to the large RMS case above.

### BDT Implementation and Results

Two BDTs were tuned for the analysis: the first allowed for the signal and background events to be normalised to 1:1 and is labelled standard BDT; the other maintained the POT and oscillation normalisation and is labelled scaled BDT. Figures 7.36 and 7.37 show the final test and training distributions for the standard BDT and the scaled BDT respectively. The fact that the training and validation distributions are similar shows that over-training was prevented. The signal KS score is 0.168 and 0.734 for the standard and scaled BDT respectively. For the background, the KS score is 0.499 and 1 for the standard and scaled respectively. Table 7.5 shows the result of the Bayesian Optimisation tuning of the hyperparameters for both BDTs.

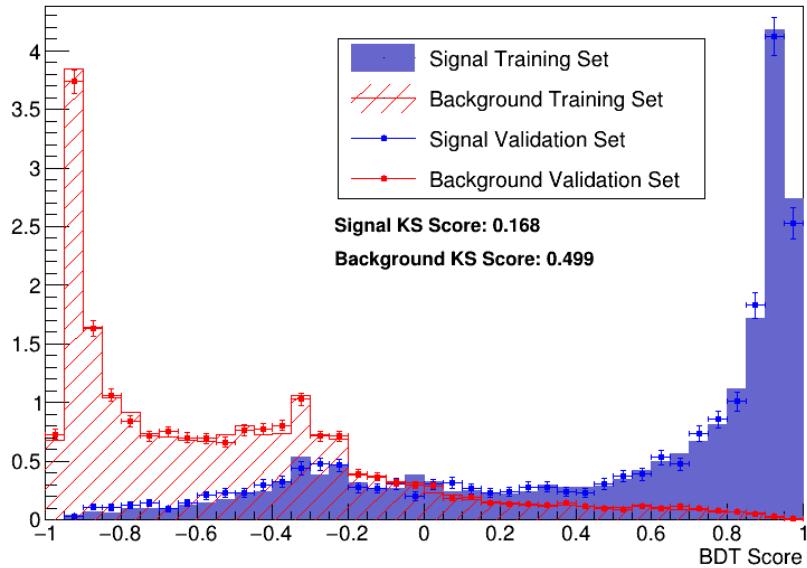


Fig. 7.36 The BDT response for the standard BDT for the validation (points) and the training (bars) data for the signal (blue) and background (red). The errors are Poissonian.

Hyperparameter	Standard BDT	Scaled BDT
Number of Trees	864	1521
Max Tree Depth	6	8
Shrinkage	0.005	0.054
Min Node Size (%)	1.0	11.99
Bagging Fraction	0.9	0.64

Table 7.5 The BDT hyperparameters for the tuned BDTs.

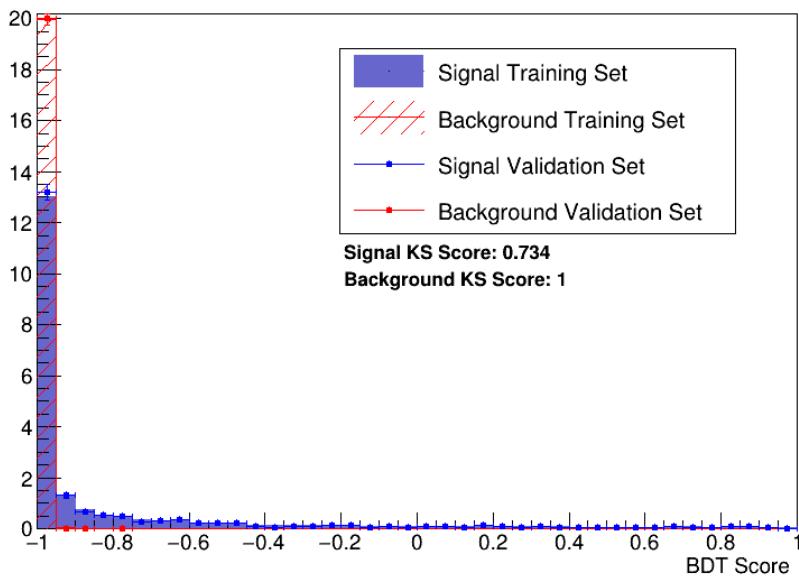


Fig. 7.37 The BDT response for the scaled BDT for the validation (points) and the training (bars) data for the signal (blue) and background (red). The errors are Poissonian.

Table 7.6 describes the input variables used within the BDTs and shows the fraction of times the variables were used within the standard BDT. Figures 7.38 show the normalised input distributions to the BDTs, which set maximum and minimum limits on the parameters. Correlation matrices of the input parameters for the signal, shown in Figure 7.39, and background, shown in Figure 7.40, show strong correlations between the energy and length of the showers. This is expected due to the logarithmic relation shown in Equation I.2. There are also correlations with other shower parameters, most of which depend on the shower energy. There is a strong correlation between the REDG and REDP, which is expected as the two parameters are from the same fit. There are also large correlations between the maximum track length and the PIDA score, which is expected from the PIDA analysis. As well as this, there are strong correlations between the reconstructed neutrino PDG and the track length as well as the shower parameters. This is because the PDG is chosen from the primary particle with the most hits. Therefore, events with large and long tracks are given a score of 14 whereas events with large showers are given a score of 12.

Input Parameter	BDT Variable Importance	Figure	Best Cut Efficiency $\times$ Background Rejection Best Cut Efficiency $\times$ Purity		
			Value	Efficiency	Background Rejection/Purity
Shower Conversion Gap	0.1259 (8)	7.25	2.1 (cm) 3.5 (cm)	71.2 $\pm$ 1.9% 77.5 $\pm$ 1.0%	61.1 $\pm$ 0.7% 1.36 $\pm$ 0.0%
Neutrino PDGcode	0.243 (3)	N/A	12.5	88.4 $\pm$ 1.1% 88.4 $\pm$ 1.1%	58.7 $\pm$ 0.7% 1.6 $\pm$ 0.0%
Number of Neutrinos	0.023 (13)	7.38	1.5 1.5	91.1 $\pm$ 1.1% 91.1 $\pm$ 1.1%	19.2 $\pm$ 0.2% 0.8 $\pm$ 0.0%
Number of Excess Showers from Residual Analysis	0.033 (11)	7.38	0.5 3.5	77.4 $\pm$ 0.1% 96.2 $\pm$ 1.1%	38.9 $\pm$ 0.4% 0.01 $\pm$ 0.01%
Shower Energy	0.312 (1)	7.28	225 (MeV) 345 (MeV)	75.1 $\pm$ 1.0% 56.9 $\pm$ 0.8%	77.1 $\pm$ 1.1% 4.4 $\pm$ 0.1%
Shower Length	0.258 (2)	I.4	37.5 (cm) 49.5 (cm)	76.6 $\pm$ 1.0% 65.9 $\pm$ 0.9%	71.3 $\pm$ 1.0% 2.6 $\pm$ 0.0%
Shower Opening Angle	0.087 (9)	I.6	0.315 (rad) 0.405 (rad)	63.9 $\pm$ 0.9% 77.1 $\pm$ 1.0 %	61.1 $\pm$ 0.8% 1.1 $\pm$ 0.0%
Shower Relative Energy Density Gradient (REDG)	0.158 (6)	I.8a	0.07 (cm $^{-4}$ ) 0.23 (cm $^{-4}$ )	64.0 $\pm$ 0.9% 89.5 $\pm$ 1.2%	53.6 $\pm$ 0.8% 1.4 $\pm$ 0.0%
Shower Relative Energy Density Power (REDP)	0.176 (5)	I.8b	1.49 1.63	55.6 $\pm$ 0.8% 91.9 $\pm$ 1.2%	61.5 $\pm$ 1.0% 1.5 $\pm$ 0.0%
Shower Track Length	0.021 (14)	I.11a	3.1 (cm) 9.9 (cm)	46.2 $\pm$ 0.7% 95.1 $\pm$ 1.1%	47.2 $\pm$ 0.5 % 0.8 $\pm$ 0.0 %
Shower Track Width	0.025 (12)	I.11b	0.35 (cm) 4.55 (cm)	56.1 $\pm$ 0.8% 99.7 $\pm$ 1.2%	56.0 $\pm$ 0.6% 0.8 $\pm$ 0.0%
Shower $dE/dx$	0.126 (7)	7.26a	2.5 (MeV/cm) 2.5 (MeV/cm)	67.6 $\pm$ 0.9% 67.6 $\pm$ 0.9%	63.5 $\pm$ 0.8% 1.4 $\pm$ 0.0%
Maximum Track Length	0.219 (4)	7.27	38 (cm) 54 (cm)	83.8 $\pm$ 1.0% 90.7 $\pm$ 1.1%	55.3 $\pm$ 0.6% 1.4 $\pm$ 0.0%
Maximum Track PIDA	0.079 (1)	7.32	9.75 0.25	48.6 $\pm$ 0.7% 100 $\pm$ 1.2%	57.2 $\pm$ 0.6% 0.1 $\pm$ 0.0 %
Number of Showers	N/A	7.38	1 10	71.8 $\pm$ 0.9 1 $\pm$ 1.2%	41.0 $\pm$ 0.4% 0.8 $\pm$ 0.0%
Standard BDT Score	N/A	7.41	0.075 0.975	77.4 $\pm$ 2.2% 39.2 $\pm$ 1.3%	87.0 $\pm$ 3.0% 31.7 $\pm$ 1.4%
Scaled BDT Score	N/A	7.42	-0.925 -0.925	34.2 $\pm$ 1.2% 34.2 $\pm$ 1.2%	99.9 $\pm$ 40.8% 72.1 $\pm$ 3.1%
Fiducial Volume	N/A	7.42	N/A	71.489 $\pm$ 0.0%	29.6 $\pm$ 0.0 % 0.1 $\pm$ 0.0%

Table 7.6 Selection parameters used in the analyses. The fraction of times the parameter is used in the standard BDT is shown (similar for the scaled BDT). The efficiency, background rejection and purity for the best cut with respect to maximising the efficiency  $\times$  background rejection and efficiency  $\times$  purity for each variable independently is shown. The Figures associated to the cuts are presented.

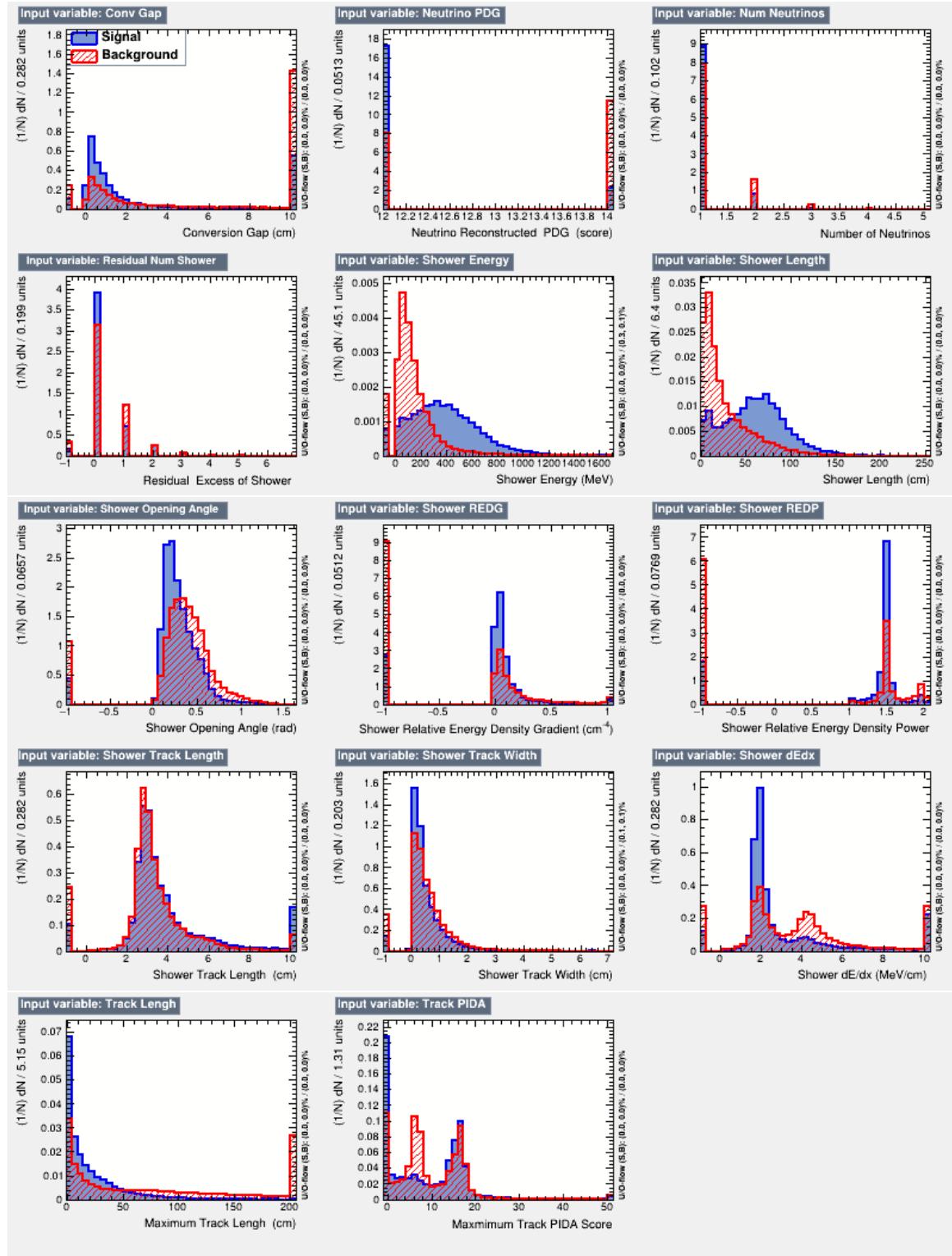


Fig. 7.38 The input parameter distribution for the BDTs. The inputs are the remaining events after the PFP neutrino and at least one shower pre-selection cuts. The blue histograms correspond to the  $\nu_e$  CC signal and the red to the  $\nu_\mu$  background. The distributions go in descending order of the Table 7.6 from the top left Figure to the bottom right Figure.

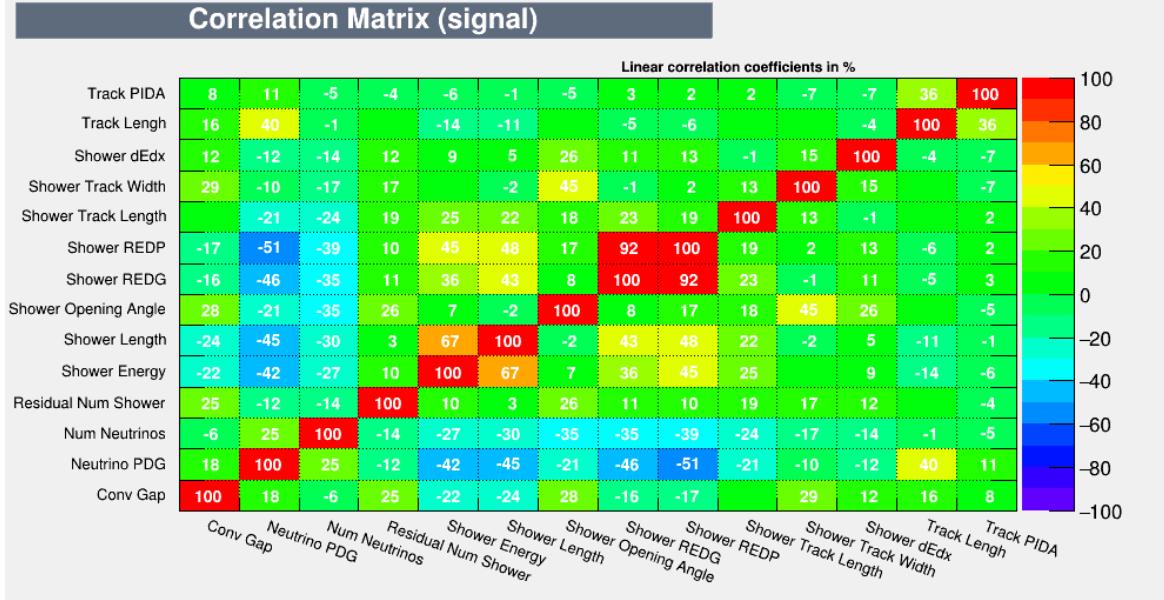


Fig. 7.39 Correlation matrix between the input parameters of the BDTs, listed in Table 7.6, for the  $\nu_e$  CC signal data. The correlation is represented as a percentage.

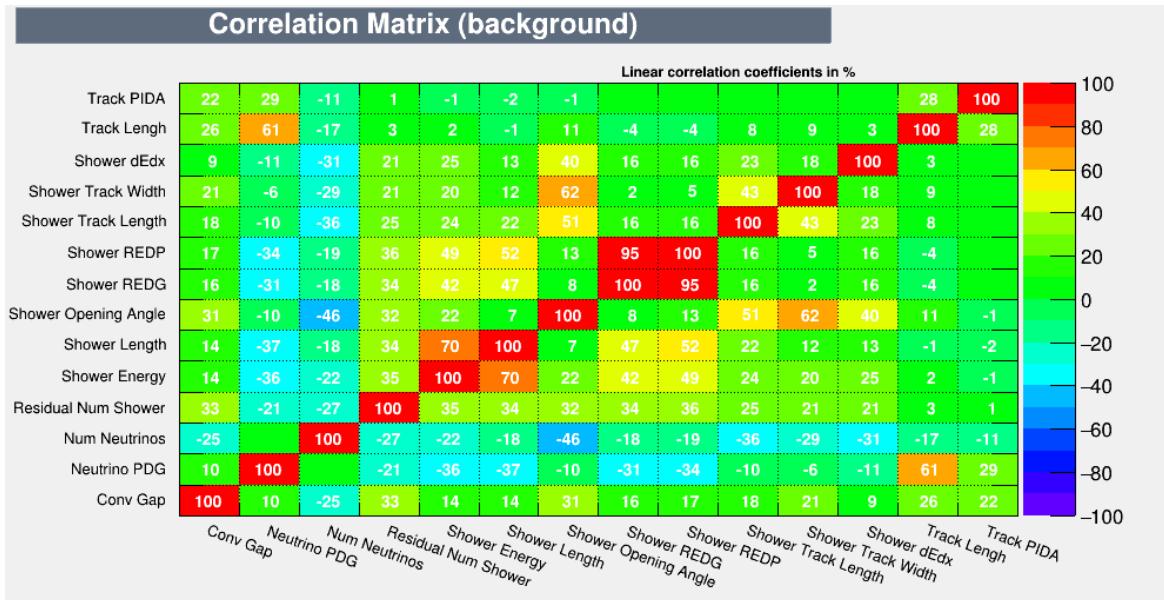


Fig. 7.40 Correlation matrix between the input parameters of the BDTs, listed in Table 7.6, for the  $\nu_\mu$  background data. The correlation is represented as a percentage.

Correlations between metrics can indicate that the separational power of the metrics is a result of the same underlying phenomena. In this case, the CPU efficiency of the BDT is reduced as additional links between variables are made, which are not necessary. As the inefficiency of the BDT was not a bottleneck, reducing the correlations was not investigated.

Figures 7.41 and 7.42 show the performance on cutting on the BDT score alone for the training dataset. For the standard BDT, the best cut for efficiency  $\times$  background rejections is  $0.67 \pm 0.03$  at a BDT score cut off of 0.075. These values result in an efficiency of  $77.4 \pm 2.1\%$  and a background rejection of  $87.0 \pm 2.9\%$ . In order to maximise the efficiency  $\times$  purity using the POT and oscillation normalised distributions, the maximum efficiency  $\times$  purity is  $0.13 \pm 0.02$  for a BDT score cut off of 0.975. This result in an efficiency of  $39.3 \pm 1.3\%$  and a purity of  $31.8 \pm 3.6\%$ . For the scaled BDT, the maximum efficiency  $\times$  background rejection is  $0.34 \pm 0.13$  at a BDT score cut of -0.985 with an efficiency of  $34.2 \pm 1.2\%$  and background rejection of  $99.9 \pm 40.8\%$ . Maximising the efficiency  $\times$  purity results in a value of  $0.25 \pm 0.03$  at a BDT cut off of -0.925, an efficiency of  $34.2 \pm 1.2\%$  and a purity of  $72.1 \pm 8.7\%$ . The cuts show powerful separation between the signal and background. However, using the values which maximise the efficiency  $\times$  background rejection results in a much larger background remaining than if the cuts were performed independently. As the efficiency  $\times$  background rejection maximum is shallow and also the errors are large, a BDT score which removes more of the background, increasing the purity of the sample, could result in a better sterile sensitivity. Using various BDT scores is investigated in Chapter 8.

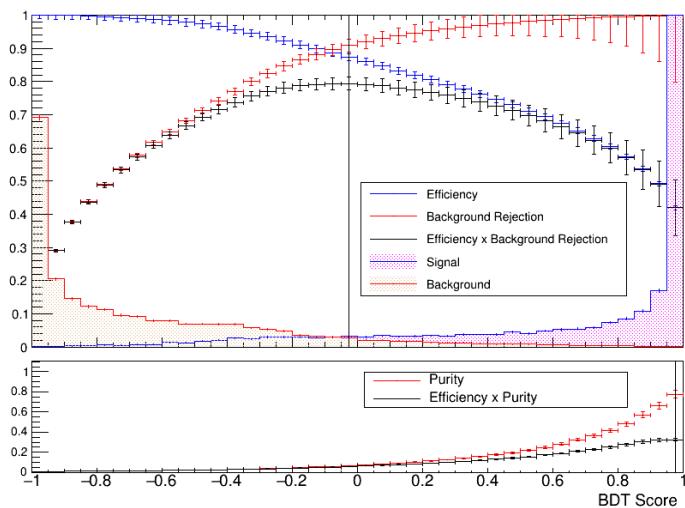


Fig. 7.41 The efficiency (blue), background rejection (red), purity (red), efficiency  $\times$  background rejection (black) and efficiency  $\times$  purity (black) for selecting events above the BDT score cut off from the standard trained BDT. The errors on the curves are the cumulative statistical error.

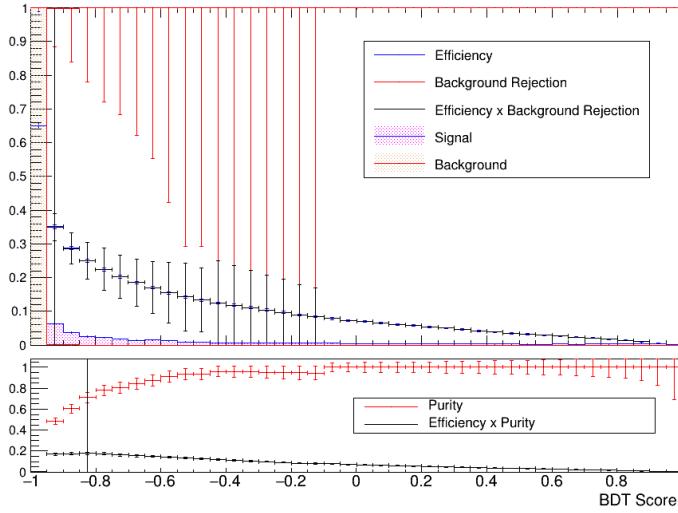
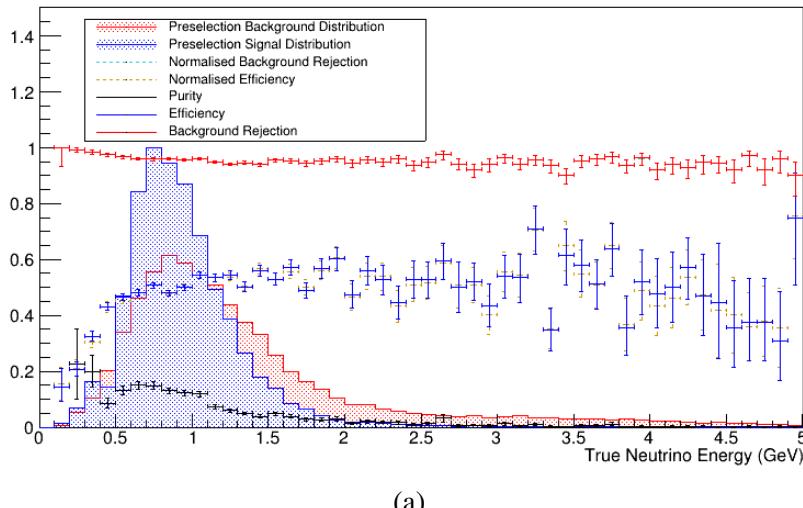


Fig. 7.42 The efficiency (blue), background rejection (red), purity (red), efficiency  $\times$  background rejection (black) and efficiency  $\times$  purity (black) for selecting events above the BDT score cut off from the scaled trained BDT. The BDT inputs have been normalised using the POT and oscillation weighting. The errors on the curves are the cumulative statistical error.

Using the standard BDT at a cut value of 0, applying the fiducial volume cut and applying a minimum shower energy cut of 100 MeV to remove cosmic background further, results in the efficiency curves shown in Figure 7.43. The BDT reduces the dependencies of the selection on the neutrino parameters, notably in the higher energy range. There is only a notable drop in efficiency in the lower energy regions due to poor reconstruction and the 100 MeV cut-off. This is a key region for sterile searches and investigations to improve the efficiency at lower energies are on going. When performing a sensitivity analysis, the flattening of the efficiency as a function of the neutrino energy makes the BDT less susceptible to changes in the event rate due to underlying oscillation parameters. For example, the oscillation values could be such that the probability is significantly lower for neutrino energy around the beam peak ( $\sim 1$  GeV). The capabilities of the analysis would then rely on the oscillations for higher (or lower) neutrino energies. Thus it is useful to maintain a high efficiency across all neutrino energies.

When maximising the efficiency  $\times$  background rejection, the standard BDT analysis improves the efficiency  $\times$  background by  $0.21 \pm 0.08$  ( $0.22 \pm 0.07$ ) due to an increase in efficiency of  $23.72 \pm 0.83\%$  ( $24.37 \pm 0.77\%$ ) compared to the new cut-based selection. This increase, however, is at the cost of a decrease of  $4.43 \pm 0.05\%$  ( $4.42 \pm 0.24\%$ ) in background rejection and a  $19.15 \pm 1.36\%$  decrease in purity. Therefore the efficiency  $\times$  purity is decreases only by  $0.006 \pm 0.004$ . For the scaled BDT, the efficiency  $\times$  background rejection

decreases by  $0.07 \pm 0.01$  ( $0.06 \pm 0.01$ ) with an decrease of  $7.30 \pm 0.67\%$  ( $5.90 \pm 0.61\%$ ) in efficiency and an increase in background rejection of  $0.60 \pm 0.01\%$  ( $0.60 \pm 0.05\%$ ). This corresponds to a resultant increase in purity of  $52.69 \pm 7.85\%$  and an increase in efficiency  $\times$  purity of  $0.85 \pm 0.07$ . As was discussed above, the BDT combines several cuts into one cutting of the efficiency  $\times$  background rejection resulting in a large amount of background. Therefore, alternative BDT scores should be considered in the sterile analysis.



(a)

Fig. 7.43 The efficiency of selecting CC  $\nu_e$  events (blue), the background rejection for removing  $\nu_\mu$  events (red) and the purity (black) of the standard BDT selection with a BDT cut off score at 0, a fiducial volume cut and a minimum shower energy cut of 100 MeV, as a function of true neutrino energy. The underlying distributions before selection for the signal (red) and background (blue) are also plotted, along with the efficiency (dashed orange) and background rejection (dashed light blue) for the POT and oscillation normalised events. For the unnormalised events, the errors are the  $1\sigma$  Clopper-Pearson [32] intervals whilst for the normalised distributions the normal approximation is used.

When using the BDT cut off which maximises the efficiency  $\times$  purity, the standard BDT reduces the efficiency by  $3.22 \pm 0.70\%$  ( $0.52 \pm 0.64\%$  (increase)) and increases the background rejection by  $0.20 \pm 0.00\%$  ( $0.15 \pm 0.00\%$ ). This increases the purity by  $17.5 \pm 8.16\%$ . Therefore the efficiency  $\times$  background rejection is reduced by  $0.032 \pm 0.007$  ( $0.005 \pm 0.006$  (increase)), but the efficiency  $\times$  purity is increased by  $0.471 \pm 0.078$ . For the scaled BDT, the efficiency is reduced by  $6.26 \pm 0.67\%$  ( $2.28 \pm 0.61\%$ ) and the background rejection is increased by  $0.29 \pm 0.01\%$  ( $0.25 \pm 0.03\%$ ), which increases the purity to  $52.69 \pm 7.85\%$ . Therefore, the efficiency  $\times$  background rejection is reduced by  $0.062 \pm 0.007$  ( $0.022 \pm 0.006$ ) and the efficiency  $\times$  purity is increased by  $0.823 \pm 0.075$ . Although both result in an increase in efficiency  $\times$  purity, the efficiency is diminished, which may cause statistical fluctuations to

be significant for certain neutrino energy ranges due to a lack of data. Therefore alternative BDT score thresholds should be studied. Figure 7.44 shows the ROC curves for the BDT score thresholds with the fiducial volume and energy cut-off. The traditional cut based results are also plotted on the Figure to indicate the improvement due to the BDT analyses.

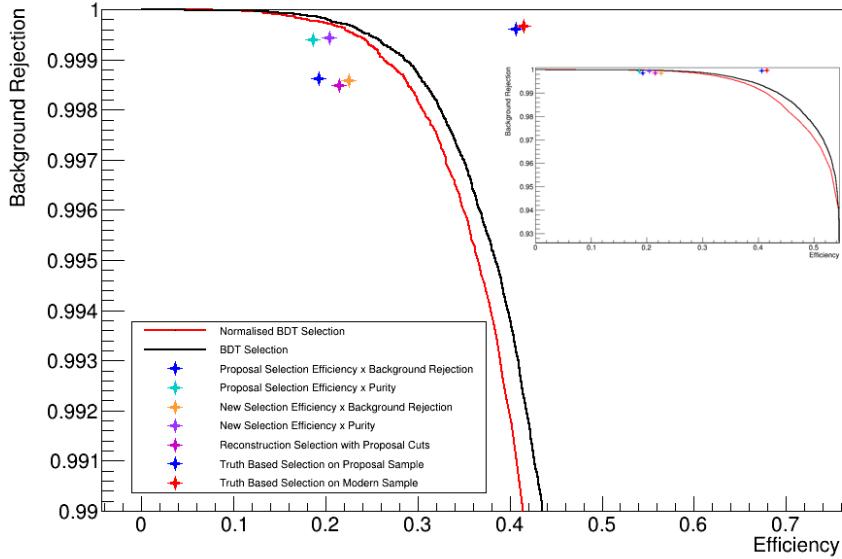


Fig. 7.44 The different selections in efficiency background rejection space for active volume events only. The cut-based selections are optimised by maximising either the efficiency  $\times$  background rejection for each cut or the efficiency  $\times$  purity. The BDT curves are calculated by varying the BDT score threshold within the selection. The subfigure is the whole curve. The 80% reconstruction efficiency weighting is applied to the truth-based analyses.

## 7.5 Comparing The Selection

Including the efficiencies of the pre-selection cuts in Section 7.4.3 in the cut based and BDT analyses results in a decrease in efficiency of 14%. However, the background rejection is increased by 0.62%, which corresponds to a factor of 5.38 reduction in the background. With the inclusion of the pre-selection efficiency, the full selection is comparable to the truth analysis for the "modern" sample discussed in Section 7.1.1. The difference in the efficiency and background rejection for the truth-based analysis and the reconstruction selections is shown in Figures 7.45 and 7.46 respectively. In addition, Figure 7.44 shows points in efficiency/background rejection space for the tuned reconstruction selection analyses and the truth-based selections for the "modern" and "proposal" samples. The BDTs are represented by ROC curves in the space which show how the efficiency and background rejection evolves

as a function of the BDT score threshold. As the ROC curves encapsulate the cut based selections, the BDT analysis outperforms the traditional cut based analyses. In the majority of the space, the standard BDT, which normalises the signal and background to 1:1, outperforms the BDT where the signal has been weighted by the oscillation probability. Only in the higher background rejection range does the scaled BDT become comparable. Therefore, various BDT scores for both BDTs are used to calculate the sensitivity to  $\nu_e$  appearance in the SBN programme.

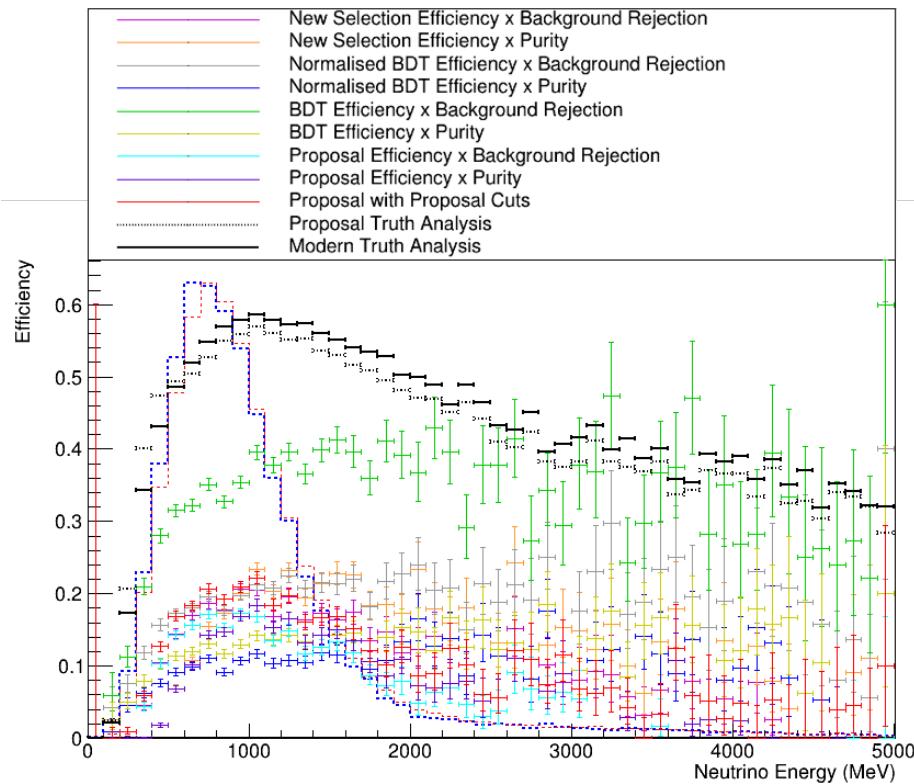


Fig. 7.45 The total selection efficiency as a function of the true neutrino energy. The underlying signal distribution (blue dashed) and background distribution (red dashed) are shown for convenience. The 80% reconstruction efficiency weighting is not applied. The errors on the reconstruction analyses are the  $1\sigma$  Clopper-Pearson intervals [32] whereas, for the truth-based analyses the normal approximation is used.

Figures 7.45, 7.46 and 7.44 show that the truth-based analysis outperforms the reconstruction analyses in both signal selection efficiency and background rejection. This is primarily due to artefacts in the reconstruction. For example, the truth-based analysis assumes 80% reconstruction efficiency on all events [17]. However, the analysis does not consider the misidentification of particles. The correctness of the 80% reconstruction efficiency can be evaluated using the correctness of the shower start position, discussed in Chapter 6, Section

6.2.2, and the vertex reconstruction efficiency, shown in Figure 7.24. The shower starting position is essential for the  $dE/dx$  cut and, combined with the vertex position, is also essential for the conversion gap cut. Combining the fact that 14% of  $\nu_e$  CC events are removed as there is no shower reconstructed in the drift window and that less than 80% of the  $\nu_e$  events with a shower have vertex a reconstructed within 5 cm of the true position, results in 31.2% of events being incorrectly reconstructed. Additionally, combining to that the fact that 41.2% of showers have a reconstructed start position greater than 5 cm away from the true start position, the fraction of events that are reconstructed "correctly" is much less than 80%. The choice of 5 cm is arbitrary; however, the majority of true track stub lengths of the showers in the "vertex" sample are within 5 cm, and thus for events above 5 cm the result is meaningless. Requiring that the reconstruction does not affect the selection significantly, one can, at best, approximate the reconstruction efficiency from this information as approximately 55%. A more detailed study of the reconstruction efficiency is required.

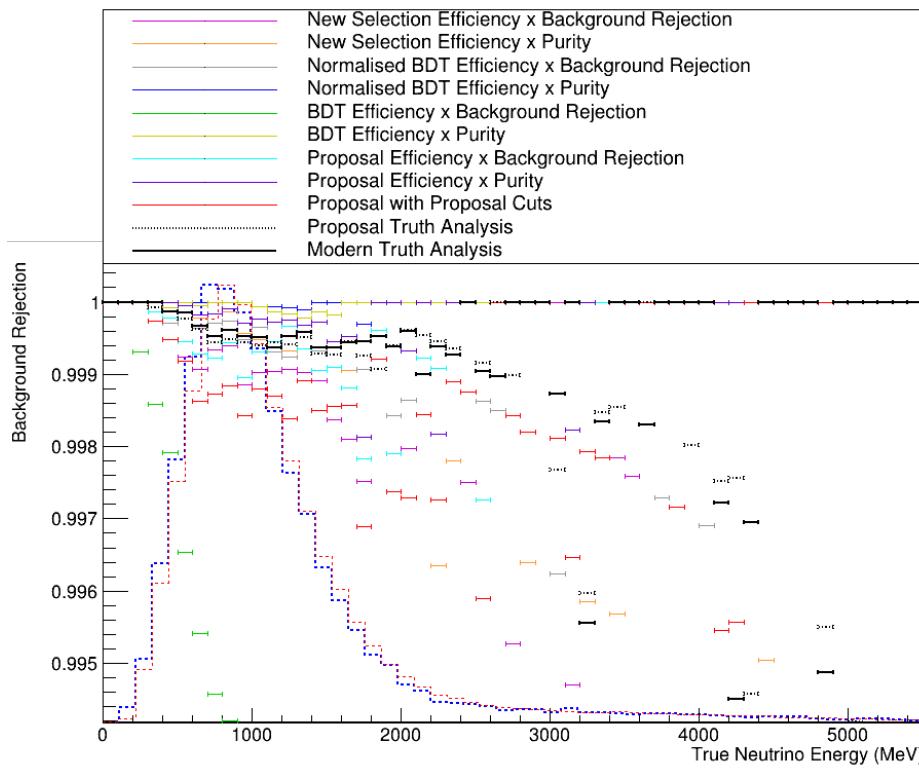


Fig. 7.46 The total background rejection as a function of the true neutrino energy. The underlying signal distribution (blue dashed) and background distribution (red dashed) are shown for convenience. The 80% reconstruction efficiency weighting is not applied. The errors are calculated using the normal approximation.

Furthermore, efficiency losses in the maximum track length and  $dE/dx$  cuts are not considered in the truth-based analysis, e.g. the  $dE/dx$  cut assumes 100%. It is known from the  $dE/dx$  truth-based analysis, discussed in Section 7.4.3, that the efficiency is not 100%. The  $dE/dx$  truth analysis also shows that an estimate of 94% in background rejection is optimistic for the BNB [246, 17]. Replacing the 94% with the 83.3% obtained from the truth-based analysis results in a more comparable background rejection of 99.9% for the truth-based analysis. Therefore, even though improvements are required in reconstruction to improve the selection, the truth-based proposal analysis is optimistic and a new estimate should be considered for future analyses.

## 7.6 Concluding Remarks

In order to progress to a full reconstruction selection in the SBN programme, a truth-based selection was created. In order to mimic the proposal selection using modern physics, a modern Monte Carlo production was created to include new physics, such as MEC events. This sample is known as the "modern" sample. GENIE [225] scaling parameters were then applied to mimic the sample used in the SBN proposal. The resultant sample is known as the "proposal" sample. A selection which followed the same selection procedure as the SBN proposal [17] was then performed on the "modern" sample and the "proposal" sample. The results of the selections were evaluated and compared to verify the proposal and see how changes in the SBN programme and neutrino physics have affected the result. It was shown that MEC events increase the signal by  $\sim 20\%$  in the "modern" sample compared to the "proposal" sample. In addition, a reduction of NC events was observed in the low reconstructed neutrino energy range. However, the new "proposal" sample did not fully correlate to the proposal spectra. This could be due to rare events. The selected events are used in the sensitivity analysis in Chapter 8.

Wire TPC information was reconstructed in SBND using the methods described in Chapter 6 with the addition of a SVM to characterise reconstructed showers and tracks. These samples were then used to identify separational metrics in the reconstruction. The proposal selection was then evaluated using reconstruction variables and compared to a new-selection which utilises additional reconstruction information, such as the particle identification score, PIDA, [29, 271] of the largest track. The new selection also attempted to remove ambiguities caused by poor reconstruction using the residual analysis. An increase in efficiency and background rejection was observed using the new selection compared to the proposal selection.

Further reconstruction metrics which weakly classified CC  $\nu_e$  and  $\nu_\mu$ , primarily by identifying electron showers and photons showers, were then used in a BDT [281], to improve the selection. Two BDTs were trained: the standard and scaled. The scaled BDT included the scaling from the POT and oscillation normalisation which mimics oscillations from the best fit global sterile parameters [39]. The BDTs improved the efficiency of the selection compared to the new cut-based selection.

Each selection is set to maximise either the efficiency  $\times$  background rejection or the efficiency  $\times$  purity of the individual cuts. It should be noted that this may not result in the best selection for the oscillation analysis as statistical fluctuations may become significant for specific energy ranges if the efficiency is too low. The selected events for all the reconstruction selections are used in the sensitivity analysis in Chapter 8.

The analysis only includes neutrinos without a cosmic overlay and a light simulation. When the simulation includes both of these, the analysis should be extended to evaluate the effects on the reconstruction efficiency of neutrino events. The identification of cosmic events and the efficiency of correctly tagging events using the photon detection system should also be considered.

# Chapter 8

## SBN Sensitivity to Sterile $\nu_e$ Appearance

One of the primary goals of the Short-Baseline Neutrino (SBN) programme is to confirm or refute the anomalous electron neutrino excesses observed by the LSND [39] and MiniBooNE [143] experiments, see Chapter 4. A sensitivity study using selected electron neutrino events is performed in order to evaluate the capabilities of the programme.

In the previous Chapter, several selections were performed on truth and reconstructed information and each provided a different selection of events for analysis. In this Chapter, each selection of events will be analysed and compared in order to calculate and identify the best potential sensitivity to  $\nu_e$  appearance due to sterile neutrinos in the SBN programme. The analysis is based on the 3+1 theory described in Chapter 2, Section 2.2.5.

The sensitivity analysis was performed using the VALEncia-Oxford-Rutherford (VALOR) neutrino oscillation framework. The method used by VALOR to calculate the sensitivity is discussed in Section 8.1. At the heart of the VALOR oscillation analysis, Monte Carlo templates are compared to the observed flux using a binned log-likelihood ratio function across a joint detector fit.

The input spectra for the analysis are presented in Section 8.2. These are the resultant spectra from the various selections described in Chapter 7. The selections were performed on the samples discussed in Chapter 7, Section 7.1, which were reconstructed using methods described in Chapter 6. Therefore both the truth-based selections and the reconstructed selections were performed.

Various flux and interaction systematic errors are included within the VALOR analyses. The main process by which flux and interaction systematic errors are included is by reweighting the selection spectra to account for changes in the underlying physics and flux constants. The specific systematic errors applied will be presented in Section 8.3.

The  $\nu_e$  appearance sensitivity was initially calculated using the events that underwent the truth-based selection. This is described in Section 8.4. The  $\nu_e$  appearance sensitivity using

reconstructed data for the reconstruction selections in Chapter 7 was then calculated and is presented in Section 8.5.

## 8.1 The VALOR Framework for the SBN Analysis

The VALOR framework was initially established in the T2K framework. Additionally, VALOR has been used in the Deep Underground Neutrino Experiment (DUNE) CP-violation analyses. Therefore, VALOR has a sophisticated and established suite of oscillation physics fitting code for use in the SBN framework. Currently, only the 3+1 sterile model with the short-baseline approximation has been used in the SBN analysis. However, the software is capable of performing a fit to the full 3+2 theory and 3+3 theory as well as having the potential to include the matter effects [291].

VALOR uses Monte Carlo (MC) templates,  $T$ , to control the data given to the framework. Separate MC templates are created for each detector ( $d$ ), beam ( $b$ ) and sample ( $s$ ) used in the analysis, e.g. SBND and intrinsic  $\nu_e$ . Furthermore, events are placed within separate MC templates depending on their interaction mode ( $m$ ), e.g. quasi-elastic. Within the MC template, events are placed within a reconstructed ( $K_r$ ) and true ( $K_t$ ) binning scheme such that a mapping between true and reconstructed information can be obtained. Therefore MC templates can be described by

$$T = T_{d;b;s;m}(r, t), \quad (8.1)$$

where  $r$  and  $t$  are bins within the kinematic spaces  $K_r$  and  $K_t$  respectively. For the SBN programme, the reconstructed energy of the neutrino is currently used as the fitting space. Using the definition of the MC template, the predicted number of events,  $n_{d;s;b}^{pred}(r, \theta, \mathbf{f})$ , for a particular detector and sample with physical parameters  $\theta = (\theta_0, \theta_1, \dots)$  and nuisance parameters  $\mathbf{f} = (f_0, f_1, \dots)$  can be calculated as

$$n_{d;s}^{pred}(r, \theta, \mathbf{f}) = \sum_m \sum_t P_{d;m}(t, \theta) \cdot R_{d;b;s;m}(r, t; \mathbf{f}) \cdot T_{d;b;s;m}(r, t), \quad (8.2)$$

where  $P_{d;m}(t, \theta)$  denotes the physics hypothesis and therefore applies a weighting to events dependent on the 3+1 model.

An oscillation scaling is applied to each true energy bin based on the average oscillation probability for the events in the bin for each mode. The individual oscillation probability for

each event is calculated as

$$P_{\nu_\mu \rightarrow \nu_e}^{3+1} = \sin^2 \theta_{\mu e} \sin^2 \left( \frac{K \Delta m_{41}^2 (\text{eV}^2) L (\text{km})}{E_\nu (\text{GeV})} \right), \quad P_{\nu_\mu \rightarrow \nu_\mu}^{3+1} = 1 - \sin^2 \theta_{\mu \mu} \sin^2 \left( \frac{K \Delta m_{41}^2 (\text{eV}^2) L (\text{km})}{E_\nu (\text{GeV})} \right), \quad (8.3)$$

where  $L$  is the baseline of the individual neutrino,  $E$  is the true energy and  $K = 1.26693281 \text{ GeV eV}^{-2} \text{ km}^{-1}$ .

For the  $\nu_e$  appearance sensitivity analysis,  $\Delta m_{41}^2$  and  $\theta_{\mu e}$  are fixed at various points in the sensitive region and  $\theta_{\mu \mu}$  is a nuisance parameter. Every  $\nu_\mu$  is given a weight corresponding to  $P_{\nu_\mu \rightarrow \nu_\mu}^{3+1}$  and every oscillated  $\nu_e$  is given a weight corresponding to  $P_{\nu_\mu \rightarrow \nu_e}^{3+1}$ . Neutral current (NC) and intrinsic  $\nu_e$  are not oscillated and contributions for  $\nu_\tau$  are neglected. Not oscillating the NC events assumes that the kinematics and the rates of NC interactions of the different neutrino flavours are identical. As the NC interactions are not oscillated, NC events arising from the oscillated sample are removed in order to avoid double counting. Oscillations of the intrinsic  $\nu_e$  and  $\nu_\mu \rightarrow \nu_\tau$  are assumed negligible.  $R_{d;b;s;m}(r, t; (f))$  parameters denote the effect of the systematic variations. The responses are treated as linear in the analysis and thus are treated independently with  $R_{d;b;s;m}^i(r, t_i, f_i) \propto f_i$ . Therefore, the responses are factorised.

A binned likelihood ratio method is used to minimise the fit to data [292–294],

$$-2 \ln \lambda(\theta, \mathbf{f}) = 2 \sum_{d;b;s;r} \left( n_{d;b;s}^{obs}(r) \ln \frac{n_{d;b;s}^{obs}(r)}{n_{d;b;s}^{pred}(r, \theta, \mathbf{f}) + (n_{d;b;s}^{pred}(r, \theta, \mathbf{f}) - n_{d;b;s}^{obs}(r))} \right) - 2 \ln \lambda_{penalty}, \quad (8.4)$$

where  $n_{d;b;s}^{obs}(r)$  is the number of events observed for the detector  $d$  in the beamline  $b$ , for the sample of events  $s$  in the reconstructed energy bin  $r$ .  $2 \ln \lambda_{penalty}$  is the penalty term which arises due to the uncertainty on nuisance parameters and can be described in terms of the systematic variables  $f$  as

$$-2 \ln \lambda_{penalty} = \left( (\mathbf{f} - \mathbf{f}_0) C_f^{-1} (\mathbf{f} - \mathbf{f}_0) \right), \quad (8.5)$$

where  $\mathbf{f}_0$  is the vector of the nominal systematic parameters. Initially, the systematic parameters  $\mathbf{f}_0$  are set to 1 so that  $n^{pred}$  is not altered.  $C_f$  is the fractional covariance matrix of the systematic parameters.  $C_f$  is calculated from the uncertainty in the selected event spectra from multiple simulated Universes with different underlying systematic parameters. More detail on this topic can be found in Section 8.3. The fit statistic  $-2 \ln \lambda(\theta, \mathbf{f})$  has a  $\chi^2$  distribution, thus can be used as a goodness-of-fit test [294, 295]. For the sensitivity analysis, the  $n^{pred}$  corresponds to the data expected for the nominal physics hypothesis of no sterile

oscillations, whilst the observed data corresponds to the weighted Monte Carlo prediction obtained from oscillation probabilities 8.3.

## 8.2 The $\nu_e$ CC Selection Spectra

For the truth-based analyses, the same binning scheme that was applied in the proposal was used to enable a direct comparison to the proposal. The spectra are depicted in Figure 8.1 for the global best fit oscillation parameters  $\sin^2(\theta \nu_e) = 0.001$  and  $\Delta m^2 = 1.32 \text{ eV}^2$  [221]. The spectra for the GENIE weighted truth sample is found in Appendix J.

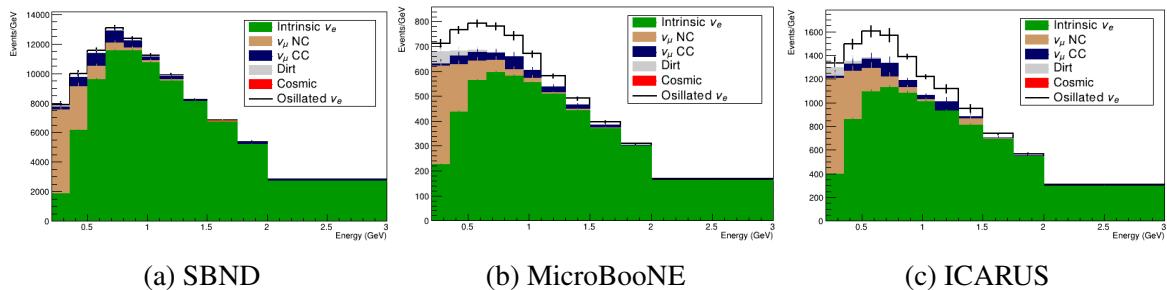


Fig. 8.1 The spectra from the three detectors for the truth-based selection. The selection is performed on the modern-like sample described in Chapter 7, Section 7.1.1. The errors are statistical.

For the reconstruction analysis, the binning scheme was altered. This was carried out because of the drop in efficiency in the reconstruction. The bins were chosen such that the statistical error for each bin was around 10% or less. The bin widths were also set to be large enough to account for the neutrino energy resolution, see Figure 8.2. Hence, a bin width at energy  $E$  with energy resolution  $x$  ranges from  $E - Ex \rightarrow E + Ex$ . This was done to mitigate bin migration. Note that the neutrino energy was discussed in Chapter 7, Section 7.4.3. The effects of changing the binning on the sensitivity is shown in Figure 8.6. As can be seen in the Figure, the sensitivity is significantly reduced. However, the difference is enhanced due to the inclusion of the Meson Exchange Current (MEC) systematic errors, see Section 8.3 for a discussion on the systematic errors.

Figure 8.3 show the spectra in each of the detectors using the proposal style reconstruction selection in Chapter 7 for the global best fit oscillation parameters  $\sin^2(\theta \nu_e) = 0.001$  and  $\Delta m^2 = 1.32 \text{ eV}^2$  [221]. The spectra for the other selections can be found in Appendix J. Note that as the dirt and cosmic backgrounds were not reconstructed and instead the spectra from the modern truth-based selection for these backgrounds are included in the analysis.

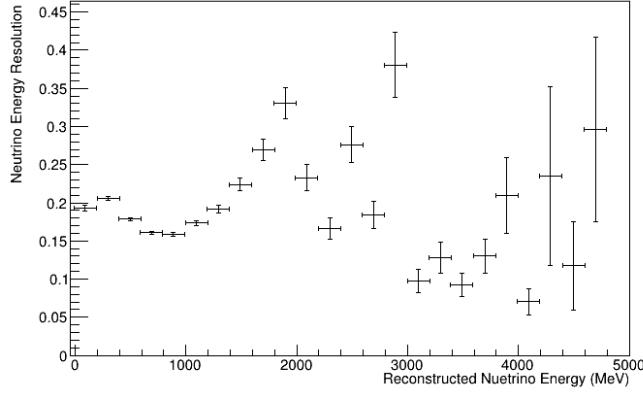


Fig. 8.2 The neutrino energy resolution for Charged Current (CC)  $\nu_e$  events after the pre-selection cuts in Chapter 7, Section 7.4.3. The analysis is performed on the Short-Baseline Near Detector (SBND) Booster Neutrino Beam (BNB) events.

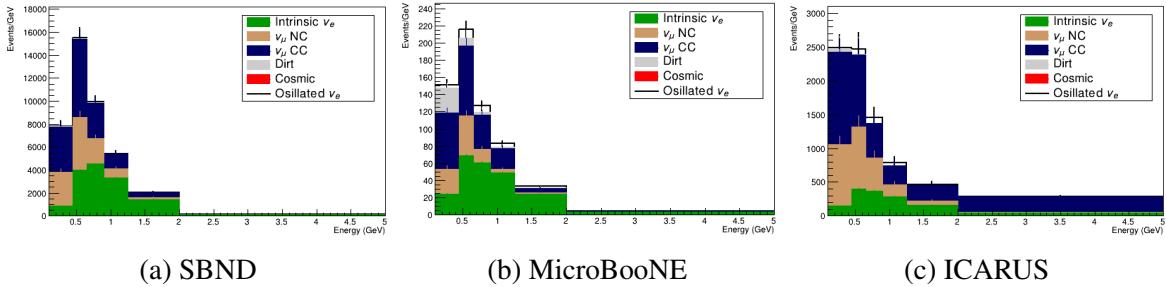


Fig. 8.3 The spectra from the three detectors for the proposal-like selection with the parameters discussed in Chapter 7, Section 7.4.3 set to maximise the efficiency  $\times$  background rejection. The spectra are created using the reconstruction information of the Time Projection Chamber (TPC) signal and backgrounds and the truth-based dirt and cosmic background. The errors are statistical.

All the spectral Figures show the peak BNB energy as approximately 700 MeV which is expected, see Chapter 4 for a description of the BNB. The  $\nu_\mu$  background is larger in the reconstructed selection compared to the truth-based selections. This is expected from the drop in background rejection presented in Chapter 7. This is due to the reduction in effectiveness of the cuts, particularly the  $dE/dx$  cut. Furthermore, due to the maximum track length cut, the  $\nu_\mu$  CC background is expected to be smaller than the NC background. This is true in the truth-based analysis; however, mischaracterisation in the reconstruction causes additional CC background events that are not accounted for in the truth-based analysis.

### 8.3 Flux and Interaction Systematics

As mentioned in the previous Section, systematic uncertainty in the analysis is accounted for via the  $\lambda_{penalty}$  term. The  $\lambda_{penalty}$  term encapsulates the uncertainty within the fractional covariance  $C_f$  which is formed from the linear combination of the flux ( $C_f$ ), interaction ( $C_i$ ) and dirt ( $C_{dirt}$ ) systematic errors:

$$C_f = C_f + C_i + C_{dirt}. \quad (8.6)$$

In the future, additional detector systematic errors, such as an error on the electron lifetime, will be considered in the analysis.

A "unisim" method was used to construct the flux and interaction covariance matrices [296]. The process involves generating many MC events for one event [297]. The new events, or "Universe", are created by reweighting the central value (CV) that arises from the best estimate of the underlying physics parameters. The reweighting accounts for randomly assigning new values to physics parameters, within the uncertainties specified in Tables 8.1 and 8.2, and producing new weights to account for these changes. The weights that are created for each systematic parameter are then multiplied together to form the total weight of the new "Universe". The weight is calculated for each event used in the matrix. The covariance matrix is then constructed from the resultant  $\mathcal{N}$  "Universes" by evaluating the changes in the reconstructed energy spectra after the selection,

$$c_{ij} = 1/\mathcal{N} \sum_m^{\mathcal{N}} \left[ N_{CV}^i - N_m^j \right] \times \left[ N_{CV}^j - N_m^i \right], \quad (8.7)$$

where  $c_{ij}$  is the covariance between the reconstructed neutrino energy bin  $i$  and bin  $j$ . The fractional covariance matrix is used in the analysis and this corresponds to the result of dividing the covariance matrices by the central values,

$$C_{ij} = \frac{c_{ij}}{N_{CV}^i N_{CV}^j}. \quad (8.8)$$

In the analysis, the mean value of the "Universes" is taken as the CV rather than the original MC.

The GENIE reweight package [225] was used to vary the interaction systematic errors displayed in Table 8.1. The MicroBooNE flux reweight package was used to vary the flux systematic errors displayed in Table 8.2. 1000 "Universes" were created for each event to construct the covariance matrices for each experiment. For the flux systematic errors, the covariance matrices were constructed for each energy bin for each detector to account for

correlations in the flux between the three experiments. The interaction covariance matrix is constructed by averaging the reconstructed energy spectra from the three detectors. Note that for selections which maximise the efficiency  $\times$  purity, no TPC events exist in the first reconstruction energy bin. Therefore, the first rows of the flux covariance matrices are empty.

There are two systematic errors which are treated differently to the above method. A 15% uncertainty on the dirt normalisation and a 100% uncertainty on the MEC normalisation are applied outside of the event reweight framework. The uncertainty is accounted for by scaling the total number of events of the interaction type (dirt/MEC) by a random number for each "Universe". A Gaussian with a mean of 1 and RMS corresponding to the uncertainty is used to generate the random number. This is performed independently for each detector. The uncertainty is then absorbed in the interaction matrix by multiplying the scale factor by the average weight of the events in each reconstructed energy bin from the "unisim" method. The matrix is then constructed as above.

Description	$1\sigma$ Uncertainty
Axial mass for CC quasi-elastic	-15% + 25%
Axial mass for CC resonance neutrino production	$\pm 20\%$
Axial mass for NC resonance neutrino production	$\pm 20\%$
Non-resonance background in $\nu p$ CC/NC $1\pi$ reaction	$\pm 50\%$
Non-resonance background in $\nu p$ CC/NC $2\pi$ reaction	$\pm 50\%$
Non-resonance background in $\nu n$ CC/NC $1\pi$ reaction	$\pm 50\%$
Non-resonance background in $\nu n$ CC/NC $2\pi$ reaction	$\pm 50\%$
Neutral Current Normalisation	$\pm 20\%$
MEC Normalisation	$\pm 100\%$
Dirt Normalisation	$\pm 15\%$

Table 8.1 The interaction systemic errors used in the analysis and the  $1\sigma$  fractional uncertainties used when creating the "Universes". Below the line, the event reweighting GENIE package is not used.

Description	$\nu_\mu$	$\nu_e$
$\pi^+$ production	14.7%	9.3%
$K^+$ production	0.9%	11.5%
$K^0$ production	0.0%	2.1%
Horn Field	2.2%	0.6%
Nucleon cross sections	2.8%	3.3%
Pion cross sections	1.2%	0.8%

Table 8.2 The flux systemic errors used in the analysis and the  $1\sigma$  fractional uncertainties used when creating the "Universes".

Figures K.2 show the flux and interaction covariance matrices for the "modern" truth selection described in Chapter 7, Section 7.1.1. The other truth-based selections are presented in Appendix K. The error in the higher energy  $\nu_\mu$  bins of the flux covariance matrices is due to the uncertainty in the  $\pi^+$  production. This uncertainty is also responsible for the larger uncertainty in the lower energy bins of the  $\nu_\mu$  and  $\nu_e$  flux.

The covariance matrices used for the reconstruction proposal style analysis which maximise the efficiency  $\times$  the background rejection are shown in Figures K.4. The other selection covariance matrices are presented in Appendix K. The matrices are similar for all of the selections. Due to the change in binning the correlations due to the MEC normalisation are more convolved with the resonant uncertainties. Also, differences occur due to the energy cut-off of the selections that maximise efficiency  $\times$  purity. The cut-off results in having only dirt events in the first energy bins for the interaction matrices and no events in the flux matrices.

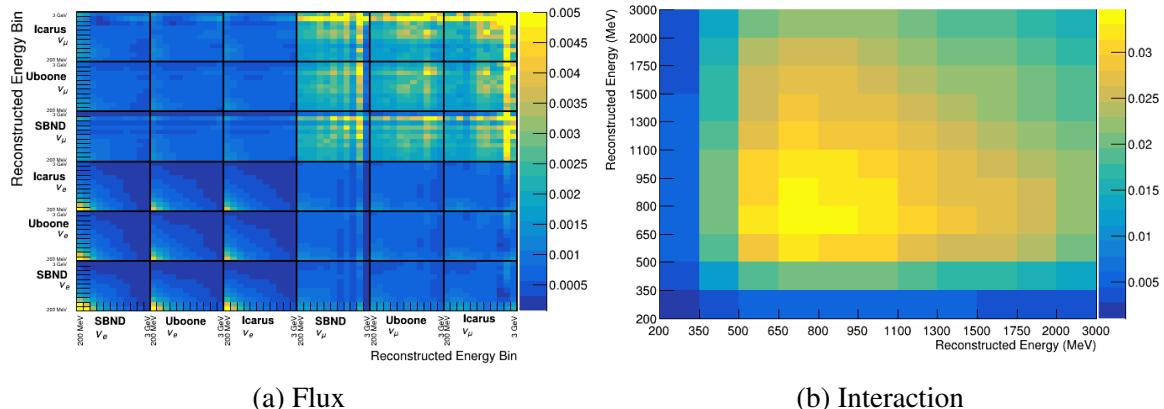


Fig. 8.4 The fractional covariance matrices for the flux (a) and interaction (b) for the truth-based modern selection discussed in Chapter 7, Section 7.1.1. The matrices are created from spectra, Figure 8.1.

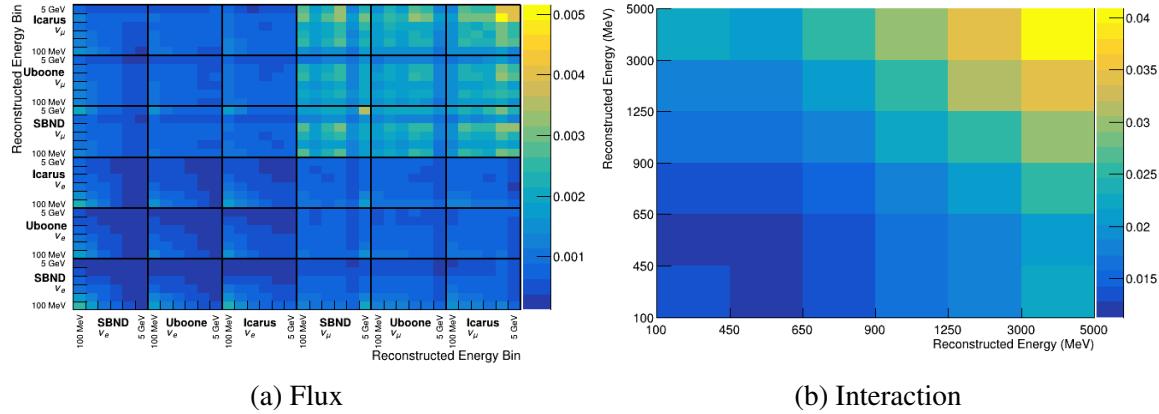


Fig. 8.5 The fractional covariance matrices for the flux (a) and interaction (b) for the reconstruction selection discussed in Chapter 7, Section 7.4.3. The matrices are created from spectra, Figure 8.3, which uses the proposal selection, maximising the efficiency  $\times$  background rejection. The new binning scheme has been used.

## 8.4 Results of the Truth-Based Analyses

Figure 8.6 shows the 90% confidence level sensitivities for  $\nu_e$  appearance and is produced by carrying out the following: performing the truth-based selection, the details of which are discussed in Chapter 7, which results in the spectra shown in Appendix J, these are discussed in Section 8.2; then performing the VALOR analysis described in Section 8.1 with the covariance matrices that can be seen in Appendix K. The figures are made by evaluating the  $\chi^2$  fit statistic when fitting to the null hypothesis of no oscillations for the selected data sample. The oscillation signal size of the data sample depends on the given values of the oscillation parameters at the specific points in the plot. The  $\chi^2$  value then indicates the probability of obtaining the oscillated spectra given the null hypothesis. This is due the relation between the  $\chi^2$  value and the p-value. The figures exclude regions to the right of the contour by the specified confidence level.

The results show that both of the 90% confidence level sensitivities are comparable to the proposal sensitivity contour performed in 2012. Figure 8.6 also shows that there is more resolving power in the "modern" sample, discussed in Chapter 7, Section 7.1.1, compared to the "proposal" sample. This difference is as a result of the increase in the signal size due to the presence of MEC events. The Figure also shows the effects of decreasing the binning. Note that the significant loss in sensitivity is due to the combined effects of the systematic errors and the binning.

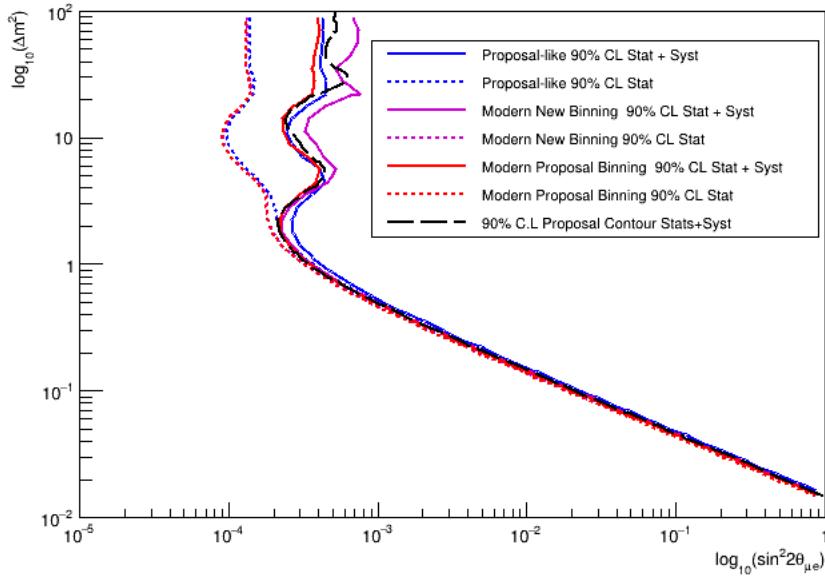


Fig. 8.6 The  $\nu_e$  appearance sensitivity for the truth-based analyses. The modern (red) and proposal-like (blue) samples are shown as well as the modern sample with the new binning scheme (purple). The sensitivities without the systematic penalty are shown via the dashed lines. The 2012 proposal sensitivity (black dashed) is shown for comparison.

## 8.5 Results of the TPC Cut Based Reconstruction Analysis

The 90% confidence level sensitivities for the  $\nu_e$  appearance using various reconstruction selections is presented in Figures 8.7 and 8.8. Like the truth-based analyses, the analysis discussed in Section 8.1 is used. The covariance matrices, displayed in Appendix K, are created for each analysis from the post-selection spectra, displayed in Appendix J. For the cut-based selections, it can be seen in Figure 8.7 that the area enclosed by the sensitivity contour is higher when maximising the efficiency  $\times$  background rejection for each of the cuts performed in the analysis. For the BDT analyses, various selections with different BDT score thresholds were performed. The area enclosed by the resultant sensitivity contour can be seen in Figure 8.8. For the BDT where the signal and background events are scaled to 1:1, a score of 0.2 provides the largest sensitivity region. Using the BDT where the signal and background events are scaled with respect to the POT and the oscillation weights, discussed in Chapter 7, the best cut is at -0.99. The selection with the largest contour for the systemic errors used in Tables 8.1 and 8.2 is the the 1:1 BDT with a score of 0.2.

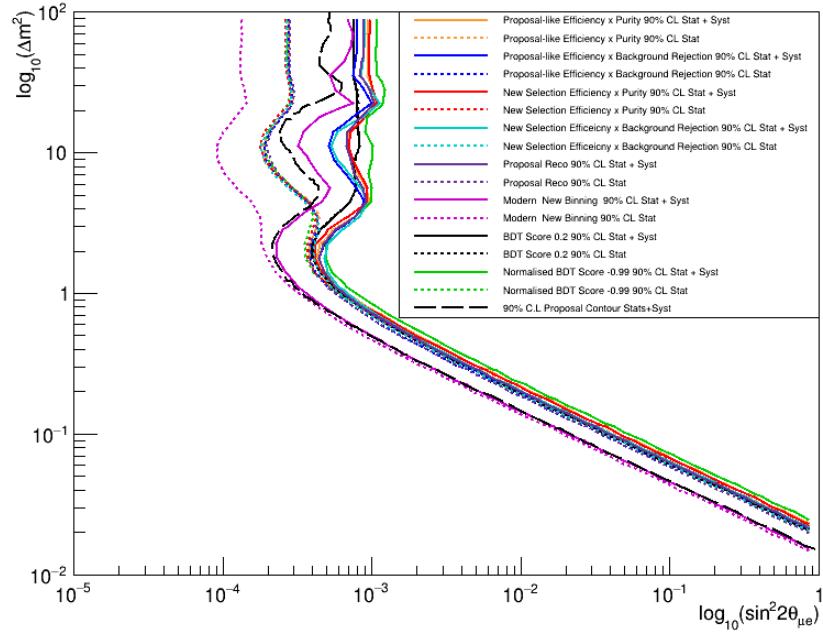


Fig. 8.7 The 90% CL  $\nu_e$  appearance sensitivity for the reconstruction analyses. The sensitivity without the systematic penalty is shown via the dashed lines and the systematic errors with the solid line. The 2012 proposal sensitivity (black dashed) and the modern truth-based analysis with the new binning scheme (purple line) is shown for comparison.

The area enclosed by the reconstructed contours is less than the modern truth-based sensitivity with the same binning. This is because the reconstruction efficiency is over-estimated and the truth-based selection does not fully consider the selection inefficiencies, see Chapter 7. Therefore, the reconstruction contours are a more realistic estimation of the capabilities of the SBN programme with the current software reconstruction in place. However, improvements in both the reconstruction efficiency and selection are still possible and therefore the truth-based selection provides a benchmark to the possible capabilities of the programme.

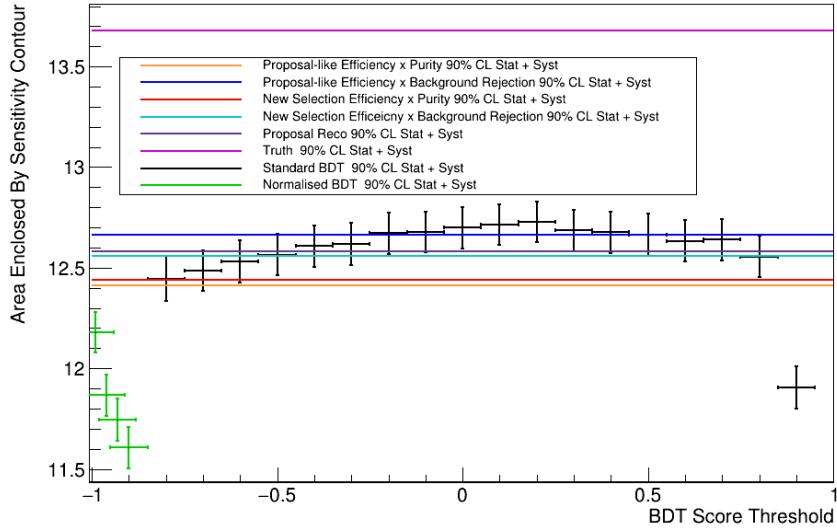


Fig. 8.8 The area covered by the 90%  $\nu_e$  appearance sensitivity for the reconstruction analyses as a function of the BDT score used. The cut-based methods are shown as a horizontal line as they do not depend on the BDT score. The  $x$  errors for each BDT area are the distance to the next point and the  $y$  errors are calculated by propagating the size of the steps used to form the sensitivity region around the contour. The modern truth-based analysis (purple line) with the new binning scheme is shown for comparison.

Figure 8.9 shows the current selection overlaid with the MiniBooNE 99% CL sensitivity and the LSND allowed regions. The Figure shows that the truth-based analysis covers most of the allowed region at  $5\sigma$  and the BDT selection covers most of the region at  $3\sigma$ . This implies that, with the current status of the reconstruction, the SBN programme can strongly imply the existence, or lack thereof, of sterile neutrinos.<sup>5</sup> Therefore, additional works should be carried out to improve the reconstruction to reduce the differences in the reconstruction and truth analyses. Furthermore additional selection procedures, should be investigated to improve the performance of the analysis. It should be noted that further systematic errors, such as ones which account for detector effect, are still required in the analysis. Also, the events are reconstructed without the light simulation or cosmic overlay, and therefore inefficiencies in the reconstruction due to triggering and cosmic slicing are not considered. Furthermore, the cosmic and dirt background spectra are not reconstructed and the truth selection is still applied. When these effects are incorporated into the analysis, the sensitivity region is expected to decrease and efforts should be made to include these and reevaluate these findings.

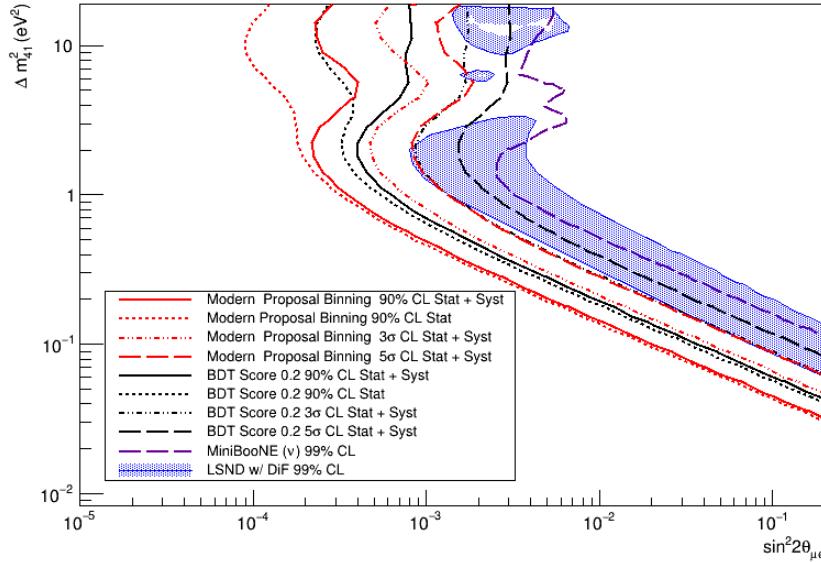


Fig. 8.9 The  $\nu_e$  appearance sensitivity for the truth-based analyses with the proposal binning (red) and the BDT reconstruction selection with a score of 0.2 (black). The 90%,  $3\sigma$  and  $5\sigma$  are presented for each analysis. The LSND allowed regions and the MiniBooNE sensitivity is also plotted to show the SBN coverage of the low energy excess.

The current significance to an excess of neutrinos in LSND is  $3.8\sigma$  [9] and in MiniBooNE is  $4.7\sigma$  [10] which combine to excess of  $6.0\sigma$  [10]. Although the significance presented in this work is lower than this value it should be noted that due to LArTPC technology the photon backgrounds can be distinguished from the  $\nu_e$  interactions. As this cannot be done in MiniBooNE and LSND it is thought that maybe the rates of the photon backgrounds in the MC are misunderstood and account for the excess. SBN can infer if this is case using the current reconstruction and selection methods and thus compliments these detectors. Furthermore, the  $\nu_e$  appearance analysis can be combined with a  $\nu_{\mu}\nu_e$  disappearance and  $\nu_e$  disappearance analyses to improve the sensitivity of the programme. Such analyses can be done using exclusive data samples as well which could also improve the sensitivity. Such combination analyses are currently being investigated.

## 8.6 Concluding Remarks

VALOR provides a sophisticated analysis framework to perform SBN analyses. Currently, SBN truth analyses exist to provide a benchmark to reconstruction and show how physics and detector changes have affected the sensitivity to  $\nu_e$  appearance since the SBN proposal. Flux and interaction systematic errors have been included in the analysis, using the GENIE

and MicroBooNE event weight frameworks. However, as the collaboration moves to GENIE v3, additional interaction systematic errors should be considered. Additionally, detector systematic errors need to be included and an effort to incorporate these is ongoing.

Simulated events contained in the TPC were reconstructed in all three detectors using the methods described in Chapter 6. These events underwent various selections, discussed in Chapter 7, and the sensitivities were compared using the VALOR fitting framework.

The best selection uses a BDT, discussed in Chapter 7, Section 7.4.5, to separate the  $\nu_e$  signal from the  $\nu_\mu$  background and therefore a BDT should be considered for future analyses. Currently, for the  $\nu_e$  excess, the LSND allowed region is almost completely covered within a  $3\sigma$  sensitivity contour of the BDT reconstruction. The simulated events used, however, do not have the light simulation incorporated within them or a cosmic overlay and are therefore over-estimations of the current sensitivity.

As there is a significant difference between the sensitivity of the truth-based analysis and the reconstruction analysis, improvements in the reconstruction should be made so that the capabilities of SBN programme can be reached.

# Chapter 9

## R&D In New Liquid Argon Readouts

As mentioned in Chapter 4, one of the goals of the Short-Baseline Near Detector (SBND) programme is to test new readout techniques in liquid argon for the Deep Underground Neutrino Experiment (DUNE) [115]. SBND is a single-phase wire Liquid Argon Time Projection Chamber (LArTPC) and is a prototype for the far detector of the DUNE experiment. At least one of the four DUNE far detectors will follow the ProtoDUNE single phase wire readout [180]. There is a concern that wire LArTPCs are not suitable for the DUNE near detector. This is due to the risk of pile up in the near detector. It is estimated that 0.2 interactions/tonne/spill [298] will occur at the near detector located 574 m downstream for the 2 MW beam [115]. If events occur at the same time, this introduces ambiguities into a wire readout [253, 254]. To alleviate the issue, pixelated readouts are being considered for the near detector where the extra degrees of freedom remove the ambiguities [299]. Pixelated detectors are also being considered for one of the far detector modules. Currently, different pixelated readouts are being developed and tested for use in the near detector. At the University of Sheffield, a liquid argon rig has been created as a testbed for research.

This Chapter describes the design concept of the Sheffield rig in Section 9.1 and the DAQ readout system implemented to develop liquid argon technologies in Section 4.3.2. Initially, a pixelated readout system, designed by the University of Bern, has been implemented into the rig to provide a LArTPC readout. Along with the TPC, other sub-detector cosmic paddles outside and a photomultiplier tube (PMT) inside the vessel have been installed, mainly for triggering. The sub-detectors are discussed in Section 9.3. Finally, software used by the DAQ will be discussed in Section 9.4 and a noise analysis of the TPC data will presented in Section 9.5.

## 9.1 The Sheffield Liquid Argon Rig Design

The Sheffield test stand is based on a 20 L cryogenic test stand designed by BNL [300]. A schematic of the test stand can be seen in Figure 9.1 and a photograph can be seen in Figure 9.2. The cryogenic system consists of a 20 L cylindrical dewar with a depth of 60.96 cm and a diameter of 24.02 cm [300]. A capacitor level gauge and a PT-100 RTD resistive temperature sensor are positioned within the dewar to monitor the pressure and temperature respectively. An inlet at the bottom of the dewar exists for filling and evacuating. The dewar is attached to a CF flange (which uses a copper gasket to seal the dewar and form a vacuum) with several feed-throughs for the purification system and the electronics. Four steel baffle plates are held approximately 2 cm apart at the top of the dewar to provide thermal insulation for the flange. The baffle plates are held in place by three threaded steel rods which extend to the bottom of the dewar and provide a support structure for the detectors in the dewar.

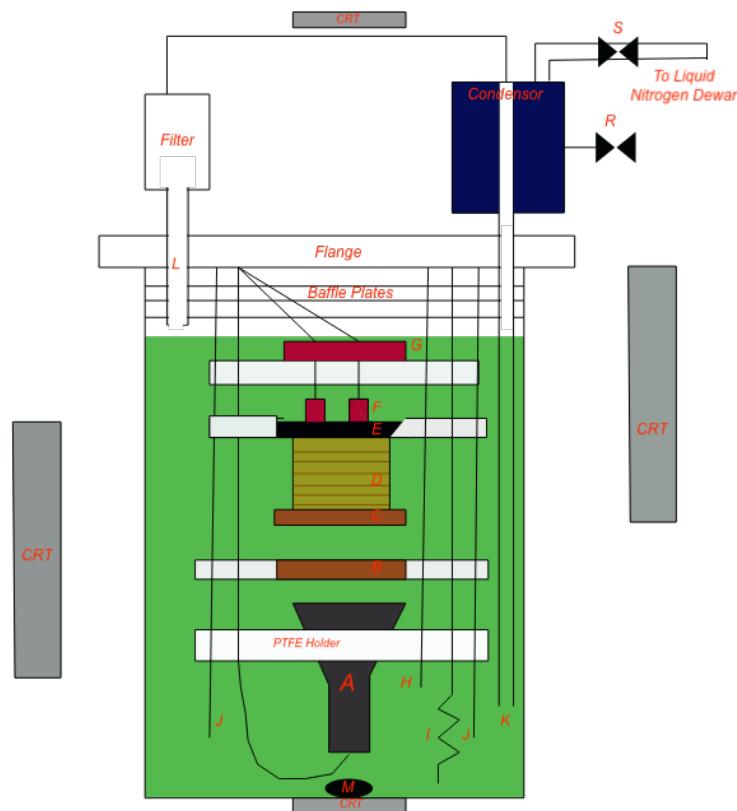


Fig. 9.1 The Sheffield liquid argon rig. A) The PMT, B) the protector mesh, C) mesh cathode, D) TPC field cage, E) pixel plane, F) LArASIC chip, G) adapter board, H) argon capacitor level gauge, I) PT-100 RTD resistive temperature sensor, J) threaded struts, K) purifier output, L) purifier input, and M) argon input.

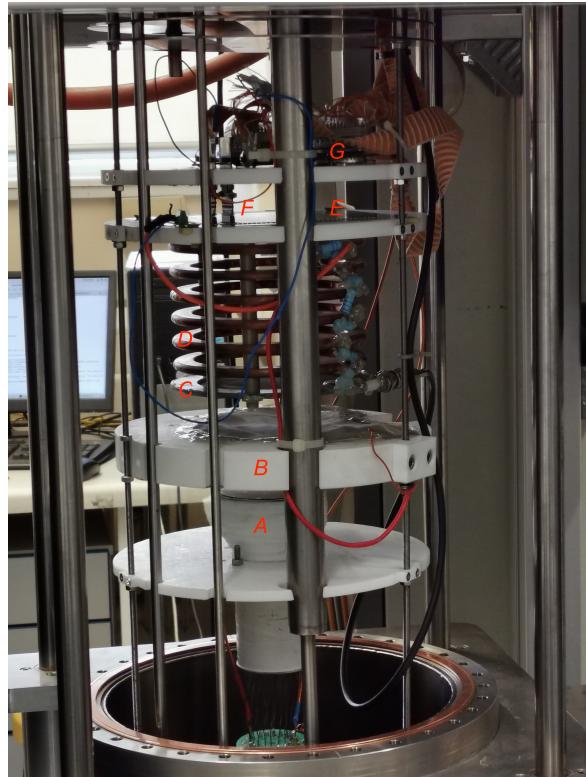


Fig. 9.2 The Sheffield liquid argon rig. A) The PMT, B) the protector mesh, C) mesh cathode, D) TPC field cage, E) pixel plane, F) LArASIC chip, and G) adapter board.

The dewar is attached to the purification system via two vacuum jacket tubes. One tube is for the incoming flow of argon gas, which occurs due to evaporation, into the purification system from the top of the dewar. The other tube is for the outgoing flow of purified argon which extends into the liquid. The incoming gas is purified by 13X molecular sieve mesh beads and a GetterMax-133 [301] copper catalyst which removes water and oxygen respectively [300]. The purified argon gas is then passed through a condenser which consists of a coaxial cylinder (outer dimensions of  $6\text{ d} \times 18.5\text{ l}$  inch) containing liquid nitrogen. An inner 2 inch cylinder containing copper wool allows for the argon to pass through. Heat is then exchanged from the argon gas to the liquid nitrogen efficiently via the wool resulting in liquid argon. The resultant liquid returns to the dewar due to gravity [300]. A back-pressure regulator maintains the pressure of the liquid nitrogen and, in turn, maintains the temperature of the nitrogen. The regulator, therefore, maintains the pressure and temperature of the argon dewar. The liquid nitrogen level is also monitored and maintained with a solenoid valve. Pressure relief valves are placed on the flange and the condensing system to prevent the pressure exceeding 8 psi. Outlets to a Servomex DF550E oxygen monitor [302] are placed in the purification pipeline to measure changes in purity.

The temperatures, liquid levels and pressures are monitored by a National Instruments (NI) SCB-68A [303] which passes the data to a NI PXI-6229 I/O card [304]. The data is then transferred to a PC via a NI PCI-8331 [305] card. The data acquisition is controlled by dedicated LABVIEW [306] software which monitors and displays the slow control data. The software also interacts with the solenoid valve.

Inside the dewar, as can be seen in Figure 9.2, a LArTPC was installed to test new readout techniques. The TPC drift field is in the vertical direction due to the geometry of the vessel. The design and the electronics of the TPC will be discussed in the next Section. Also, as can be seen in Figure 9.2, a PMT is located behind the cathode of the TPC in order to detect scintillation for triggering purposes. The PMT faces the cathode as it is a wire mesh and is thus transparent, unlike the anode. As can be seen in Figure 9.1, 4 additional sub-detectors, which are cosmic ray scintillator paddles, are positioned on the outside of the detector. These are used for triggering and to identify cosmic events which pass through the TPC. The sub-detectors will be discussed in Section 9.3.

## 9.2 TPC Design and Electronic Readout

A cylindrical LArTPC was created for the Sheffield rig. As can be seen in Figures 9.1 and 9.2, the TPC drift direction is situated in the vertical direction with a 10 cm drift and a 10 cm diameter.

The anode consists of a pixelated readout, designed by the University of Bern, with a total of 1008 pixels [33, 307]. The  $900 \mu m$  pixels have a pitch of 2.54 mm and are separated into 28  $6 \times 6$  squares, known as Regions Of Interests (ROIs) [33]. As can be seen in Figure 9.3, each ROI has an induction grid which surrounds the pixels. The grid is biased such that the readout charge is focused onto the pixels. The voltage for the induction grid is supplied via the field cage using a  $56 M\Omega$  resistor.

To reduce the number of readout channels required, multiplexing is used. There are 36 pixel channels (one for each pixel in the ROI square) and 28 ROI channels. Combining the signal induced on an ROI channel and the signal on a collection pixel channel allows identification of the pixel which detected the charge. The multiplexing reduces the channels needed from  $N \times N$  to a manageable  $2N$ .

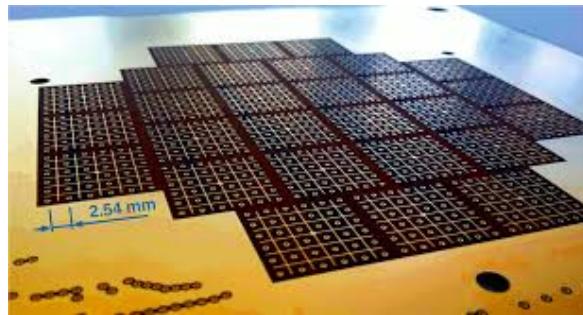


Fig. 9.3 The face of the Bern pixel readout. From Reference [33].

Figure 9.6 shows the electronic readout chain. Once there is a signal on the pixel board it is then passed to an electronics chip which holds a LArASIC7 [208] to amplify and shape the signal within the liquid argon (see Chapter 4, Section 4.3.2 for more information on LArASIC7). The LArASIC7 shaping time and gain settings are adjusted using an Arduino nano [308], which sends bit-wise signals to the ASIC via ribbon cables. The system is designed to run with a shaping time of  $2 \mu\text{s}$  and gain of  $25 \text{ mV/fC}$  for the LArASIC7. The processed pixel signal is then passed to an adapter board which passes the channels to the feedthrough via kapton ribbon cables. The pixel signals are passed to a V1740 CAEN DAQ [309] via two ribbon cables. The readout has been tested by sending a square test pulse from the Arduino to the LArASIC7, the response of which is shown in Figure 9.4. A 1 V square pulse has also been sent through a wire in air to a region below the readout to test the readout. The response is shown in Figure 9.5.

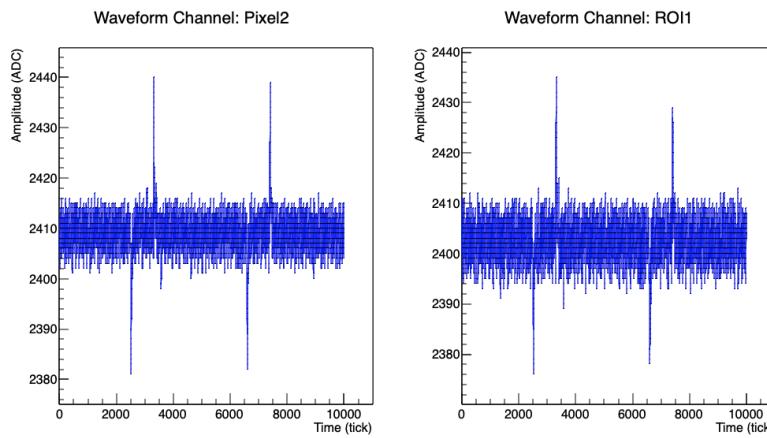


Fig. 9.4 The response to a test pulse being supplied to the ASIC chip via the Arduino nano.

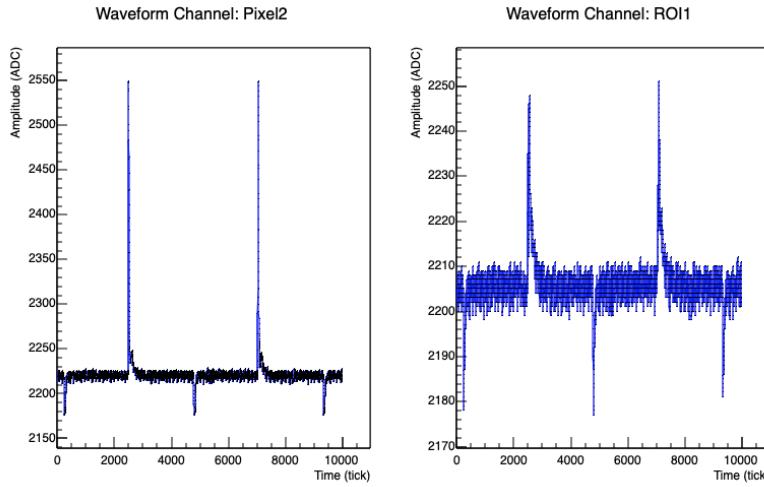


Fig. 9.5 The response to a test pulse being supplied to the ASIC chip via a charged wire.

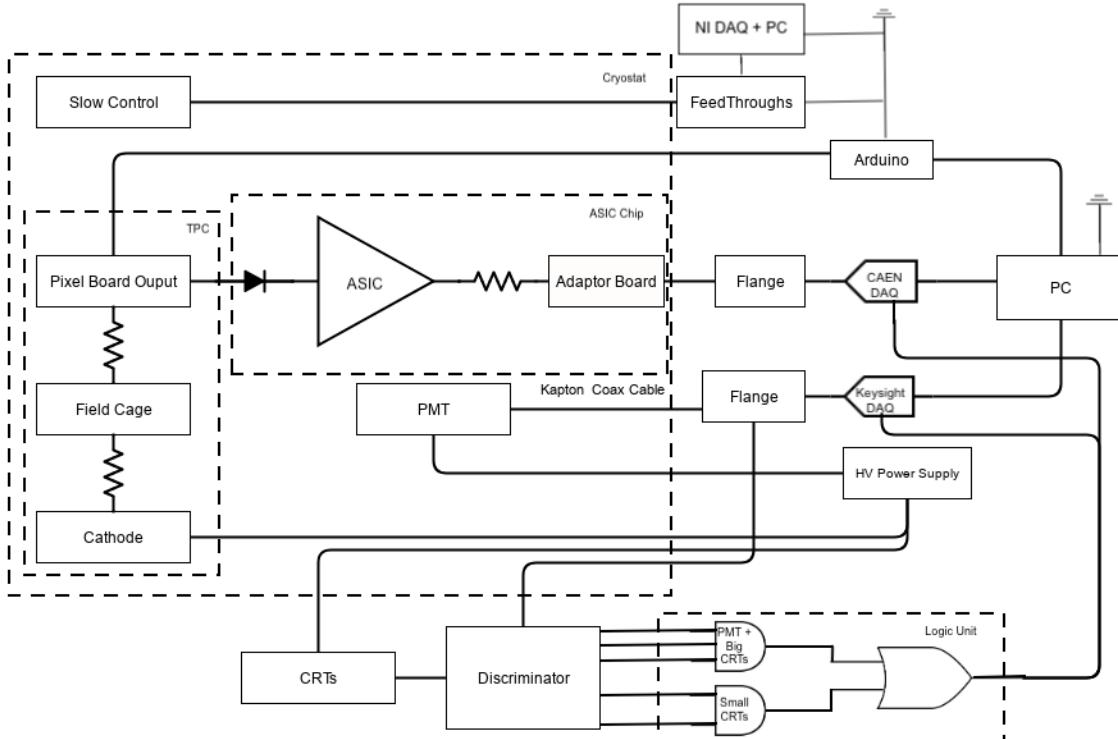


Fig. 9.6 A flow diagram of the electronics chain in the Sheffield rig.

The field cage of the TPC is made of six 6.35 mm thick copper rings evenly spaced and connected by a  $56\text{ M}\Omega$  resistor. Various field cage designs using the Ansys Maxwell [310] software were created. The cathode is made from a wire mesh to allow scintillation light to

pass through. The voltage is supplied to the cathode via a CAEN V6521 [311] using SHV cables outside and inside the vessel. The full setup can be seen in Figure 9.2. Successful tests of the field cage up to 5 kV have occurred in air with a stable current.

## 9.3 Sub-detector Setup and Electronic Readout

The sub-detector system consists of a PMT coated in TPB which faces the cathode within the LAR and 4 cosmic ray paddles (CRTs). The CRTs are positioned outside of the detector to identify cosmic events which pass through the TPCs. This Section will discuss the sub-detectors and their electronics setup.

### PMT

As is shown Figures 9.1 and 9.2 on the inside of the vessel a Hamamatsu R11065 3-inch PMT [312] faces the cathode to detect argon scintillation light. To protect the PMT field from the cathode field, an additional grounded mesh is placed between the cathode and the PMT. 1700 V is supplied to the PMT from a CAEN V6521 power supply [311] via an SHV cable outside the vessel and a coaxial kapton cable inside the vessel. The PMT output signal is passed through coaxial cables to a Keysight U5309A DAQ [313] for readout and a CAEN N417 discriminator [314] for triggering purposes via BNC cables.

Initial tests of the PMT have been implemented with various sources. Argon gas scintillation from an americium-241 alpha source and cosmic muons have been detected. Runs with a LED and a caesium-137 gamma source, when coupled to a sodium iodide crystal, have also been undertaken to provide further calibration data. The spectrum from the caesium-137 run is shown in Figure 9.7, where the gamma peak is visible. Additional analyses of the purity of gaseous argon using the PMT data have been undertaken within the rig setup. This was completed by evaluating the slow and fast scintillation components of argon.

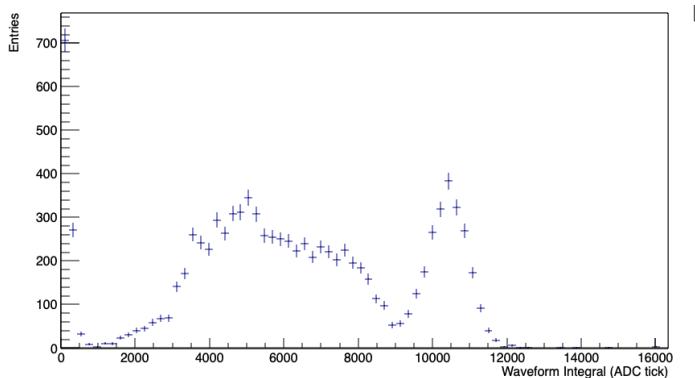


Fig. 9.7 The response of the Hamamatsu R11065 PMT to a Cs-137 source when coupled to sodium iodide crystal. A tick is 1 ns in the DAQ. Credit: Ed Tyley for redoing the analysis as the initial data was lost during PC hardware upgrades.

## CRT

Outside of the vessel 4 cosmic ray paddles (CRT) are installed. These are highlighted in Figure 9.2. The paddles' primary purpose is to provide a trigger for cosmic muons that traverse the TPC. The paddles are made by coupling scintillator plastic with PMTs.

The larger two paddles (50 cm by 50 cm) are positioned on the sides of the rig, 65 cm apart, such that the angle of cosmic muons which pass through the two paddles is less than  $52^\circ$  with respect to the pixel readout. This prevents events having a drift time in a pixel larger than the 2  $\mu$ s shaping time of the LArASIC amplifier, and hence also prevents saturation. The side paddles have been calibrated and shown to have a 5 ns offset between the signals, as is visible in Figure 9.8.

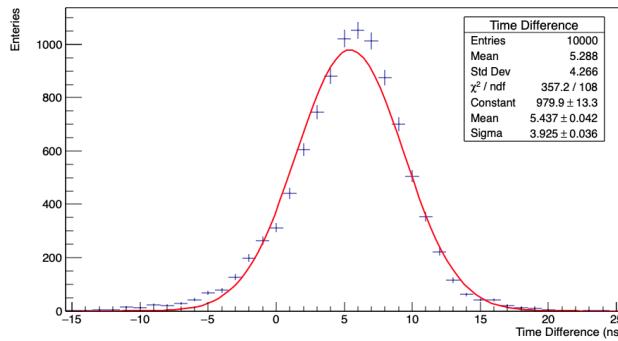


Fig. 9.8 The time distribution of the peak heights from one of the larger cosmic paddles when triggering with the other.

The other two paddles have an area of  $10 \times 10 \text{ cm}^2$  and are placed at the top and the bottom of the rig ( $\sim 2.5 \text{ m}$  apart). They are positioned such that if a vertical muon passes through both paddles, the muon must pass through the pixel readout. This ensures that ionisation depositions are detected by the readout regardless of the purity of the argon. This trigger pathway allows for the first test of the system during a run. Both sets of cosmic paddles trigger the detector in this setup.

### Electronics

The sub-detector electronics setup is depicted in Figure 9.6. The outputs of the cosmic ray paddles are passed to the Keysight U5309A DAQ for data readout and the N417 CAEN discriminator for triggering. The output of the PMT in the vessel and the two large CRTs are coupled together via AND logic using a CAEN N405 [315] logic module. This setup ensures the trigger is a result of a cosmic muon within the TPC. Similarly, the two small paddles are coupled using AND logic. The two triggers are then coupled and sent as a trigger to the CAEN 1740 DAQ and the Keysight U5309A DAQ for the acquisition of the pixel data and sub-detector data.

## 9.4 Data Acquisition and Readout Monitoring Software

A software infrastructure was also created so that acquisition and analysis of data can be easily maintained. The code basis is split into three sections: the Keysight data acquisition, the CAEN data acquisition and analysis software. Additional LABVIEW code is used to monitor the slow control data, such as the pressure and temperature of the liquids in the rig.

The data acquisition software saves the data in a binary format which is decoded by software offline. For each event, a header is stored which provides metadata, such as a timestamp for the event. The metadata allows for the sub-detector and TPC data to be linked together offline if required. Information is passed from the PC through an optical cable to a CAEN V1718 [316], which interacts via a VME Bus with CAEN 1740 DAQ. Dedicated software exists within the CAEN pixel DAQ system to initialise the Arduino via a C++ API. This API allows the Arduino to be easily configured with the same file as the DAQ configuration.

To monitor the running of the system, an online monitoring system has been developed. For the Keysight DAQ, system this consists of a waveform display and integration spectrum during runtime. A simple baseline calculator finds the average ADC count over the waveform before the triggered event time. This is then subtracted from the sum of the bins just before the pulse peak until an endpoint which is defined by the user.

For the CAEN DAQ, the software was designed to calculate the baseline, average pulse height and noise RMS. This information is then passed to an event display and a Postgres SQL [239] server for monitoring changes (via the libpq API [317]). The averages are then displayed as a function of time for the user to identify changes during runs. This is shown in Figure 9.9. The monitoring can be switched to an offline system if the interaction rate is higher for a specific experiment. For cosmic muon triggering the time required to run the online monitoring process is significantly less than the CPU time of the monitoring system.

There also exists a monitor for the CAEN V6521 high voltage and an email is sent out if the voltage trips. Finally, a webcam pointing towards the detector has been setup to monitor major catastrophes.

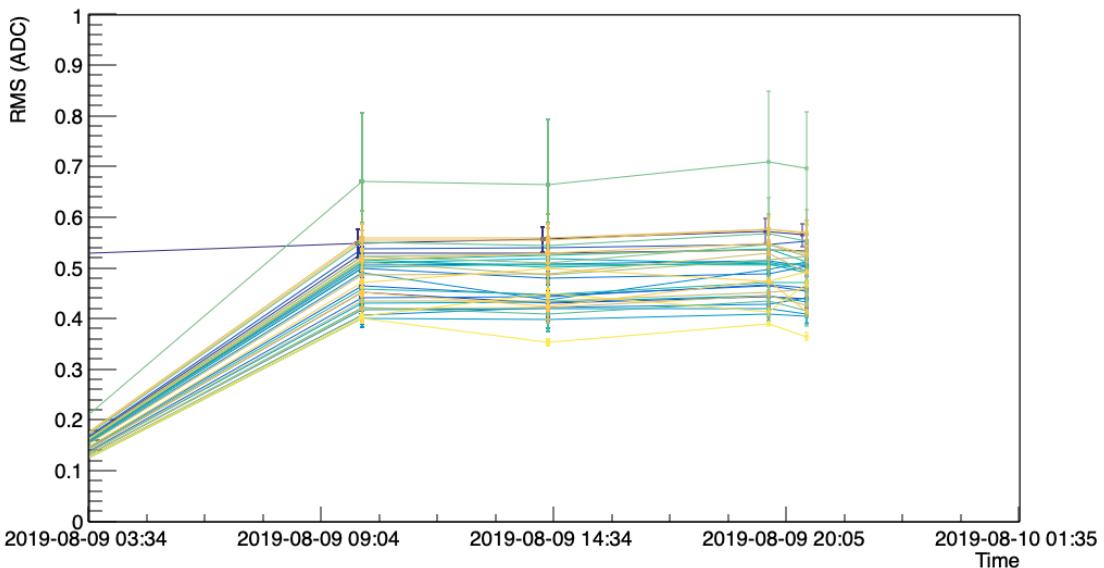


Fig. 9.9 Online monitoring diagram showing the average noise RMS on the CAEN DAQ channels every hour.

After data acquisition, the data from each DAQ is decoded into root format. The Keysight decoding process and analysis was originally designed for tests of a gaseous photomultiplier [318]. The code was updated for use with the new subsystem. An iterative procedure is undertaken to find the pulse and baseline of the waveform. Initially, the waveform baseline is taken as the average ADC of the waveform. The noise RMS is taken as the standard deviation from a Gaussian fit to the distribution. A simple pulse height algorithm is then employed to find the highest point in the waveform. The pulse width is found by iteratively calculating the mean of five ticks from the highest point in the pulse until the average is less than three times the size of the noise RMS. The highest point on a waveform and the width are shown

in event displays such as Figure 9.10. The integral of the waveform between the peak point and the stopping point is calculated and saved. The waveform outside the pulse width is used to define the baseline and the noise RMS and then the process is repeated.

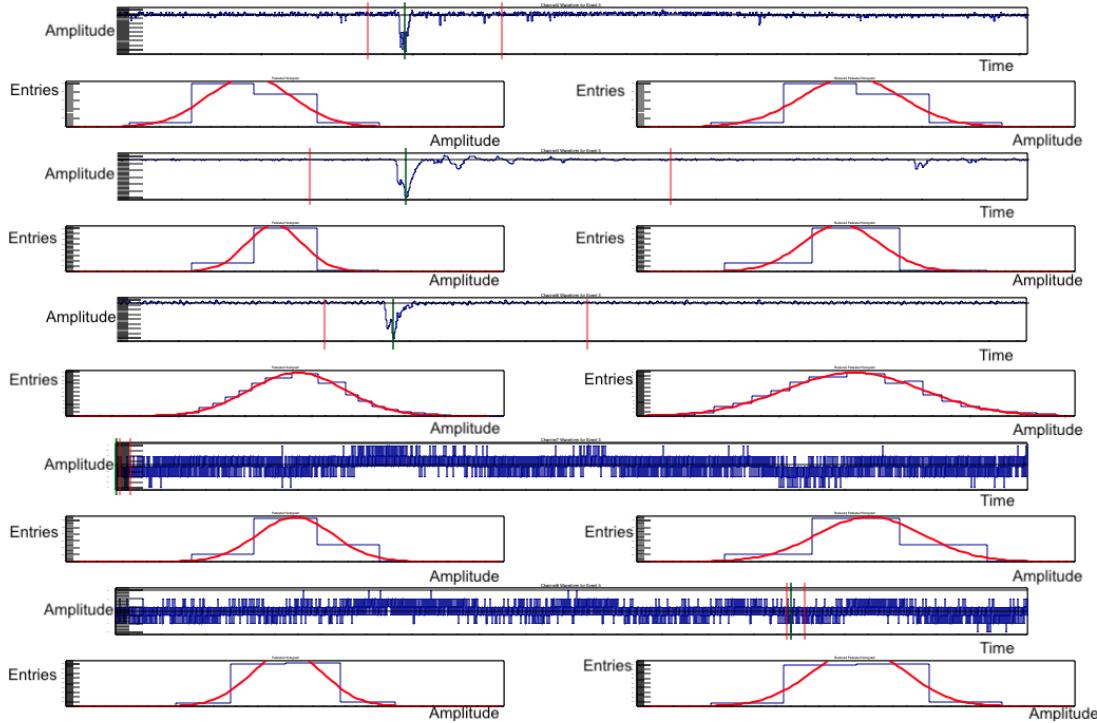


Fig. 9.10 The sub-detector event display for an event which triggered the PMT and two large CRTs. The top waveform (in ticks = 1 ns) corresponds to the internal PMT. The next two waveforms are the side CRTs and the bottom two are the small vertical CRTs. Underneath each waveform is the distribution of the waveform in ADC. A Gaussian is fitted to the distribution in the first (left-hand side) and second (right-hand side) iteration of the analysis. The integral of the waveform is taken within the two vertical red lines, whereas the baseline is calculated using the information outside the lines. The green vertical line represents the peak height point and the black horizontal line represents the baseline.

For the CAEN decoder, the pulse finding is not undertaken and the pedestal is calculated by taking the mean of a Gaussian fit to the waveform, removing points  $3\sigma$  away from the mean and then repeating the fit. In the future, hit finding and clustering algorithms are to be implemented as well as analyses such as electron lifetime and signal to noise. Noise analyses have been performed and are presented in the next Section.

## 9.5 Noise Analysis

A noise study was performed to evaluate the potential of the system. The noise was evaluated by calculated RMS of the noise from the system. This was done by taking data using the software trigger. A Gaussian fit was then performed around the baseline for each pixel/ROI for each event to extract the RMS. The fitting is performed for a second time with entries outside of  $5\sigma$  from the mean of the first fit removed, where  $\sigma$  is the standard deviation of the first Gaussian fit. The standard deviation of the Gaussian from the second fit is taken as the noise RMS. The RMS is then converted into electrons ( $\text{ENC}/e^-$ ) by using the shaping time of the LArASIC

$$1 \text{ ADC} = 1 \text{ ADC} \frac{2 \text{ V}}{2^{12} \text{ ADC}} \frac{1 \text{ fC}}{0.025 \text{ V}} \frac{6241 e^-}{1 \text{ fC}} = 121.89 e^-, \quad (9.1)$$

where 2 V is the peak to peak of the 12-bit DAQ and  $2^{12}$  is the number of ADC counts from the DAQ. The average over each event and the ROI/Pixel channels is then defined as the RMS of the system. The current noise RMS in air is  $276 \pm 46 e^-$  for the pixels and the RMS for each pixel is shown in Figure 9.11a. The noise RMS in air for the ROIs is  $330 \pm 37 e^-$  and the RMS for each ROI is shown in Figure 9.11b. The shape of the ROI distribution was not expected and investigations into this are ongoing. A reduction in the RMS is expected in cryogenic temperatures due to the reduction of the inherent noise of the pre-amplifiers. By converting the deposition of a Minimum Ionising Particle (MIP) travelling parallel to the pixel plane, the expected signal size of the system in ENC is

$$1 \text{ MIP} = 1.7 \text{ MeV/cm} \frac{1e^-}{0.254 \text{ cm}} \frac{0.68}{23.6 \text{ eV}} = 12,441.69 e^-, \quad (9.2)$$

where 1.7 MeV/cm is the typical energy deposited by a MIP, 0.254 cm is the width of a pixel, 23.6 eV is the energy required to ionise one argon atom, and 0.68 is the recombination factor. Therefore, a signal to noise ratio of 44.95 for the pixels and 37.7 for the ROIs at the anode should be expected for the system when operational. Note that this analysis is only performed on a small set of test data and should be repeated more rigorously. Methods described in Chapter 5 should also be employed to confirm this expectation during the first analysis runs.

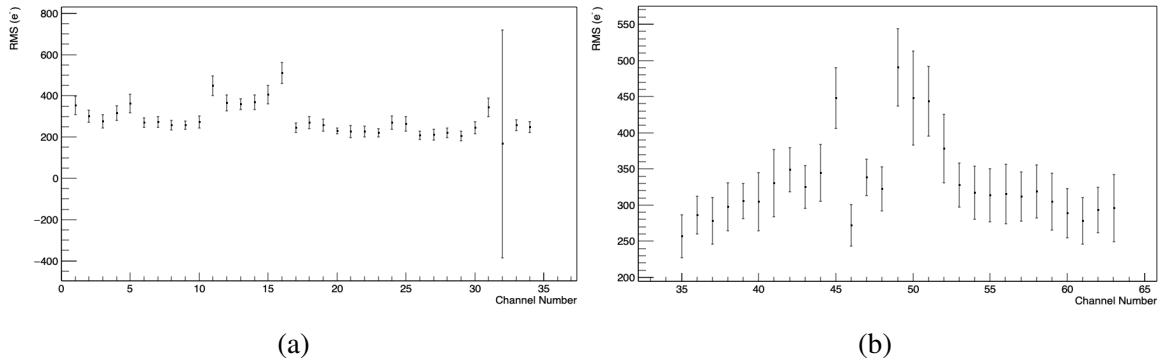


Fig. 9.11 The mean noise RMS on the pixel channels (a) and the ROI channels (b). The errors are the standard deviation on the RMS calculated for each event.

Additional studies of the noise were also undertaken. A Fourier transform analysis, the result of which are shown in Figure 9.12, was performed using the sample described above. This was performed individually for each event and the average is presented. The result of the Fourier transform analysis shows three noise sources at 6000, 8000 and 12500 Hz. Studies should be performed to identify the noise sources. However, as is shown by Equation 9.1, the largest peak is below 1.51 ADC and therefore is only just detected by the ADC due to the ADC resolution.

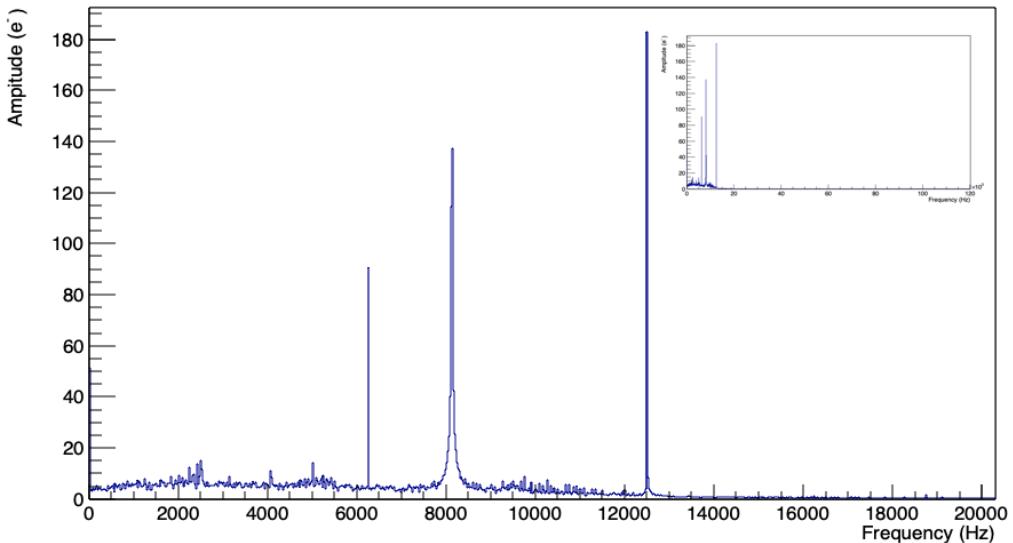


Fig. 9.12 FFT from software triggered data. The sub-figure is the entire frequency range available.

In case additional noise occurs in the future, a FFT noise filter was created. The filter works by using the same FFT analysis described above and identifying points in the frequency

spectrum above a user set amplitude and frequency. These frequencies are then removed from the FFT of each event and an inverse FFT is applied to obtain the noise filtered waveform. The result of this procedure is shown in Figure 9.13 where a sinusoidal noise (that has been since been removed from the system) is identified and removed.

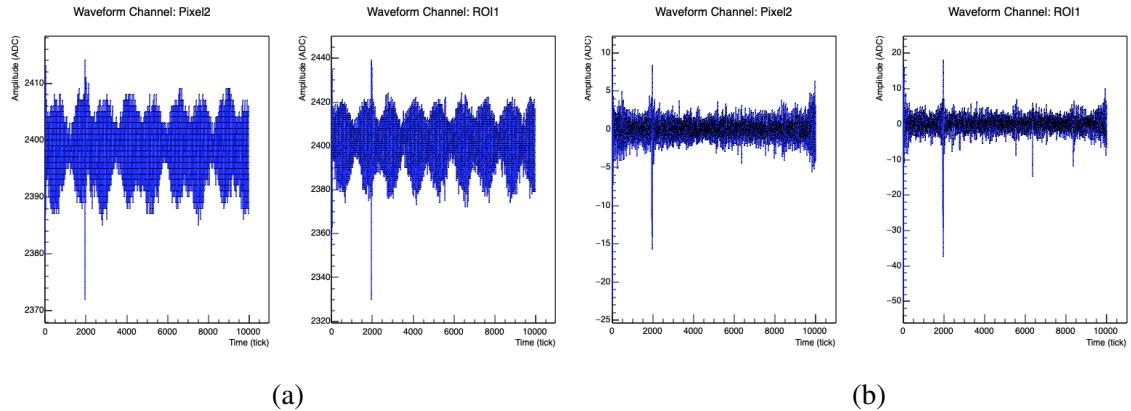


Fig. 9.13 a) A test pulse on a pixel (left) and an ROI (right) from the Ardunio. Since the image, alterations to the grounding scheme were undertaken and the oscillatory noise has been removed. b) The same waveforms but with an FFT noise filter applied.

## 9.6 Outlook

Numerous tests have been undertaken on the Sheffield liquid argon rig in order to prepare it for use. The entire rig has been tested with gaseous argon. As well as this, the purification system has been tested successfully with liquid argon over a period of several days. Testing undertaken on the sub-detectors using gaseous argon has been completed and resulted in successful data acquisition and triggering. The data acquisition system and online monitoring have also been tested in preparation for a liquid argon run.

Therefore, the rig is ready for the first liquid argon test and physics analysis run. During this, it will be necessary to re-tune the online monitoring. If successful, the rig is expected to show at least a signal to noise ratio of 44.95 for the pixels and 37.7 for the ROIs when using the Bern Pixel design [33] and BNL LArASIC [208].

Whilst the rig is currently set up for the above technologies, it was designed to have the flexibility to incorporate new developments and electronics. Therefore other technologies, such as LArPix [183] which is being considered for the DUNE near detector [115], can be easily tested and compared using the Sheffield rig in the future. These pixelated designs are to be tested in the ArgonCube project [319] to verify the scalability and performance within a neutrino beam.

# Chapter 10

## Conclusion

This thesis has presented developments towards an electron neutrino oscillation appearance analysis in the Short Baseline Neutrino (SBN) programme. Such an analysis is an important objective for the SBN collaboration and will confirm or refute the low energy excess anomalies observed at Liquid Scintillator Neutrino Detector (LSND) and Mini Booster Neutrino Experiment (MiniBooNE). The measurement could provide evidence for one or more sterile neutrinos, which have been suggested as a potential cause of the excess. If the sterile hypothesis is true, then it would introduce new physics beyond the standard model.

Construction of the Short Baseline Near Detector (SBND) and Commissioning of the Imaging Cosmic And Rare Underground Signals (ICARUS) detector, the far detector, is underway. The the Micro Booster Neutrino Experiment (MicroBooNE), which is the middle detector, has been running since 2015 and will search for the MiniBooNE excess independently as well as part of the SBN programme. All the detectors were discussed in Chapter 4. The details of the detectors,

For the SBND detector, in particular, new Cold Off The Shelf (COTS) ADCs were implemented to minimise the number of channels coming out of the cryostat and reduce thermal noise. In order to establish the COTS ADCs with the SBND electronics chain and to ensure the COTS ADCs reach the nominal signal to noise ratio required by the collaboration to perform the physics goals, the Vertical Slice Test (VST) was performed. The VST measured the signal to noise ratio with the Liquid Argon In A Test beam (LArIAT) Time Projection Chamber (TPC). The procedure was presented in Chapter 5. The VST average noise levels were found to be 398 ENC for the induction plane and 254 ENC for the collection plane before noise filtering. The signal to noise ratio was found to be approximately 58 for the collection plane and 15 for the induction plane. This was then extrapolated to the SBND detector dimensions using MicroBooNE noise data. The upper limit on the noise RMS in the SBND detector was then found to be 412 ENC for the collection plane and 630 ENC for

the induction plane. Therefore, the cold electronics used in the VST meet the SBND noise requirement of  $< 580$  ENC for the collection plane and  $< 700$  ENC for the induction plane. An alternative test-stand was also created from a hybrid of ICARUS and SBND electronics as a potential backup for the COTS ADC. The signal to noise ratio of the collection plane for the CERN test-stand was found to be  $34.60 \pm 0.21$  and for the induction plane, it was found to be  $12.08 \pm 0.08$ .

In order for the VST to run effectively, an online monitoring system was created and was presented in Chapter 5. This measured properties of the TPC such as the noise RMS and hit heights. The information was used in a website to monitor the detectors change during run time. In the online monitoring, an electron lifetime analysis was performed in order to measure the purity of the argon. This measurement was used in the signal to noise ratio analysis and was derived from previous simulated analyses of the lifetime in the SBND simulation. Different methods were developed to calculate the lifetime in SBND and have been used either in test-stand data or are to be employed in the online monitoring for SBND. The methods were discussed in Chapter 5.

A robust reconstruction framework is required for the SBN programme. The LArSoft framework has been employed in the SBN programme such that developments with other liquid argon experiments can be shared easily. The current status of the SBND reconstruction chain was presented in Chapter 6, where the Pandora pattern recognition software is used to cluster hits (charge depositions) recorded in the TPC into particle flow objects. Reconstruction modules downstream of the Pandora software then identify the particles based on the topology and calorimetric information. The main focus of this thesis was an electron neutrino appearance analysis from the muon neutrino Booster Neutrino Beam. Therefore, a new framework, the Tool based Reconstruction Algorithm for Characterising Showers (TRACS), was created to characterise the electromagnetic showers that are recorded in the TPC. TRACS was developed to provide users with a characterisation algorithm where it is easy to swap out calculations of the specific characteristics of the showers using a configuration file. TRACS built upon previous developments by incorporating algorithms into the framework and using them as a basis to develop the shower characterisation. There are now 31 different algorithms to calculate the six key properties of the shower. This allows experiments to choose the pathway that is best for the different analyses within each experiment. A new method to tune the reconstruction, called Bayesian Optimisation, was also introduced to maximise the selection efficiency in SBND.

Once the best reconstruction pathway was chosen, a  $\nu_e$  CC selection was undertaken on simulated data. This was initially done by using truth information from the Monte Carlo (MC). The truth-based selection set a benchmark for selections. As it has been several years

---

since the initial SBN proposal, the proposal selection was repeated in Chapter 7. The analysis incorporated the changes in the physics models and detector designs. Weights were applied to remove the changes since the proposal and mimic the proposal spectra. The differences due to the physics and detector design were then evaluated.

Various other reconstruction selections were also performed in SBND and presented in Chapter 7. Initially, the proposal selection was performed and then new metrics were introduced to improve the selection. The new selection, which took advantage of additional tracking calorimetric information, topological information and artifacts due to the reconstruction showed a slight increase in  $\nu_e$  CC selection efficiency when maximising the efficiency  $\times$  background rejection of 3.25%. The background rejection decreased by 0.0037% but this small change can have a large impact due to the large ratio of background compared to signal from the current best fits of the sterile parameters.

Further metrics were created from the shower characteristics to select electrons from the  $\nu_e$  CC signal and remove photons that occur in the background. However, these metrics had a small separational power. Therefore, two Boosted Decision Trees (BDT) were implemented to improve the selection. The first BDT scaled the signal and background to a 1:1 normalisation and the second scaled the signal and background using oscillation weights which mimicked the signal size expected in the far detector for a current global best fit for the sterile parameters. It was shown in Chapter 7 that a ROC curve of both BDTs encompassed the cut-based method. Hence the BDTs improve the selection power. However, it was shown that the truth-based selection had a significantly higher efficiency. This is primarily due to a lower reconstruction efficiency. Therefore, the reconstruction must improve in order to support analysis.

Furthermore, the reconstruction selections were performed on simulated data without an optical simulation and a cosmic simulation. Recent studies show that optical reconstruction is not 100% efficient and overlaying cosmic muon reduces the reconstruction efficiency. Therefore the current efficiency is expected to be less than stated.

Once the selections were optimised in SBND, SBN reconstruction was performed on simulated data in all three detectors. A  $\nu_e$  CC appearance study was then performed for the truth-based selection using the VALOR oscillation fitting framework, assuming a 3+1 sterile model, in Chapter 8. The sensitivity between the weighted proposal sample and the modern sample was compared to show how the changes since 2012 affect the sensitivity. Due to the addition of meson exchange current events in the simulation, the selection efficiency increased without an increase in the background. This, in turn, allowed for a better sensitivity for  $\nu_e$  appearance.

The truth selections were compared to the reconstruction selection in the analysis. The comparison showed the truth-based selection to have a larger sensitivity region than the reconstruction selections. The reconstructed cut-based methods have similar performances, but optimising the selections using the efficiency  $\times$  background rejection, resulted in a higher sensitivity. The analysis was then performed for various BDT selection thresholds for both BDTs. A BDT cut off score of 0.2 for the 1:1 normalised BDT resulted in the largest sensitivity region. This was an increase of  $0.0634 \text{ eV}^2$  in the area compared to the best cut based method.

Events were simulated without a cosmic overlay and optical simulation due to computing constraints. Therefore the expected sensitivity will decrease. Furthermore, several systematic errors were applied to the analysis; however, no detector systematic errors were applied and there are further interaction systematic errors to apply. Hence, the sensitivity will not be as powerful as presented. Major work is currently ongoing to improve the simulation, the reconstruction, to include detector systematic errors and to move to a more current interaction model (GENIE v\_3).

Finally, other research and development activities were shown in Chapter 9. This involved developing a liquid argon rig at the University of Sheffield to test new electronic readouts for future liquid argon experiments. The initial focus has been testing a 3D pixelated readout for the near detector of the Deep Underground Neutrino Experiment (DUNE). A LArTPC with sub-detectors has been designed and tested for the University of Sheffield rig. Initial testing of the purification system has also been successful. A software framework has been designed so that data can be taken using the DAQ, then decoded and analysed. An online monitoring system has been implemented such that the current running status of the rig can be checked. The rig is ready for a liquid argon test run before it is fully operational. The current setup should provide an easy basis for future research to be undertaken.

# References

- [1] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson, “Experimental test of parity conservation in beta decay,” *Phys. Rev.*, vol. 105, pp. 1413–1415, Feb 1957.
- [2] G. Danby, J.-M. Gaillard, K. Goulian, L. M. Lederman, N. Mistry, M. Schwartz, and J. Steinberger, “Observation of high-energy neutrino reactions and the existence of two kinds of neutrinos,” *Phys. Rev. Lett.*, vol. 9, pp. 36–44, Jul 1962.
- [3] K. Hirata *et al.*, “Results from one thousand days of real time directional solar neutrino data,” *Phys. Rev. Lett.*, vol. 65, pp. 1297–1300, 1990.
- [4] Y. e. Fukuda, “Evidence for oscillation of atmospheric neutrinos,” *Phys. Rev. Lett.*, vol. 81, pp. 1562–1567, Aug 1998.
- [5] Q. Ahmad *et al.*, “Direct evidence for neutrino flavor transformation from neutral current interactions in the Sudbury Neutrino Observatory,” *Phys. Rev. Lett.*, vol. 89, p. 011301, 2002.
- [6] F. An *et al.*, “Neutrino Physics with JUNO,” *J. Phys.*, vol. G43, no. 3, p. 030401, 2016.
- [7] L. Zhan, Y. Wang, J. Cao, and L. Wen, “Determination of the Neutrino Mass Hierarchy at an Intermediate Baseline,” *Phys. Rev. D*, vol. 78, p. 111103, 2008.
- [8] Z. Hou, R. Keisler, L. Knox, M. Millea, and C. Reichardt, “How massless neutrinos affect the cosmic microwave background damping tail,” *Phys. Rev. D*, vol. 87, p. 083008, Apr 2013.
- [9] A. Aguilar-Arevalo *et al.*, “Evidence for neutrino oscillations from the observation of anti-neutrino(electron) appearance in a anti-neutrino(muon) beam,” *Phys. Rev.*, vol. D64, p. 112007, 2001.
- [10] A. Aguilar-Arevalo *et al.*, “Significant Excess of ElectronLike Events in the Mini-BooNE Short-Baseline Neutrino Experiment,” *Phys. Rev. Lett.*, vol. 121, no. 22, p. 221801, 2018.
- [11] C. Giunti and T. Lasserre, “eV-scale Sterile Neutrinos,” *Ann. Rev. Nucl. Part. Sci.*, vol. 69, pp. 163–190, 2019.
- [12] M. Dentler, A. Hernández-Cabezudo, J. Kopp, P. A. N. Machado, M. Maltoni, I. Martínez-Soler, and T. Schwetz, “Updated Global Analysis of Neutrino Oscillations in the Presence of eV-Scale Sterile Neutrinos,” *JHEP*, vol. 08, p. 010, 2018.

- [13] R. Acciarri *et al.*, "Design and Construction of the MicroBooNE Detector," *JINST*, vol. 12, no. 02, p. P02017, 2017.
- [14] L. Landau, "On the energy loss of fast particles by ionization," *J. Phys.(USSR)*, vol. 8, pp. 201–205, 1944. [Translated by Haar, D. Ter, *Collected Papers of L.D. Landau*, Pergamon , 1965].
- [15] D. Casper, "The Nuance neutrino physics simulation, and the future," *Nucl. Phys. B Proc. Suppl.*, vol. 112, pp. 161–170, 2002.
- [16] J. Formaggio and G. Zeller, "From eV to EeV: Neutrino Cross Sections Across Energy Scales," *Rev. Mod. Phys.*, vol. 84, pp. 1307–1341, 2012.
- [17] R. Acciarri *et al.*, "A proposal for a three detector short-baseline neutrino oscillation program in the fermilab booster neutrino beam," 2015.
- [18] A. Aguilar-Arevalo *et al.*, "The Neutrino Flux prediction at MiniBooNE," *Phys. Rev. D*, vol. 79, p. 072002, 2009.
- [19] The MicroBooNE Collaboration, "Booster neutrino flux prediction at microboone," tech. rep., July 2018.
- [20] P. Abratenko *et al.*, "Measurement of the Flux-Averaged Inclusive Charged-Current Electron Neutrino and Antineutrino Cross Section on Argon using the NuMI Beam and the MicroBooNE Detector," 1 2021.
- [21] M. Auger *et al.*, "A Novel Cosmic Ray Tagger System for Liquid Argon TPC Neutrino Detectors," *Instruments*, vol. 1, no. 1, p. 2, 2017.
- [22] B. Fleming, "The MicroBooNE Technical Design Report," tech. rep., 2 2012.
- [23] M. Antonello *et al.*, "Experimental search for the "LSND anomaly" with the ICARUS detector in the CNGS neutrino beam," *Eur. Phys. J. C*, vol. 73, no. 3, p. 2345, 2013.
- [24] I. Nutini, "The LArIAT experiment," *Journal of Physics: Conference Series*, vol. 888, p. 012168, sep 2017.
- [25] Square, Inc. , "Cubism." <http://square.github.io/cubism/>. Accessed: 12-06-20.
- [26] F. JAMES and M. WINKLER, "Minuit function minimization and error analysis: Reference manual version 94.1." <https://root.cern.ch/root/html/doc/guides/minuit2/Minuit2.pdf>. Accessed: 01/06/2017.
- [27] 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*, vol. 78, no. 1, p. 82, 2018.
- [28] The MicroBooNE Collaboration, "Selection of numu charged-current induced interactions with  $N=0$  protons and performance of events with  $N=2$  protons in the final state in the MicroBooNE detector from the BNB, Public Note," 10 2018.
- [29] R. Acciarri *et al.*, "A Study of Electron Recombination Using Highly Ionizing Particles in the ArgoNeuT Liquid Argon TPC," *JINST*, vol. 8, p. P08005, 2013.

- [30] E. Brochu, V. M. Cora, and N. de Freitas, “A Tutorial on Bayesian Optimization of Expensive Cost Functions, with Application to Active User Modeling and Hierarchical Reinforcement Learning,” *arXiv e-prints*, p. arXiv:1012.2599, Dec. 2010.
- [31] J. Kopp, P. A. N. Machado, M. Maltoni, and T. Schwetz, “Sterile Neutrino Oscillations: The Global Picture,” *JHEP*, vol. 05, p. 050, 2013.
- [32] C. J. Clopper and E. S. Pearson, “The use of confidence or fiducial limits illustrated in the case of the binomial,” *Biometrika*, vol. 26, no. 4, pp. 404–413, 1934.
- [33] J. Asaadi *et al.*, “First Demonstration of a Pixelated Charge Readout for Single-Phase Liquid Argon Time Projection Chambers,” *Instruments*, vol. 4, no. 1, p. 9, 2020.
- [34] I. Esteban, M. C. Gonzalez-Garcia, A. Hernandez-Cabezudo, M. Maltoni, and T. Schwetz, “Global analysis of three-flavour neutrino oscillations: synergies and tensions in the determination of  $\theta_{23}$ ,  $\delta_{CP}$ , and the mass ordering,” *JHEP*, vol. 01, p. 106, 2019.
- [35] B. Pontecorvo, “Neutrino Experiments and the Problem of Conservation of Leptonic Charge,” *Sov. Phys. JETP*, vol. 26, pp. 984–988, 1968. [Zh. Eksp. Teor. Fiz. 53, 1717(1967)].
- [36] Z. Maki, M. Nakagawa, and S. Sakata, “Remarks on the unified model of elementary particles,” *Progress of Theoretical Physics*, vol. 28, no. 5, p. 870, 1962.
- [37] R. Davis, “A review of the homestake solar neutrino experiment,” *Progress in Particle and Nuclear Physics*, vol. 32, pp. 13 – 32, 1994.
- [38] J. Boger *et al.*, “The Sudbury neutrino observatory,” *Nucl. Instrum. Meth. A*, vol. 449, pp. 172–207, 2000.
- [39] C. Athanassopoulos *et al.*, “Candidate events in a search for  $\bar{\nu}_\mu \rightarrow \bar{\nu}_e$  oscillations,” *Phys. Rev. Lett.*, vol. 75, pp. 2650–2653, Oct 1995.
- [40] J. N. Abdurashitov *et al.*, “Measurement of the solar neutrino capture rate with gallium metal. III: Results for the 2002–2007 data-taking period,” *Phys. Rev.*, vol. C80, p. 015807, 2009.
- [41] P. Anselmann *et al.*, “First results from the 51cr neutrino source experiment with the gallex detector,” *Physics Letters B*, vol. 342, no. 1, pp. 440 – 450, 1995.
- [42] G. Mention, M. Fechner, T. Lasserre, T. Mueller, D. Lhuillier, M. Cribier, and A. Letourneau, “The Reactor Antineutrino Anomaly,” *Phys. Rev. D*, vol. 83, p. 073006, 2011.
- [43] E. Majorana, “Teoria simmetrica dell’elettrone e del positrone,” *Il Nuovo Cimento (1924-1942)*, vol. 14, no. 4, p. 171, 2008. [Translated from Italian by Luciano Maiani “Il Nuovo Cimento” 14(1937) 171-184].
- [44] K. N. Abazajian *et al.*, “Light sterile neutrinos: A white paper,” 2012.

- [45] W. Pauli. Letter from W. Pauli to Physical Institute of the Federal Institute of Technology, Zürich, 4 December 1930, <http://microboone-docdb.fnal.gov/cgi-bin/RetrieveFile?docid=953;filename=pauli%20letter1930.pdf>, Accessed on 20/07/2017 [Translated by: Kurt Riesselmann].
- [46] E. Fermi, "Versuch einer theorie der  $\beta$ -strahlen. i," *Zeitschrift für Physik*, vol. 88, no. 3-4, pp. 161–177, 1934. [Translated by Wilson, Fred L., "Fermi's Theory of Beta Decay", *Am. J. Phys.*, vol. 36, no. 12, pp 1150-1160].
- [47] S. Bilenky, *Introduction to the Physics of Massive and Mixed Neutrinos*. Lecture Notes in Physics, Springer Berlin Heidelberg, 2010.
- [48] F. Reines and C. L. Cowan, "Detection of the free neutrino," *Phys. Rev.*, vol. 92, pp. 830–831, Nov 1953.
- [49] C. L. Cowan, F. Reines, F. B. Harrison, H. W. Kruse, and A. D. McGuire, "Detection of the free neutrino: a confirmation," *Science*, vol. 124, no. 3212, pp. 103–104, 1956.
- [50] T.-D. Lee and C.-N. Yang, "Question of parity conservation in weak interactions," *Physical Review*, vol. 104, no. 1, p. 254, 1956.
- [51] A. Salam, "On parity conservation and neutrino mass," *Nuovo Cim.*, vol. 5, pp. 299–301, 1957.
- [52] L. Landau, "On the conservation laws for weak interactions," *Nucl. Phys.*, vol. 3, pp. 127–131, 1957.
- [53] T. D. Lee and C. N. Yang, "Parity nonconservation and a two-component theory of the neutrino," *Phys. Rev.*, vol. 105, pp. 1671–1675, Mar 1957.
- [54] S. Bilenky, "The History of neutrino oscillations," *Phys. Scripta T*, vol. 121, pp. 17–22, 2005.
- [55] M. Goldhaber, L. Grodzins, and A. W. Sunyar, "Helicity of neutrinos," *Phys. Rev.*, vol. 109, pp. 1015–1017, Feb 1958.
- [56] R. P. Feynman and M. Gell-Mann, "Theory of the fermi interaction," *Phys. Rev.*, vol. 109, pp. 193–198, Jan 1958.
- [57] E. C. G. Sudarshan and R. E. Marshak, "Chirality invariance and the universal fermi interaction," *Phys. Rev.*, vol. 109, pp. 1860–1862, Mar 1958.
- [58] H. L. Anderson, T. Fujii, R. H. Miller, and L. Tau, "Electron decay of the positive pion," *Phys. Rev. Lett.*, vol. 2, pp. 53–56, Jan 1959.
- [59] L. M. Lederman, "Neutrino physics technical report," 1 1963.
- [60] S. Frankel, J. Halpern, L. Holloway, W. Wales, M. Yearian, O. Chamberlain, A. Lemonick, and F. M. Pipkin, "New limit on the  $e + \gamma$  decay mode of the muon," *Phys. Rev. Lett.*, vol. 8, pp. 123–125, Feb 1962.
- [61] S. Weinberg, "A model of leptons," *Phys. Rev. Lett.*, vol. 19, pp. 1264–1266, Nov 1967.

- [62] A. Salam and J. C. Ward, “Electromagnetic and weak interactions,” *Physics Letters*, vol. 13, pp. 168–171, Nov. 1964.
- [63] S. L. Glashow, “Partial Symmetries of Weak Interactions,” *Nucl. Phys.*, vol. 22, pp. 579–588, 1961.
- [64] M. Schwartz, *Quantum Field Theory and the Standard Model*. Quantum Field Theory and the Standard Model, Cambridge University Press, 2014.
- [65] K. Kodama *et al.*, “Observation of tau neutrino interactions,” *Phys. Lett. B*, vol. 504, pp. 218–224, 2001.
- [66] M. Aker *et al.*, “Improved Upper Limit on the Neutrino Mass from a Direct Kinematic Method by KATRIN,” *Phys. Rev. Lett.*, vol. 123, no. 22, p. 221802, 2019.
- [67] B. Pontecorvo, “Mesonium and antimesonium,” *Soviet Phys. JETP*, vol. 6, 2 1958.
- [68] B. Pontecorvo, “Heavy leptons and neutrino astronomy,” *JETP Lett.*, vol. 13, pp. 199–201, 1971.
- [69] S. Bilenky and B. Pontecorvo, “Quark-lepton analogy and neutrino oscillations,” *Physics Letters B*, vol. 61, no. 3, pp. 248 – 250, 1976.
- [70] S. M. Bilenky and S. T. Petcov, “Massive neutrinos and neutrino oscillations,” *Rev. Mod. Phys.*, vol. 59, pp. 671–754, Jul 1987.
- [71] S. Bilenky and B. Pontecorvo, “Lepton mixing and neutrino oscillations,” *Physics Reports*, vol. 41, no. 4, pp. 225 – 261, 1978.
- [72] C. Giunti and C. W. Kim, *Fundamentals of Neutrino Physics and Astrophysics*. 4 2007.
- [73] S. Boyd, “Lecture notes in neutrino physics.” <https://warwick.ac.uk/fac/sci/physics/staff/academic/boyd/stuff/neutrinolectures/>. 2019.
- [74] L. Wolfenstein, “Neutrino oscillations in matter,” *Phys. Rev. D*, vol. 17, pp. 2369–2374, May 1978.
- [75] E. Akhmedov, “Quantum mechanics aspects and subtleties of neutrino oscillations,” in *Proceedings of the International Conference on History of the Neutrino: 1930-2018: Paris, France. September 5-7, 2018*, 2019.
- [76] B. Cleveland, T. Daily, J. Davis, Raymond, J. R. Distel, K. Lande, C. Lee, P. S. Wildenhain, and J. Ullman, “Measurement of the solar electron neutrino flux with the Homestake chlorine detector,” *Astrophys. J.*, vol. 496, pp. 505–526, 1998.
- [77] J. N. Bahcall, M. H. Pinsonneault, and G. J. Wasserburg, “Solar models with helium and heavy-element diffusion,” *Reviews of Modern Physics*, vol. 67, pp. 781–808, Oct. 1995.

- [78] K. S. Hirata, T. Kajita, T. Kifune, K. Kihara, M. Nakahata, K. Nakamura, S. Ohara, Y. Oyama, N. Sato, M. Takita, Y. Totsuka, Y. Yaginuma, M. Mori, A. Suzuki, K. Takahashi, T. Tanimori, M. Yamada, M. Koshiba, T. Suda, K. Miyano, H. Miyata, H. Takei, K. Kaneyuki, H. Nagashima, Y. Suzuki, E. W. Beier, L. R. Feldscher, E. D. Frank, W. Frati, S. B. Kim, A. K. Mann, F. M. Newcomer, R. Van Berg, and W. Zhang, “Observation of  ${}^8\text{B}$  solar neutrinos in the kamiokande-ii detector,” *Phys. Rev. Lett.*, vol. 63, pp. 16–19, July 1989.
- [79] K. Abe *et al.*, “Solar Neutrino Measurements in Super-Kamiokande-IV,” *Phys. Rev.*, vol. D94, no. 5, p. 052010, 2016.
- [80] Y. Fukuda *et al.*, “Measurement of a small atmospheric muon-neutrino / electron-neutrino ratio,” *Phys. Lett. B*, vol. 433, pp. 9–18, 1998.
- [81] S. Fukuda *et al.*, “Constraints on neutrino oscillations using 1258 days of Super-Kamiokande solar neutrino data,” *Phys. Rev. Lett.*, vol. 86, pp. 5656–5660, 2001.
- [82] K. Abe *et al.*, “Atmospheric neutrino oscillation analysis with external constraints in Super-Kamiokande I-IV,” *Phys. Rev. D*, vol. 97, no. 7, p. 072001, 2018.
- [83] K. Zuber, “The discovery of neutrino oscillations,” *Annalen Phys.*, vol. 528, pp. 452–457, 2016.
- [84] A. Achterberg *et al.*, “First Year Performance of The IceCube Neutrino Telescope,” *Astropart. Phys.*, vol. 26, pp. 155–173, 2006.
- [85] M. G. Aartsen *et al.*, “Determining neutrino oscillation parameters from atmospheric muon neutrino disappearance with three years of IceCube DeepCore data,” *Phys. Rev.*, vol. D91, no. 7, p. 072004, 2015.
- [86] Q. Ahmad *et al.*, “Measurement of the rate of  $\nu_e + d \rightarrow p + p + e^-$  interactions produced by  ${}^8\text{B}$  solar neutrinos at the Sudbury Neutrino Observatory,” *Phys. Rev. Lett.*, vol. 87, p. 071301, 2001.
- [87] W. Hampel *et al.*, “GALLEX solar neutrino observations: Results for GALLEX III.,” *Phys. Lett. B*, vol. 388, pp. 384–396, 1996.
- [88] M. Altmann *et al.*, “GNO solar neutrino observations: Results for GNO I,” *Phys. Lett. B*, vol. 490, pp. 16–26, 2000.
- [89] J. N. Abdurashitov *et al.*, “Solar neutrino flux measurements by the Soviet-American Gallium Experiment (SAGE) for half the 22 year solar cycle,” *J. Exp. Theor. Phys.*, vol. 95, pp. 181–193, 2002. [Zh. Eksp. Teor. Fiz. 122, 211(2002)].
- [90] G. Bellini *et al.*, “Final results of Borexino Phase-I on low energy solar neutrino spectroscopy,” *Phys. Rev. D*, vol. 89, no. 11, p. 112007, 2014.
- [91] S. Abe *et al.*, “Precision Measurement of Neutrino Oscillation Parameters with KamLAND,” *Phys. Rev. Lett.*, vol. 100, p. 221803, 2008.
- [92] M. Tanabashi *et al.*, “Review review of particle physics,” *Phys. Rev. D*, vol. 98, p. 030001, 2018 and 2019 update.

- [93] P. Adamson *et al.*, “A Study of Muon Neutrino Disappearance Using the Fermilab Main Injector Neutrino Beam,” *Phys. Rev.*, vol. D77, p. 072002, 2008.
- [94] M. A. Acero *et al.*, “First Measurement of Neutrino Oscillation Parameters using Neutrinos and Antineutrinos by NOvA,” *Phys. Rev. Lett.*, vol. 123, no. 15, p. 151803, 2019.
- [95] The MicroBooNE Collaboration, “Automated selection of electron neutrinos from the numi beam in the microboone detector and prospects for a measurement of the charged-current inclusive cross section, public note,” tech. rep., November 2018.
- [96] S. E. Kopp, “The NuMI neutrino beam at Fermilab,” *AIP Conf. Proc.*, vol. 773, no. 1, pp. 276–278, 2005.
- [97] K. Abe *et al.*, “Measurement of Neutrino Oscillation Parameters from Muon Neutrino Disappearance with an Off-axis Beam,” *Phys. Rev. Lett.*, vol. 111, no. 21, p. 211803, 2013.
- [98] M. Ahn *et al.*, “Measurement of Neutrino Oscillation by the K2K Experiment,” *Phys. Rev. D*, vol. 74, p. 072003, 2006.
- [99] N. Agafonova *et al.*, “Observation of a first  $\nu_\tau$  candidate in the OPERA experiment in the CNGS beam,” *Phys. Lett. B*, vol. 691, pp. 138–145, 2010.
- [100] K. Abe *et al.*, “Observation of Electron Neutrino Appearance in a Muon Neutrino Beam,” *Phys. Rev. Lett.*, vol. 112, p. 061802, 2014.
- [101] P. Adamson *et al.*, “Electron neutrino and antineutrino appearance in the full MINOS data sample,” *Phys. Rev. Lett.*, vol. 110, no. 17, p. 171801, 2013.
- [102] M. Acero *et al.*, “New constraints on oscillation parameters from  $\nu_e$  appearance and  $\nu_\mu$  disappearance in the NOvA experiment,” *Phys. Rev. D*, vol. 98, p. 032012, 2018.
- [103] T. Nakaya and R. K. Plunkett, “Neutrino Oscillations with the MINOS, MINOS+, T2K, and NOvA Experiments,” *New J. Phys.*, vol. 18, no. 1, p. 015009, 2016.
- [104] X. Guo *et al.*, “A Precision measurement of the neutrino mixing angle  $\theta_{13}$  using reactor antineutrinos at Daya-Bay,” 2007.
- [105] F. An *et al.*, “Independent measurement of the neutrino mixing angle  $\theta_{13}$  via neutron capture on hydrogen at Daya Bay,” *Phys. Rev. D*, vol. 90, no. 7, p. 071101, 2014.
- [106] H. de Kerret *et al.*, “First Double Chooz  $\theta_{13}$  Measurement via Total Neutron Capture Detection,” *Nature Phys.*, vol. 16, no. 5, pp. 558–564, 2020.
- [107] M. Apollonio *et al.*, “Search for neutrino oscillations on a long baseline at the CHOOZ nuclear power station,” *Eur. Phys. J.*, vol. C27, pp. 331–374, 2003.
- [108] J. K. Ahn *et al.*, “Observation of Reactor Electron Antineutrino Disappearance in the RENO Experiment,” *Phys. Rev. Lett.*, vol. 108, p. 191802, 2012.
- [109] F. P. An *et al.*, “A side-by-side comparison of Daya Bay antineutrino detectors,” *Nucl. Instrum. Meth.*, vol. A685, pp. 78–97, 2012.

- [110] F. P. An *et al.*, “Observation of electron-antineutrino disappearance at daya bay,” *Phys. Rev. Lett.*, vol. 108, p. 171803, Apr 2012.
- [111] D. A. Dwyer, “Antineutrinos from nuclear reactors: recent oscillation measurements,” *New Journal of Physics*, vol. 17, p. 025003, feb 2015.
- [112] K. Abe *et al.*, “Search for CP Violation in Neutrino and Antineutrino Oscillations by the T2K Experiment with  $2.2 \times 10^{21}$  Protons on Target,” *Phys. Rev. Lett.*, vol. 121, no. 17, p. 171802, 2018.
- [113] A. Chatterjee, M. M. Devi, M. Ghosh, R. Moharana, and S. K. Raut, “Probing CP violation with the first three years of ultrahigh energy neutrinos from IceCube,” *Phys. Rev.*, vol. D90, no. 7, p. 073003, 2014.
- [114] K. Abe *et al.*, “Hyper-Kamiokande Design Report,” 5 2018.
- [115] B. Abi *et al.*, “Deep Underground Neutrino Experiment (DUNE), Far Detector Technical Design Report, Volume II DUNE Physics,” 2 2020.
- [116] S. J. Parke, “Determining the neutrino mass hierarchy,” in *Third NO-VE International Workshop on Neutrino Oscillations in Venice : Fifty years after the neutrino experimental discovery : Venezia, February 7-10, 2006, Istituto Veneto di Scienze, Lettere ed Arti, Campo Santo Stefano*, pp. 115–125, 7 2006.
- [117] A. Giaz, “Status and perspectives of the JUNO experiment,” *Journal of Physics: Conference Series*, vol. 1056, p. 012048, Jul 2018.
- [118] M. G. Aartsen *et al.*, “PINGU: A Vision for Neutrino and Particle Physics at the South Pole,” *J. Phys.*, vol. G44, no. 5, p. 054006, 2017.
- [119] Hofestadt, Jannik, “Prospects for measuring the neutrino mass hierarchy with KM3NeT/ORCA,” in *25th European Cosmic Ray Symposium*, 1 2017.
- [120] S. Dell’Oro, S. Marcocci, M. Viel, and F. Vissani, “Neutrinoless double beta decay: 2015 review,” *Adv. High Energy Phys.*, vol. 2016, p. 2162659, 2016.
- [121] M. C. Chen, “Neutrinoless double beta decay experiments,” *Journal of Physics: Conference Series*, vol. 598, p. 012008, apr 2015.
- [122] S. M. Bilenky, A. Faessler, T. Gutsche, and F. Simkovic, “Neutrinoless double beta decay and neutrino mass hierarchies,” *Phys. Rev.*, vol. D72, p. 053015, 2005.
- [123] C. Patrignani *et al.*, “Review of Particle Physics,” *Chin. Phys. C*, vol. 40, no. 10, p. 100001, 2016.
- [124] S. Mertens, “Direct neutrino mass experiments,” *Journal of Physics: Conference Series*, vol. 718, p. 022013, may 2016.
- [125] N. Aghanim *et al.*, “Planck 2018 results. VI. Cosmological parameters,” 2018.
- [126] J. Lesgourgues and S. Pastor, “Neutrino mass from Cosmology,” *Adv. High Energy Phys.*, vol. 2012, p. 608515, 2012.

- [127] M. Gell-Mann, P. Ramond, and R. Slansky, “Complex Spinors and Unified Theories,” *Conf. Proc. C*, vol. 790927, pp. 315–321, 1979.
- [128] M. Srednicki, *Quantum Field Theory*. Cambridge University Press, 2007.
- [129] S. Raby and R. Slansky, “Neutrino masses: How to add them to the standard model,” *Los Alamos Sci.*, vol. 25, pp. 64–71, 1997.
- [130] K. N. Abazajian *et al.*, “Neutrino Physics from the Cosmic Microwave Background and Large Scale Structure,” *Astropart. Phys.*, vol. 63, pp. 66–80, 2015.
- [131] R. K. Sachs and A. M. Wolfe, “Perturbations of a Cosmological Model and Angular Variations of the Microwave Background,” *apj*, vol. 147, p. 73, Jan. 1967.
- [132] S. Schael *et al.*, “Precision electroweak measurements on the Z resonance,” *Phys. Rept.*, vol. 427, pp. 257–454, 2006.
- [133] M. Carena, A. de Gouvea, A. Freitas, and M. Schmitt, “Invisible Z boson decays at e+ e- colliders,” *Phys. Rev.*, vol. D68, p. 113007, 2003.
- [134] J. Conrad and M. Shaevitz, *Sterile Neutrinos: An Introduction to Experiments*, vol. 28, pp. 391–442. 2018.
- [135] S. Böser, C. Buck, C. Giunti, J. Lesgourgues, L. Ludhova, S. Mertens, A. Schukraft, and M. Wurm, “Status of Light Sterile Neutrino Searches,” *Prog. Part. Nucl. Phys.*, vol. 111, p. 103736, 2020.
- [136] A. A. Aguilar-Arevalo *et al.*, “A Combined  $\nu_\mu \rightarrow \nu_e$  and  $\bar{\nu}_\mu \rightarrow \bar{\nu}_e$  Oscillation Analysis of the MiniBooNE Excesses,” 2012.
- [137] C. Athanassopoulos *et al.*, “The Liquid scintillator neutrino detector and LAMPF neutrino source,” *Nucl. Instrum. Meth.*, vol. A388, pp. 149–172, 1997.
- [138] J. M. Conrad, W. C. Louis, and M. H. Shaevitz, “The LSND and MiniBooNE Oscillation Searches at High  $\Delta m^2$ ,” *Ann. Rev. Nucl. Part. Sci.*, vol. 63, pp. 45–67, 2013.
- [139] S. M. Bilenky, C. Giunti, and W. Grimus, “Phenomenology of neutrino oscillations,” *Prog. Part. Nucl. Phys.*, vol. 43, pp. 1–86, 1999.
- [140] W. Louis, “Lsnd neutrino oscillation results and implications,” *Progress in Particle and Nuclear Physics*, vol. 40, pp. 151 – 165, 1998. Neutrinos in Astro, Particle and Nuclear Physics.
- [141] A. Aguilar-Arevalo *et al.*, “The MiniBooNE Detector,” *Nucl. Instrum. Meth. A*, vol. 599, pp. 28–46, 2009.
- [142] C. Giunti, A. Ioannisian, and G. Ranucci, “A new analysis of the MiniBooNE low-energy excess,” 2019.
- [143] M. Sorel, J. M. Conrad, and M. Shaevitz, “Combined analysis of short-baseline neutrino experiments in the (3+ 1) and (3+ 2) sterile neutrino oscillation hypotheses,” *Physical Review D*, vol. 70, no. 7, p. 073004, 2004.

- [144] S. Gariazzo, “Light sterile neutrinos: the current picture from neutrino oscillations,” in *16th International Conference on Topics in Astroparticle and Underground Physics (TAUP 2019) Toyama, Japan, September 9-13, 2019*, 2019.
- [145] P. Adamson *et al.*, “Search for sterile neutrinos in minos and minos+ using a two-detector fit,” *Phys. Rev. Lett.*, vol. 122, p. 091803, Mar 2019.
- [146] P. Adamson *et al.*, “Search for active-sterile neutrino mixing using neutral-current interactions in nova,” *Phys. Rev. D*, vol. 96, p. 072006, Oct 2017.
- [147] K. Abe *et al.*, “Limits on sterile neutrino mixing using atmospheric neutrinos in Super-Kamiokande,” *Phys. Rev. D*, vol. 91, p. 052019, 2015.
- [148] M. G. Aartsen *et al.*, “Searches for sterile neutrinos with the icecube detector,” *Phys. Rev. Lett.*, vol. 117, p. 071801, Aug 2016.
- [149] M. Aartsen *et al.*, “Search for sterile neutrino mixing using three years of IceCube DeepCore data,” *Phys. Rev. D*, vol. 95, no. 11, p. 112002, 2017.
- [150] B. Armbruster *et al.*, “Upper limits for neutrino oscillations muon-anti-neutrino  $\rightarrow$  electron-anti-neutrino from muon decay at rest,” *Phys. Rev. D*, vol. 65, p. 112001, 2002.
- [151] P. Astier *et al.*, “Search for nu(mu)  $\rightarrow$  nu(e) oscillations in the NOMAD experiment,” *Phys. Lett. B*, vol. 570, pp. 19–31, 2003.
- [152] N. Agafonova *et al.*, “Search for  $\nu_\mu \rightarrow \nu_e$  oscillations with the OPERA experiment in the CNGS beam,” *JHEP*, vol. 07, p. 004, 2013. [Addendum: JHEP 07, 085 (2013)].
- [153] L. Borodovsky, C. Y. Chi, Y. Ho, N. Kondakis, W. Lee, J. Mechalakos, B. Rubin, R. Seto, C. Stoughton, G. Tzanakos, B. Blumenfeld, L. Chichura, J. Krizmanic, E. Lincke, L. Lueking, W. Lyle, A. Pevsner, W. P. Hogan, E. O’Brien, K. Reardon, P. D. Sheldon, and G. W. Sullivan, “Search for muon-neutrino oscillations  $\nu_\mu \rightarrow \nu_e$   $\bar{\nu}_\mu \rightarrow \bar{\nu}_e$  in a wide-band neutrino beam,” *Phys. Rev. Lett.*, vol. 68, pp. 274–277, Jan 1992.
- [154] C. Rott and on behalf of the JSNS2 collaboration, “Status of *jsns<sup>2</sup>* - j-parc sterile neutrino search at j-parc spallation neutron source,” *Journal of Physics: Conference Series*, vol. 1468, p. 012175, Sep 2019.
- [155] C. Rubbia, “The liquid-argon time projection chamber: a new concept for neutrino detectors,” Tech. Rep. CERN-EP-INT-77-8, CERN, Geneva, 1977.
- [156] B. Baller *et al.*, “Liquid Argon Time Projection Chamber Research and Development in the United States,” *JINST*, vol. 9, p. T05005, 2014.
- [157] C. Adams *et al.*, “Ionization electron signal processing in single phase LArTPCs. Part II. Data/simulation comparison and performance in MicroBooNE,” *JINST*, vol. 13, no. 07, p. P07007, 2018.
- [158] R. Acciarri *et al.*, “Noise Characterization and Filtering in the MicroBooNE Liquid Argon TPC,” *JINST*, vol. 12, no. 08, p. P08003, 2017.

- [159] S. Amerio *et al.*, “Design, construction and tests of the ICARUS T600 detector,” *Nucl. Instrum. Meth. A*, vol. 527, pp. 329–410, 2004.
- [160] P. Cennini *et al.*, “Detection of scintillation light in coincidence with ionizing tracks in a liquid argon time projection chamber,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 432, no. 2, pp. 240 – 248, 1999.
- [161] P. K. and, “Liquid argon scintillation light studies in LArIAT,” *Journal of Physics: Conference Series*, vol. 888, p. 012184, sep 2017.
- [162] L. Onsager, “Initial recombination of ions,” *Phys. Rev.*, vol. 54, pp. 554–557, Oct 1938.
- [163] G. Jaffé, “Zur theorie der ionisation in kolonnen,” *Annalen der Physik*, vol. 347, no. 12, pp. 303–344, 1913.
- [164] H. Kramers, “On a modification of jaffé’s theory of column-ionization,” *Physica*, vol. 18, no. 10, pp. 665 – 675, 1952.
- [165] J. Thomas and D. A. Imel, “Recombination of electron-ion pairs in liquid argon and liquid xenon,” *Phys. Rev. A*, vol. 36, pp. 614–616, Jul 1987.
- [166] S. Amoruso *et al.*, “Study of electron recombination in liquid argon with the icarus tpc,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 523, no. 3, pp. 275 – 286, 2004.
- [167] J. B. Birks, “Scintillations from organic crystals: Specific fluorescence and relative response to different radiations,” *Proceedings of the Physical Society. Section A*, vol. 64, pp. 874–877, oct 1951.
- [168] “LArSoft Physics Constants.” [https://nusoft.fnal.gov/larsoft/doxsvn/html/PhysicalConstants\\_8h\\_source.html](https://nusoft.fnal.gov/larsoft/doxsvn/html/PhysicalConstants_8h_source.html). Accessed: 11-06-2020.
- [169] M. Antonello *et al.*, “Detection of cherenkov light emission in liquid argon,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 516, no. 2, pp. 348 – 363, 2004.
- [170] D. W. Swan, “Drift velocity of electrons in liquid argon, and the influence of molecular impurities,” *Proceedings of the Physical Society*, vol. 83, pp. 659–666, apr 1964.
- [171] BNL, “Liquid argon properties (tables and calculators).” <https://lar.bnl.gov/properties/>. Accessed: 11/06/2020.
- [172] W. Walkowiak, “Drift velocity of free electrons in liquid argon,” *Nucl. Instrum. Meth. A*, vol. 449, pp. 288–294, 2000.
- [173] Y. Li, T. Tsang, C. Thorn, X. Qian, M. Diwan, J. Joshi, S. Kettell, W. Morse, T. Rao, J. Stewart, W. Tang, and B. Viren, “Measurement of longitudinal electron diffusion in liquid argon,” *Nuclear Instruments and Methods in Physics Research A*, vol. 816, pp. 160–170, Apr. 2016.

- [174] M. Schenk, *Studies with a Liquid Argon Time Projection Chamber: Addressing Technological Challenges of Large-Scale Detectors*. BestMasters, Springer Fachmedien Wiesbaden, 2015.
- [175] S. Amoruso *et al.*, “Analysis of the liquid argon purity in the icarus t600 tpc,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 516, no. 1, pp. 68 – 79, 2004.
- [176] M. Antonello *et al.*, “Experimental observation of an extremely high electron lifetime with the ICARUS-T600 LAr-TPC,” *JINST*, vol. 9, no. 12, p. P12006, 2014.
- [177] C. Anderson *et al.*, “The ArgoNeuT Detector in the NuMI Low-Energy beam line at Fermilab,” *JINST*, vol. 7, p. P10019, 2012.
- [178] The MicroBooNE Collaboration, “A measurement of the attenuation of drifting electrons in the microboone larpc, public note,” tech. rep., August 2017.
- [179] D. Adams *et al.*, “Design and performance of a 35-ton liquid argon time projection chamber as a prototype for future very large detectors,” *JINST*, vol. 15, no. 03, p. P03035, 2020.
- [180] B. Abi *et al.*, “First results on ProtoDUNE-SP liquid argon time projection chamber performance from a beam test at the CERN Neutrino Platform,” *JINST*, vol. 15, no. 12, p. P12004, 2020.
- [181] M. Mooney, “The MicroBooNE Experiment and the Impact of Space Charge Effects,” in *Proceedings, Meeting of the APS Division of Particles and Fields (DPF 2015): Ann Arbor, Michigan, USA, 4-8 Aug 2015*, 2015.
- [182] S. Palestini and K. McDonald, “Space charge in ionization detectors,” 01 2010.
- [183] J. Asaadi *et al.*, “A New Concept for Kilotonne Scale Liquid Argon Time Projection Chambers,” *Instruments*, vol. 4, no. 1, p. 6, 2020.
- [184] A. Scarpelli, “ProtoDUNE and a Dual-phase LArTPC,” *PoS*, vol. NOW2018, p. 036, 2019.
- [185] L. Zambelli and S. Murphy, “WA105: A large demonstrator of a liquid argon dual phase TPC,” *Journal of Physics: Conference Series*, vol. 888, p. 012202, sep 2017.
- [186] L. Alvarez-Ruso *et al.*, “NuSTEC White Paper: Status and challenges of neutrino–nucleus scattering,” *Prog. Part. Nucl. Phys.*, vol. 100, pp. 1–68, 2018.
- [187] G. F. Chew and G. C. Wick, “The impulse approximation,” *Phys. Rev.*, vol. 85, pp. 636–642, Feb 1952.
- [188] T. Katori, “Meson Exchange Current (MEC) Models in Neutrino Interaction Generators,” *AIP Conf. Proc.*, vol. 1663, no. 1, p. 030001, 2015.
- [189] E. A. Paschos and D. Schalla, “Coherent Pion Production by Neutrinos,” *Phys. Rev.*, vol. D80, p. 033005, 2009.

- [190] D. Rein and L. M. Sehgal, “Neutrino Excitation of Baryon Resonances and Single Pion Production,” *Annals Phys.*, vol. 133, pp. 79–153, 1981.
- [191] S. Dytman, “Final state interactions in neutrino-nucleus experiments,” *Acta Phys. Polon. B*, vol. 40, pp. 2445–2460, 2009.
- [192] J. Cugnon, “Intranuclear cascade model. a review,” *Nuclear Physics A*, vol. 387, no. 1, pp. 191 – 203, 1982.
- [193] I. Stancu *et al.*, “Technical Design Report for the 8 GeV Beam,” tech. rep., 5 2001.
- [194] A. A. Aguilar-Arevalo *et al.*, “The Neutrino Flux prediction at MiniBooNE,” *Phys. Rev.*, vol. D79, p. 072002, 2009.
- [195] P. Adamson *et al.*, “The NuMI Neutrino Beam,” *Nucl. Instrum. Meth. A*, vol. 806, pp. 279–306, 2016.
- [196] P. Adamson *et al.*, “Neutrino and antineutrino inclusive charged-current cross section measurements with the minos near detector,” *Phys. Rev. D*, vol. 81, p. 072002, Apr 2010.
- [197] P. A. Machado, O. Palamara, and D. W. Schmitz, “The Short-Baseline Neutrino Program at Fermilab,” *Ann. Rev. Nucl. Part. Sci.*, vol. 69, pp. 363–387, 2019.
- [198] N McConkey, “SBND: Status of the fermilab short-baseline near detector,” *Journal of Physics: Conference Series*, vol. 888, p. 012148, sep 2017.
- [199] A. Machado and E. Segreto, “ARAPUCA a new device for liquid argon scintillation light detection,” *JINST*, vol. 11, no. 02, p. C02004, 2016.
- [200] Hamamatsu, *Large Photocathode Area Photomultiper Tubes*. [https://www.hamamatsu.com/resources/pdf/etd/LARGE\\_AREA\\_PMT TPMH1376E.pdf](https://www.hamamatsu.com/resources/pdf/etd/LARGE_AREA_PMT TPMH1376E.pdf).
- [201] T. Kaptanoglu, “Characterization of the Hamamatsu 8" R5912-MOD Photomultiplier Tube,” *Nucl. Instrum. Meth.*, vol. A889, pp. 69–77, 2018.
- [202] SensL, *Low Noise, Blue-Sensitive Silicon Photomultipliers*, 2014. <http://sensl.com/downloads/ds/DS-MicroCseries.pdf>.
- [203] E. Segreto, A. Machado, L. Paulucci, F. Marinho, D. Galante, S. Guedes, A. Fauth, V. Teixeira, B. Gelli, M. Guzzo, W. Araujo, C. Ambrósio, M. Bissiano, and A. L. Filho, “Liquid argon test of the ARAPUCA device,” *Journal of Instrumentation*, vol. 13, pp. P08021–P08021, aug 2018.
- [204] L. Paulucci, F. Marinho, A. A. Machado, and E. Segreto, “A complete simulation of the X-ARAPUCA device for detection of scintillation photons,” *JINST*, vol. 15, no. 01, p. C01047, 2020.
- [205] A. Machado, E. Segreto, D. Warner, A. Fauth, B. Gelli, R. Maximo, A. Pizolatti, L. Paulucci, and F. Marinho, “The X-ARAPUCA: An improvement of the ARAPUCA device,” 4 2018.

- [206] V. Radeka *et al.*, “Cold electronics for ‘Giant’ Liquid Argon Time Projection Chambers,” *J. Phys. Conf. Ser.*, vol. 308, p. 012021, 2011.
- [207] H. Chen, G. D. Geronimo], F. Lanni, D. Lissauer, D. Makowiecki, V. Radeka, S. Rescia, C. Thorn, and B. Yu, “Front end readout electronics of the microboone experiment,” *Physics Procedia*, vol. 37, pp. 1287 – 1294, 2012. Proceedings of the 2nd International Conference on Technology and Instrumentation in Particle Physics (TIPP 2011).
- [208] G. De Geronimo, A. D’Andragora, S. Li, N. Nambiar, S. Rescia, E. Vernon, H. Chen, F. Lanni, D. Makowiecki, V. Radeka, C. Thorn, and B. Yu, “Front-end asic for a liquid argon tpc,” *IEEE Nuclear Science Symposium Medical Imaging Conference*, pp. 1658–1666, 2010.
- [209] S. Ohkawa, M. Yoshizawa, and K. Husimi, “Direct synthesis of the gaussian filter for nuclear pulse amplifiers,” *Nuclear Instruments and Methods*, vol. 138, no. 1, pp. 85 – 92, 1976.
- [210] Anlog Devices Inc, *3 MSPS,10-/12-Bit ADCs in 8-Lead TSOT AD7273/AD7274*. [https://www.analog.com/media/en/technical-documentation/data-sheets/AD7273\\_7274.pdf](https://www.analog.com/media/en/technical-documentation/data-sheets/AD7273_7274.pdf).
- [211] H. Chen *et al.*, “Lifetime Study of COTS ADC for SBND LAr TPC Readout Electronics,” in *Proceedings, 21st IEEE Real Time Conference (RT2018): Williamsburg, USA, June 11-15, 2018*, 6 2018.
- [212] R. Acciarri *et al.*, “Noise Characterization and Filtering in the MicroBooNE Liquid Argon TPC,” *JINST*, vol. 12, no. 08, p. P08003, 2017.
- [213] P. Abratenko *et al.*, “Measurement of Space Charge Effects in the MicroBooNE LArTPC Using Cosmic Muons,” *JINST*, vol. 15, no. 12, p. P12037, 2020.
- [214] P. Abratenko *et al.*, “Neutrino Event Selection in the MicroBooNE Liquid Argon Time Projection Chamber using Wire-Cell 3-D Imaging, Clustering, and Charge-Light Matching,” 11 2020.
- [215] C. Adams *et al.*, “Rejecting cosmic background for exclusive charged current quasi elastic neutrino interaction studies with Liquid Argon TPCs; a case study with the MicroBooNE detector,” *Eur. Phys. J. C*, vol. 79, no. 8, p. 673, 2019.
- [216] P. Abratenko *et al.*, “Measurement of differential cross sections for  $\nu_\mu$  -Ar charged-current interactions with protons and no pions in the final state with the MicroBooNE detector,” *Phys. Rev. D*, vol. 102, no. 11, p. 112013, 2020.
- [217] M. Antonello *et al.*, “Precise 3D track reconstruction algorithm for the ICARUS T600 liquid argon time projection chamber detector,” *Adv. High Energy Phys.*, vol. 2013, p. 260820, 2013.
- [218] Electron Tubes Ltd, *9357FLA photomultiplier no. 1116 data sheet*.
- [219] D.G. Michael *et al.*, “The magnetized steel and scintillator calorimeters of the minos experiment,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 596, no. 2, pp. 190 – 228, 2008.

- [220] F. Ardellier *et al.*, “Double Chooz: A Search for the neutrino mixing angle theta(13),” 2006.
- [221] A. Diaz, C. Argüelles, G. Collin, J. Conrad, and M. Shaevitz, “Where Are We With Light Sterile Neutrinos?,” 5 2019.
- [222] CERN, “Root a data analysis framework | root a data analysis framework.” <https://root.cern.ch>. Accessed: 01/06/2018.
- [223] “Art.” <https://art.fnal.gov>. Accessed: 07-04-2020.
- [224] “LArSoft.” <https://larsoft.org>. Accessed: 07-04-2020.
- [225] C. Andreopoulos *et al.*, “The GENIE Neutrino Monte Carlo Generator,” *Nucl. Instrum. Meth.*, vol. A614, pp. 87–104, 2010.
- [226] “Corsika – an air shower simulation program.” <https://github.com/LArSoft/larg4>. Accessed: 07-04-2020.
- [227] D. Heck, J. Knapp, J. N. Capdevielle, G. Schatz, and T. Thouw, “CORSIKA: A Monte Carlo code to simulate extensive air showers,” 1998.
- [228] “LArG4.” <https://github.com/LArSoft/larg4>. Accessed: 07-04-2020.
- [229] S. Agostinelli *et al.*, “Geant4—a simulation toolkit,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 506, no. 3, pp. 250 – 303, 2003.
- [230] “Geant4 website.” <https://geant4.web.cern.ch>. Accessed: 10-02-2020.
- [231] F. Cavanna, M. Kordosky, J. Raaf, and B. Rebel, “LArIAT: Liquid Argon In A Testbeam,” 6 2014.
- [232] R. Acciarri *et al.*, “The Liquid Argon In A Testbeam (LArIAT) Experiment,” *JINST*, vol. 15, no. 04, p. P04026, 2020.
- [233] K. Biery, E. Flumerfelt, J. Freeman, W. Ketchum, G. Lukhanin, and R. Rechenmacher, “artdaq: DAQ software development made simple,” *Journal of Physics: Conference Series*, vol. 898, p. 032013, oct 2017.
- [234] “ArtDAQ.” <https://cdcvn.fnal.gov/redmine/projects/artdaq>. Accessed: 12-06-2020.
- [235] “ArtDAQ-daqinterface.” <https://cdcvn.fnal.gov/redmine/projects/artdaq-utilities/wiki/Artdaq-daqinterface>. Accessed: 12-06-2020.
- [236] “Experimental physics and industrial control system.” <https://epics.anl.gov/index.php>. Accessed: 12-06-20.
- [237] Redis Labs, “Redis website.” <https://redis.io>. Accessed: 12-06-20.
- [238] SNO+, “Minard - SNO+ Monitoring.” <https://github.com/snoplus/minard>. Accessed: 12-06-20.

- [239] PostgreSQL Global Development Group, “Postgresql.” <https://www.postgresql.org>. Accessed: 15/06/2020.
- [240] “SAM.” <https://cdcvns.fnal.gov/redmine/projects/sam/wiki>. Accessed: 12-06-20.
- [241] “File transfer service information.” [https://cdcvns.fnal.gov/redmine/projects/sam/wiki/File\\_Transfer\\_Service\\_Information](https://cdcvns.fnal.gov/redmine/projects/sam/wiki/File_Transfer_Service_Information). Accessed: 12-06-20.
- [242] A. Booth and M. Stancari, “Raw Hit Finder LArSoft Module.” <http://larsoft.org/single-record/?pdb=179>.
- [243] B. Baller, “LineCluster LArSoft Module.” <http://larsoft.org/single-record/?pdb=98>.
- [244] D. S. Robert Sulej, “PMA Track Algorithm LArSoft Module.” <http://larsoft.org/single-record/?pdb=102>.
- [245] CERN, “Root principle component analysis.” <https://root.cern.ch/doc/master/classTPrincipal.html>.
- [246] R. Acciarri *et al.*, “First Observation of Low Energy Electron Neutrinos in a Liquid Argon Time Projection Chamber,” *Phys. Rev.*, vol. D95, no. 7, p. 072005, 2017. [Phys. Rev.D95,072005(2017)].
- [247] CAEN, “Caen a2795.” <http://www.caen.it/csite/CaenProd.jsp?parent=10&idmod=1004>. Accessed: 01/06/2017.
- [248] CERN, “Tvirtualfft class reference.” <https://root.cern.ch/doc/master/classTVirtualFFT.html>.
- [249] “LArReco LArSoft Module.” <https://github.com/LArSoft/larreco>.
- [250] S. Ramo, “Currents induced by electron motion,” *Proceedings of the IRE*, vol. 27, no. 9, pp. 584–585, 1939.
- [251] C. Adams *et al.*, “Ionization electron signal processing in single phase LArTPCs. Part I. Algorithm Description and quantitative evaluation with MicroBooNE simulation,” *JINST*, vol. 13, no. 07, p. P07006, 2018.
- [252] N. Wiener, *Extrapolation, Interpolation, and Smoothing of Stationary Time Series*. The MIT Press, 1964.
- [253] J. Joshi and X. Qian, “Signal Processing in the MicroBooNE LArTPC,” in *Meeting of the APS Division of Particles and Fields*, 11 2015.
- [254] X. Qian, C. Zhang, B. Viren, and M. Diwan, “Three-dimensional Imaging for Large LArTPCs,” *JINST*, vol. 13, no. 05, p. P05032, 2018.
- [255] “GausHitFinder LArSoft Module.” [https://github.com/LArSoft/larreco/blob/develop/larreco/HitFinder/GausHitFinder\\_module.cc](https://github.com/LArSoft/larreco/blob/develop/larreco/HitFinder/GausHitFinder_module.cc). Accessed: 11/06/2020.
- [256] B. Baller, “Liquid argon TPC signal formation, signal processing and reconstruction techniques,” *JINST*, vol. 12, no. 07, p. P07010, 2017.

- [257] CERN, “Tf1 class reference.” <https://root.cern.ch/doc/master/classTF1.html>. Accessed: 01/06/2018.
- [258] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, “Scikit-learn: Machine learning in Python,” *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
- [259] Y. Freund and R. E. Schapire, “A decision-theoretic generalization of on-line learning and an application to boosting,” *Journal of Computer and System Sciences*, vol. 55, no. 1, pp. 119 – 139, 1997.
- [260] C. Cortes and V. Vapnik, “Support-vector networks,” in *Machine Learning*, pp. 273–297, 1995.
- [261] M. Wallbank, *Reconstruction and Analysis for the DUNE 35-ton Liquid Argon Prototype*. PhD thesis, University of Sheffield, 2018.
- [262] M. Wallbank, “Blurred Clustering LArSoft Module.” [https://github.com/LArSoft/larreco/blob/develop/larreco/ClusterFinder/BlurredClustering\\_module.cc](https://github.com/LArSoft/larreco/blob/develop/larreco/ClusterFinder/BlurredClustering_module.cc).
- [263] M. Wallbank, “The blurred clustering reconstruction technique,” tech. rep., University of Sheffield, August 2015.
- [264] M. Wallbank, “The emshower reconstruction technique,” tech. rep., University of Sheffield, June 2016.
- [265] “EMShower Module.” [https://github.com/LArSoft/larreco/blob/develop/larreco/ShowerFinder/EMShower\\_module.cc](https://github.com/LArSoft/larreco/blob/develop/larreco/ShowerFinder/EMShower_module.cc). Accessed: 15-02-2020.
- [266] “LArPandora Shower Creation Module.” [https://github.com/PandoraPFA/larpandora/blob/develop/larpandora/LArPandoraEventBuilding/LArPandoraShowerCreation\\_module.cc](https://github.com/PandoraPFA/larpandora/blob/develop/larpandora/LArPandoraEventBuilding/LArPandoraShowerCreation_module.cc). Accessed: 15-02-2020.
- [267] “LArPandora Track Creation Module.” [https://github.com/PandoraPFA/larpandora/blob/develop/larpandora/LArPandoraEventBuilding/LArPandoraTrackCreation\\_module.cc](https://github.com/PandoraPFA/larpandora/blob/develop/larpandora/LArPandoraEventBuilding/LArPandoraTrackCreation_module.cc). Accessed: 07-04-2020.
- [268] “Calorimetry Class Module.” [https://github.com/LArSoft/larreco/blob/develop/larreco/Calorimetry/Calorimetry\\_module.cc](https://github.com/LArSoft/larreco/blob/develop/larreco/Calorimetry/Calorimetry_module.cc). Accessed: 10-02-2020.
- [269] P. Abratenko *et al.*, “Determination of muon momentum in the MicroBooNE LArTPC using an improved model of multiple Coulomb scattering,” *JINST*, vol. 12, no. 10, p. P10010, 2017.
- [270] A. A. Aguilar-Arevalo, C. E. Anderson, A. O. Bazarko, S. J. Brice, B. C. Brown, L. Bugel, J. Cao, L. Coney, J. M. Conrad, D. C. Cox, A. Curioni, Z. Djurcic, D. A. Finley, B. T. Fleming, R. Ford, F. G. Garcia, G. T. Garvey, J. Grange, C. Green, J. A. Green, T. L. Hart, E. Hawker, R. Imlay, R. A. Johnson, G. Karagiorgi, P. Kasper, T. Katori, T. Kobilarcik, I. Kourbanis, S. Koutsoliotas, E. M. Laird, S. K. Linden, J. M. Link, Y. Liu, Y. Liu, W. C. Louis, K. B. M. Mahn, W. Marsh, C. Mauger, V. T. McGary, G. McGregor, W. Metcalf, P. D. Meyers, F. Mills, G. B. Mills, J. Monroe,

- C. D. Moore, J. Mousseau, R. H. Nelson, P. Nienaber, J. A. Nowak, B. Osmanov, S. Ouedraogo, R. B. Patterson, Z. Pavlovic, D. Perevalov, C. C. Polly, E. Prebys, J. L. Raaf, H. Ray, B. P. Roe, A. D. Russell, V. Sandberg, R. Schirato, D. Schmitz, M. H. Shaevitz, F. C. Shoemaker, D. Smith, M. Soderberg, M. Sorel, P. Spentzouris, J. Spitz, I. Stancu, R. J. Stefanski, M. Sung, H. A. Tanaka, R. Tayloe, M. Tzanov, R. G. Van de Water, M. O. Wascko, D. H. White, M. J. Wilking, H. J. Yang, G. P. Zeller, and E. D. Zimmerman, “First measurement of the muon neutrino charged current quasielastic double differential cross section,” *Phys. Rev. D*, vol. 81, p. 092005, May 2010.
- [271] “Chi2ParticleID Module.” [https://github.com/LArSoft/larana/blob/develop/larana/ParticleIdentification/Chi2ParticleID\\_module.cc](https://github.com/LArSoft/larana/blob/develop/larana/ParticleIdentification/Chi2ParticleID_module.cc). Accessed: 10-02-2020.
- [272] “TRACS Module.” [https://github.com/LArSoft/larreco/blob/develop/larreco/ShowerFinder/TRACS\\_module.cc](https://github.com/LArSoft/larreco/blob/develop/larreco/ShowerFinder/TRACS_module.cc). Accessed: 15-02-2020.
- [273] Artist, “Art tools.” [https://cdcvfs.fnal.gov/redmine/projects/art/wiki/Guide\\_to\\_writing\\_and\\_using\\_tools](https://cdcvfs.fnal.gov/redmine/projects/art/wiki/Guide_to_writing_and_using_tools). Accessed: 30/07/2020.
- [274] G. Guennebaud, B. Jacob, *et al.*, “Eigen.” <http://eigen.tuxfamily.org>, 2020.
- [275] “Pandora LArThreeDSlidingFitResult.” <https://github.com/PandoraPFA/LArContent/blob/master/larpandoracontent/LArObjects/LArThreeDSlidingFitResult.cc>. Accessed: 15-02-2020.
- [276] J. Snoek, H. Larochelle, and R. P. Adams, “Practical bayesian optimization of machine learning algorithms,” in *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 2*, NIPS’12, (Red Hook, NY, USA), p. 2951–2959, Curran Associates Inc., 2012.
- [277] F. Nogueira, “Bayesian Optimization: Open source constrained global optimization tool for Python,” 2014–.
- [278] P. I. Frazier, “A Tutorial on Bayesian Optimization,” *arXiv e-prints*, p. arXiv:1807.02811, July 2018.
- [279] K. O. D. Görtler, “A visual exploration of gaussian processes.” <https://distill.pub/2019/visual-exploration-gaussian-processes/>. Accessed: 11/06/2020.
- [280] N. Srinivas, A. Krause, S. M. Kakade, and M. Seeger, “Gaussian Process Optimization in the Bandit Setting: No Regret and Experimental Design,” *arXiv e-prints*, p. arXiv:0912.3995, Dec. 2009.
- [281] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, and H. Voss, “TMVA: Toolkit for Multivariate Data Analysis,” *PoS*, vol. ACAT, p. 040, 2007.
- [282] LArSoft Collaboration, “Larsoft collaboration.” <http://larsoft.org>. Accessed: 17/07/2020.
- [283] “GENIE Release Notes.” <http://releases.genie-mc.org>. Accessed: 20-04-2020.
- [284] “LArSoft Interaction Types ENUM.” [https://nusoft.fnal.gov/larsoft/doxsvn/html/MCNeutrino\\_8h\\_source.html](https://nusoft.fnal.gov/larsoft/doxsvn/html/MCNeutrino_8h_source.html). Accessed: 20-04-2020.

- [285] J. H. Friedman, “Greedy function approximation: A gradient boosting machine.,” *Ann. Statist.*, vol. 29, pp. 1189–1232, 10 2001.
- [286] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*. Springer Series in Statistics, New York, NY, USA: Springer New York Inc., 2001.
- [287] S. J. Russell and P. Norvig, *Artificial Intelligence: a modern approach*. Pearson, 3 ed., 2009.
- [288] L. Breiman, “Bagging predictors,” *Machine Learning*, vol. 24, no. 2, pp. 123–140, 1996.
- [289] C. Schaffer, “Selecting a classification method by cross-validation,” in *Machine Learning*, pp. 135–143, 1993.
- [290] “ROOT TH1 Source Code TH1::KolmogorovTest.” [https://root.cern.ch/doc/master/TH1\\_8cxx\\_source.html#l07683](https://root.cern.ch/doc/master/TH1_8cxx_source.html#l07683). Accessed: 08-06-2020.
- [291] C. Andreopoulos, C. Barry, F. Bench, A. Chappell, T. Dealtry, S. Dennis, L. Escudero, R. Jones, N. Grant, M. Roda, D. Sgalaberna, and R. Shah, “VALOR joint oscillation analysis using multiple LAr-TPCs in the booster neutrino beam at fermilab,” *Journal of Physics: Conference Series*, vol. 888, p. 012254, sep 2017.
- [292] T. Devlin, “Correlations from systematic corrections to poisson-distributed data in log-likelihood functions,” tech. rep., May 1999.
- [293] G. Fogli, E. Lisi, A. Marrone, D. Montanino, and A. Palazzo, “Getting the most from the statistical analysis of solar neutrino oscillations,” *Phys. Rev. D*, vol. 66, p. 053010, 2002.
- [294] G. Cowan, *Statistical data analysis*. Oxford University Press, USA, 1998.
- [295] L. Demortier, “Equivalence of the best-fit and covariance-matrix methods for comparing binned data with a model in the presence of correlated systematic uncertainties,” tech. rep., April 1999.
- [296] B. Roe, “Statistical errors in Monte Carlo estimates of systematic errors,” *Nucl. Instrum. Meth. A*, vol. 570, pp. 159–164, 2007.
- [297] M. Del Tutto, *First Measurements of Inclusive Muon Neutrino Charged Current Differential Cross Sections on Argon at 0.8 GeV Average Neutrino Energy with the MicroBooNE Detector*. PhD thesis, Oxford U., 2019.
- [298] Mike Kordosky, Kam-Biu Luk, Steven Manly, and A. DUNE, “Report from the DUNE Near Detector Concept Study Group,” 05 2018.
- [299] C. Adams, M. Del Tutto, J. Asaadi, M. Bernstein, E. Church, R. Guenette, J. M. Rojas, H. Sullivan, and A. Tripathi, “Enhancing Neutrino Event Reconstruction with Pixel-Based 3D Readout for Liquid Argon Time Projection Chambers,” *JINST*, vol. 15, no. 04, p. P04009, 2020.
- [300] Y. Li *et al.*, “A 20-Liter Test Stand with Gas Purification for Liquid Argon Research,” *JINST*, vol. 11, no. 06, p. T06001, 2016.

- [301] Catalyst Central, “Products/gettermax 133.” <http://www.catalyst-central.com/product>. Accessed: 05/05/2020.
- [302] Servomex, “Df-550e nanotrace oxygen analyzer.” <https://www.servomex.com/gas-analyzers/product-ranges/df/550e>. Accessed: 15/06/2020.
- [303] NI, “Scb-68a terminal block.” <https://www.ni.com/en-gb/support/model.scb-68a.html>. Accessed: 15/06/2020.
- [304] NI, “Pxi-6229 i/o card.” <https://www.ni.com/en-gb/support/model.pxi-6229.html>. Accessed: 15/06/2020.
- [305] NI, “Pci-8331 card.” <https://www.ni.com/pdf/manuals/370840b.pdf>. Accessed: 15/06/2020.
- [306] NI, “Labview.” <https://www.ni.com/en-gb/shop/labview/labview-details.html>. Accessed: 15/06/2020.
- [307] D. Göldi, *A Novel Liquid Argon Time Projection Chamber Detector: The ArgonCube Concept*. PhD thesis, Bern U., 2018.
- [308] Arduino, “Arduinonano.” <https://www.arduino.cc/en/Guide/ArduinoNano>. Accessed: 15/06/2020.
- [309] CAEN, “V1740 64 channel 12 bit 62.5 ms/s digitizer.” <https://www.caen.it/products/v1740/>. Accessed: 15/06/2020.
- [310] Ansys, “Maxwell.” <https://www.ansys.com/products/electronics/ansys-maxwell>. Accessed: 15/06/2020.
- [311] CAEN, “V6521 6 channel 6 kv/300  $\mu$ a vme hv power supply module.” <https://www.caen.it/products/v6521/>. Accessed: 15/06/2020.
- [312] Hamamatsu, “R11065 photomultiplier tube.” <https://www.hamamatsu.com/eu/en/product/type/R11065-20/index.html>. Accessed: 15/06/2020, Note: Older version used.
- [313] Keysight, “U5309a pcie high-speed digitizer - adc card, 8-bit, 2 gs/s, up to 8-ch, fpga signal processing.” <https://www.keysight.com/en/pd-2350477-pn-U5309A/8-bit-pcie-gen2-high-speed-digitizer-with-on-board-signal-processing?pm=PL&nid=-35502.1072169&cc=CY&lc=eng>. Accessed: 15/06/2020.
- [314] CAEN, “N417 8-channel low-threshold discriminator.” [http://moby dick.mib.infn.it/brofferio/didattica/datasheet/caen\\_N417-8chDiscriminator.pdf](http://moby dick.mib.infn.it/brofferio/didattica/datasheet/caen_N417-8chDiscriminator.pdf). Accessed: 15/06/2020.
- [315] CAEN, “Triple 4-fold logic unit/majority with veto.” <https://www.caen.it/products/n405/>. Accessed: 15/06/2020.
- [316] CAEN, “V1718 vme-usb2.0 bridge.” <https://www.caen.it/products/v1718/>. Accessed: 15/06/2020.

- [317] PostgreSQL Global Development Group, “libpq.” <https://www.postgresql.org/docs/9.5/libpq.html>. Accessed: 15/06/2020.
- [318] M. Thiesse, *Research and Development Toward Massive Liquid Argon Time Projection Chambers for Neutrino Detection*. PhD thesis, Sheffield U., 2017.
- [319] M. Auger, A. Ereditato, and J. Sinclair, “ArgonCube: a Modular Approach for Liquid Argon TPC Neutrino Detectors for Near Detector Environments,” Tech. Rep. CERN-SPSC-2017-025. SPSC-I-246, CERN, Geneva, Jun 2017.
- [320] C. Andreopoulos, C. Barry, S. Dytman, H. Gallagher, T. Golan, R. Hatcher, G. Perdue, and J. Yarba, “The GENIE Neutrino Monte Carlo Generator: Physics and User Manual,” 10 2015.
- [321] E. Longo and I. Sestili, “Monte Carlo Calculation of Photon Initiated Electromagnetic Showers in Lead Glass,” *Nucl. Instrum. Meth.*, vol. 128, p. 283, 1975. [Erratum: Nucl.Instrum.Meth. 135, 587 (1976)].
- [322] W. R. Nelson, T. M. Jenkins, R. C. McCall, and J. K. Cobb, “Electron-induced cascade showers in copper and lead at 1 gev,” *Phys. Rev.*, vol. 149, pp. 201–208, Sep 1966.



# Appendix A

## Full Hit Finding Efficiencies

In principle, a MIP-like muon deposits charge on every wire it crosses. The first definition (Figures A.1, A.4 and A.7) describes how many of these energy depositions are missed due to the thresholds set in the *GausHitFinder\_module.cc*. The first definition limits the use of the MC hit creation algorithm, removing the ambiguities from defining an MC hit. The efficiency definition can also be altered for implementation on real data.

The second definition (Figures A.2, A.5 and A.8) describes the efficiency of identifying the MC hits in reconstruction without the presence of delta rays. This is done by removing wires and the corresponding neighbouring wires when some of the charge on the wire occurs from a delta ray ionisation process. This definition is similar to the first efficiency definition; however, the process identifies the efficiency of detecting MC hit muon energy depositions only.

The third definition (Figures A.3, A.6 and A.9) is the efficiency of identifying energy depositions in the reconstruction. The definition, therefore, indicates the efficiency of identifying delta rays and muon hits separately and thus provides information on the fundamental efficiency level of vertex reconstruction in the detector.

In Figures A.4, A.5 and A.6 one can see the effect of altering the MC merging threshold. The efficiency increases as the resolution of the TPC becomes less important and more hits are merged into one hit. Increasing the MC amplitude threshold also increases the efficiency as more true energy depositions are not considered in the analysis since they are too small to see in the detector, see Figures A.1, A.2 and A.6. Finally Figures A.7, A.8 and A.9 show the effect of increasing the reconstructed hit finding threshold. As the threshold is increased, there is a decrease in efficiency, as expected. However, the decrease for the majority of topologies in the detector is minimal.

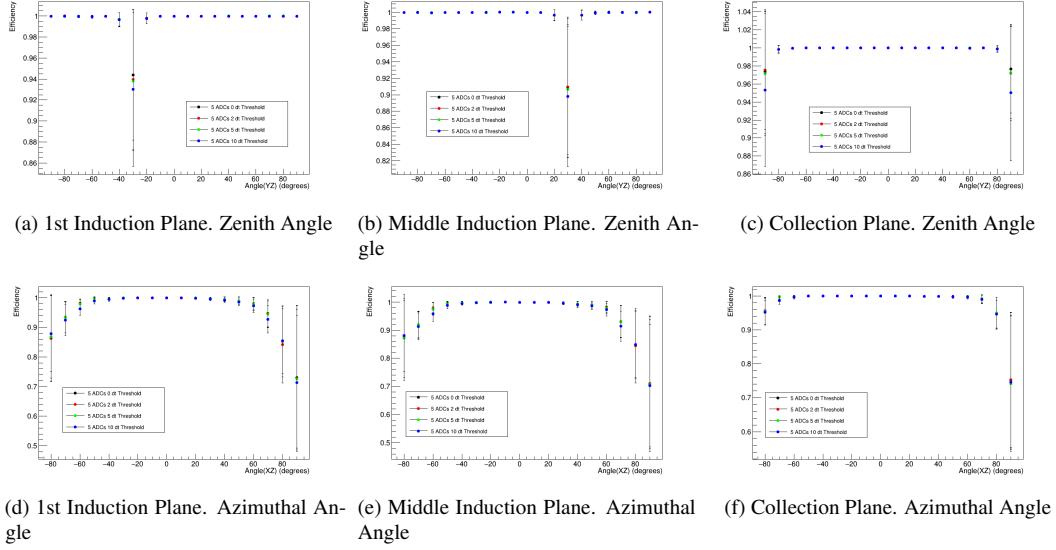


Fig. A.1 The 1st efficiency definition in Equation 6.1.2 for various MC  $dt$  thresholds as a function of the azimuthal (XZ) and zenith (YZ) angles. The MC ADC threshold is set to 5 ADC. Errors are the standard deviation of the efficiency for the corresponding angle.

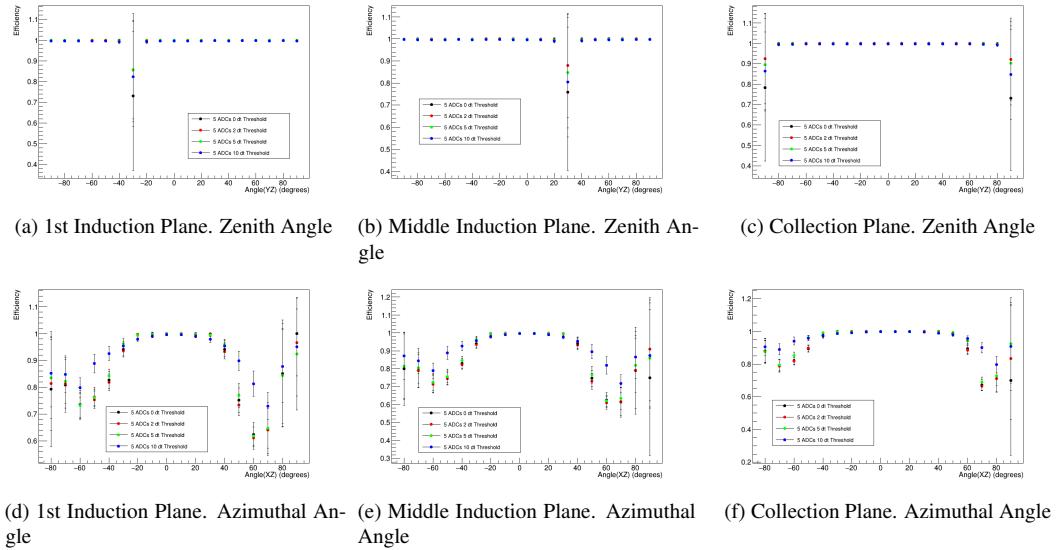


Fig. A.2 The 2nd efficiency definition in Equation 6.1.2 for various MC  $dt$  thresholds as a function of the azimuthal (XZ) and zenith (YZ) angles. The MC ADC threshold is set to 5 ADC. Errors are the standard deviation of the efficiency for the corresponding angle.

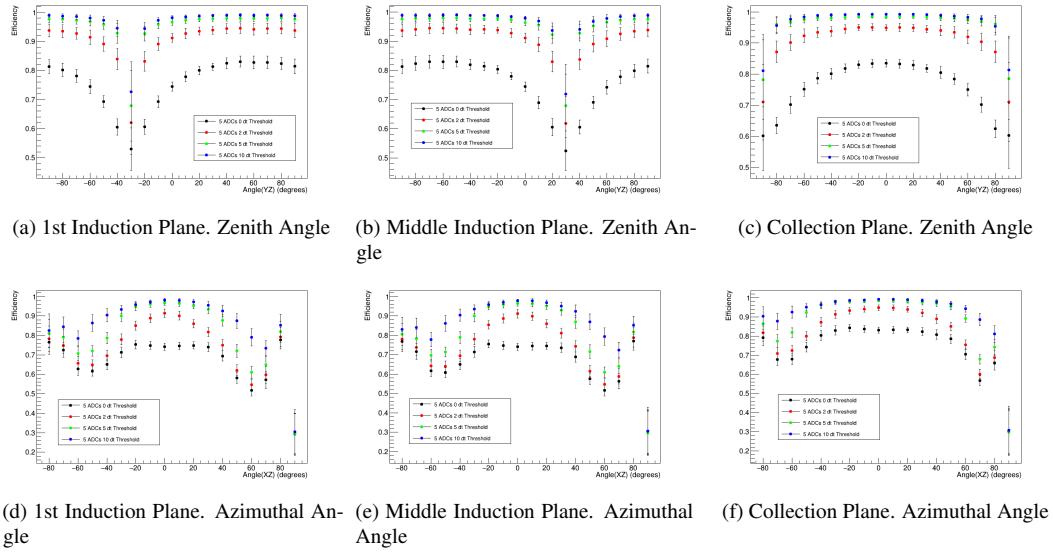


Fig. A.3 The 3rd efficiency definition in Equation 6.1.2 for various MC  $dt$  thresholds as a function of the azimuthal (XZ) and zenith (YZ) angles. The MC ADC threshold is set to 5 ADC. Errors are the standard deviation of the efficiency for the corresponding angle.

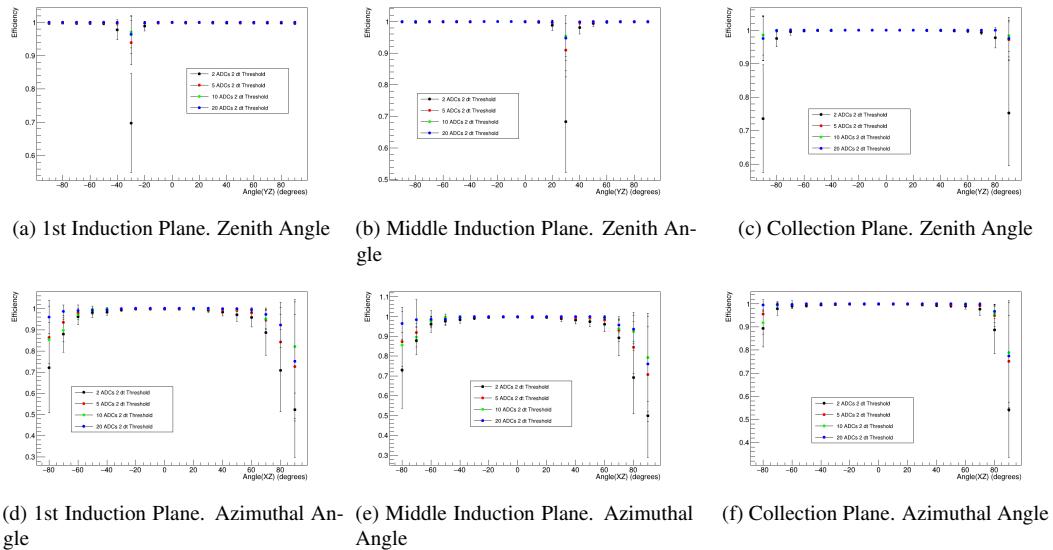


Fig. A.4 The 1st efficiency definition in Equation 6.1.2 for various MC ADC thresholds as a function of the azimuthal (XZ) and zenith (YZ) angles. The  $dt$  merging threshold is set to 2 ticks. Errors are the standard deviation of the efficiency for the corresponding angle.

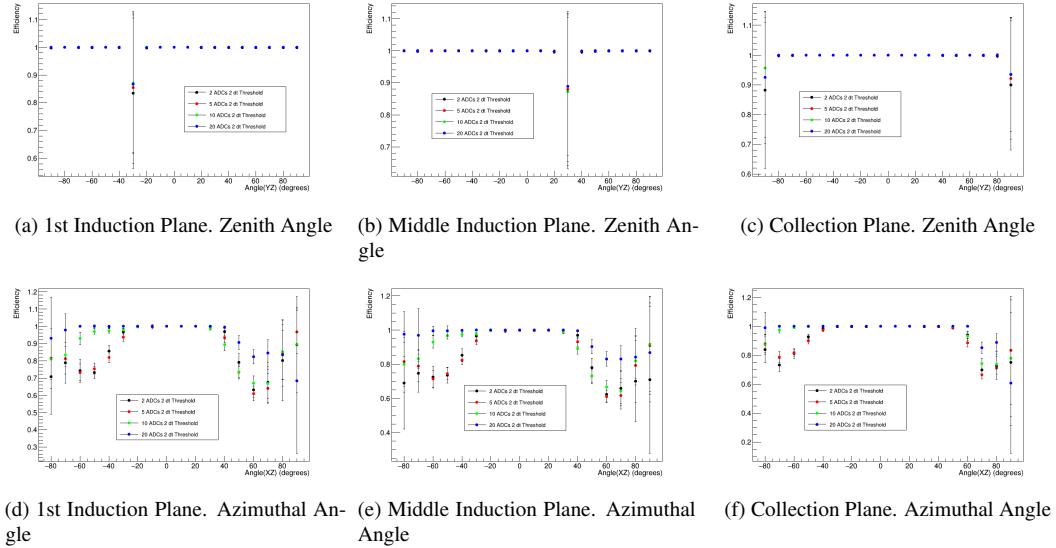


Fig. A.5 The 2nd efficiency definition in Equation 6.1.2 for various MC ADC thresholds as a function of the azimuthal (XZ) and zenith (YZ) angles. The  $dt$  merging threshold is set to 2 ticks. Errors are the standard deviation of the efficiency for the corresponding angle.

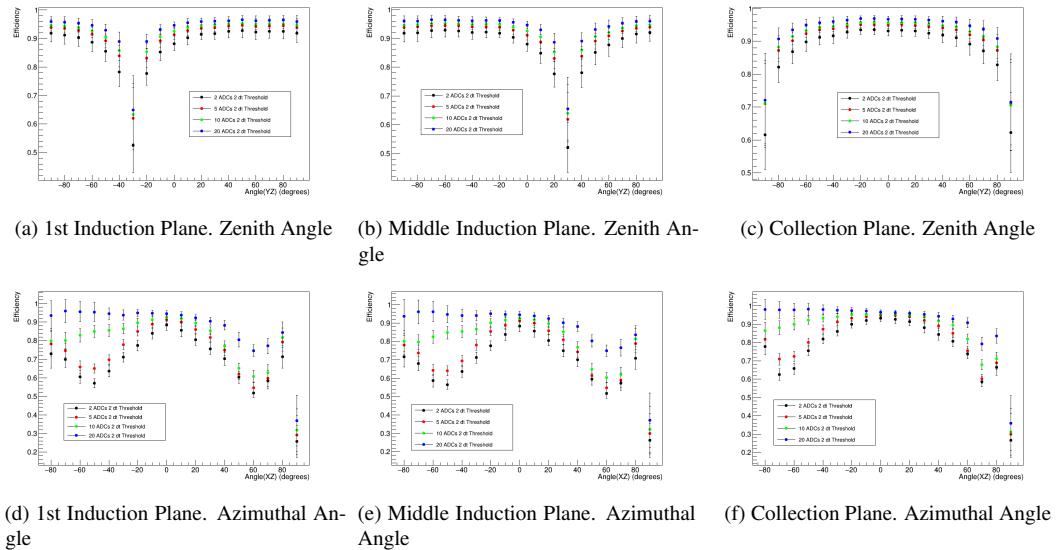


Fig. A.6 The 3rd efficiency definition in Equation 6.1.2 for various MC ADC thresholds as a function of the azimuthal (XZ) and zenith (YZ) angles. The  $dt$  merging threshold is set to 2 ticks. Errors are the standard deviation of the efficiency for the corresponding angle.

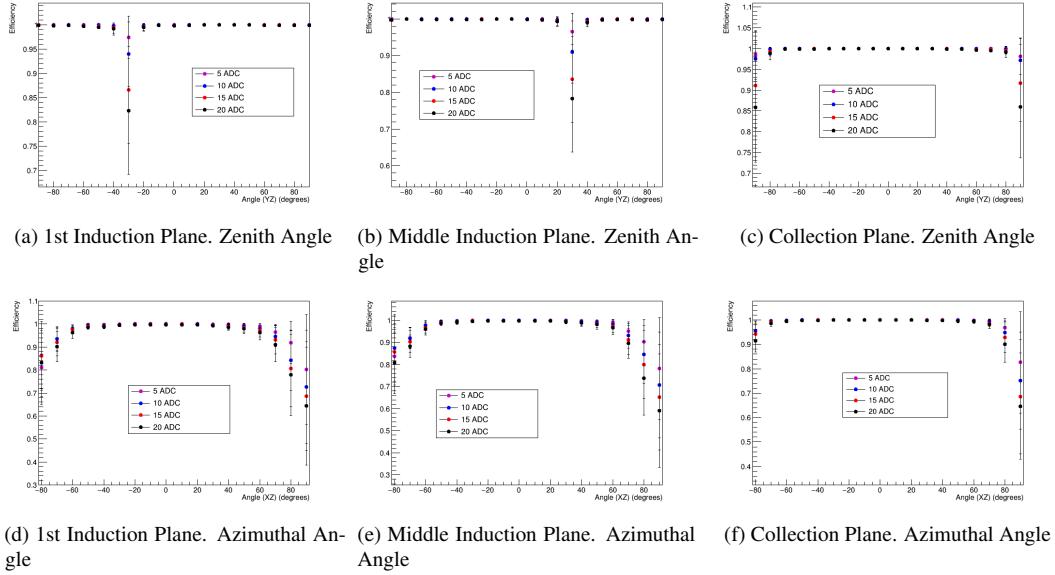


Fig. A.7 The 1st efficiency definition in Equation 6.1.2 for various ADC thresholds set in the *GausHit\_Module* as a function of the azimuthal (XZ) and zenith (YZ) angles. The MC ADC threshold is set to 5 ADC and the *dt* merging threshold is set to 2. Errors are the standard deviation of the efficiency for the corresponding angle.

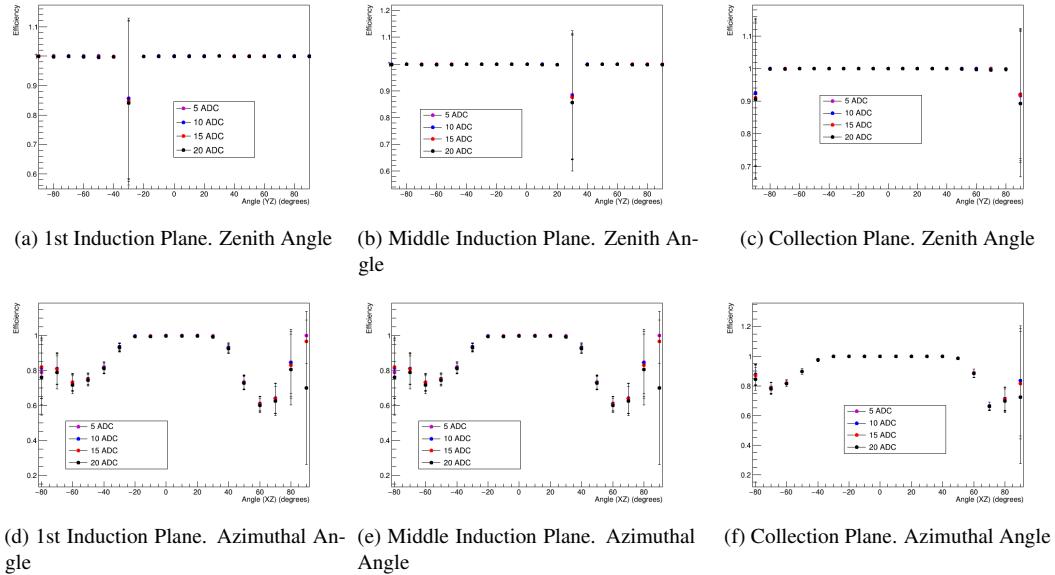


Fig. A.8 The 2nd efficiency definition in Equation 6.1.2 for various ADC thresholds set in the *GausHit\_Module* as a function of the azimuthal (XZ) and zenith (YZ) angles. The MC ADC threshold is set to 5 ADC and the *dt* merging threshold is set to 2. Errors are the standard deviation of the efficiency for the corresponding angle.

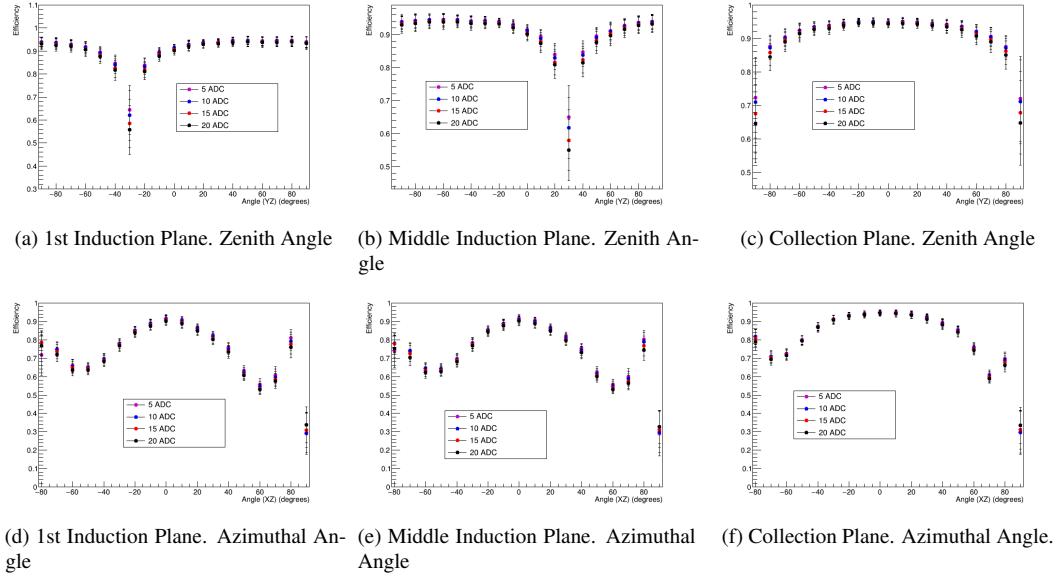


Fig. A.9 The 3rd efficiency definition in Equation 6.1.2 for various ADC thresholds set in the *GausHit\_Module* as a function of the azimuthal (XZ) and zenith (YZ) angles. The MC ADC threshold is set to 5 ADC and the *dt* merging threshold is set to 2. Errors are the standard deviation of the efficiency for the corresponding angle.

## Appendix B

### Details of the TRACS framework

More specifically, the base of the framework is a list of `art::Tools` [273] which modularise the shower characterisation. Each tool calculates properties known as elements of the shower. Elements which are values that make up a `recob::Shower` (e.g. start position) are known as properties whilst other elements are known as data product. An arbitrary number of tools can be run in the reconstruction and a second pass of the tools can be implemented. The shower tools derive from a C++ class known as `IShowerTool` such that the base driver module, `TRACS_module.cc`, does not have dependencies on the tools themselves and so does not have to be altered by the user. Therefore each tool overrides a set of base functions that are run in the driver module.

The primary function that is overwritten is the `CalculateElement` function. This is where data products and properties of the showers are calculated. `CalculateElement` takes as an input the `art::Event` so that the user has access to all the previous reconstructed and raw information. It is also given the `PFParticle` that is being characterised and the `ShowerElementHolder`. The structure of the `ShowerElementHolder` is shown in Figure B.1. The `ShowerElementHolder` is a class which stores data products such that users can access elements of any type that are calculated in downstream tools. To handle the elements in the holder, there exists a set of Get/Set/Check functions which use the unique name of the element to access the underlying data object. The driver module also accesses the shower properties in the `ShowerElementHolder` to create the `recob::Shower`. The objects stored in the `ShowerElementHolder` can be either of class type `ShowerDataProduct<T>` or `ShowerProperty<T, T2>`. As is shown in Figure B.2, the two data products are derived classes of `ShowerElementBase`. Elements defined as a `ShowerProperty` are set with a data object of type `T`, e.g. a vector of the shower energies calculated per wire plane and a corresponding error data object `T2`. The properties used to define the `recob::Shower` must be set as `ShowerProperty<T, T2>` objects. A `ShowerProperty<T, T2>` data object does

not have to be exclusively the properties set in the `recob::Shower`. However, the user can only store a `ShowerDataProduct<T>` in the `art::Event`.

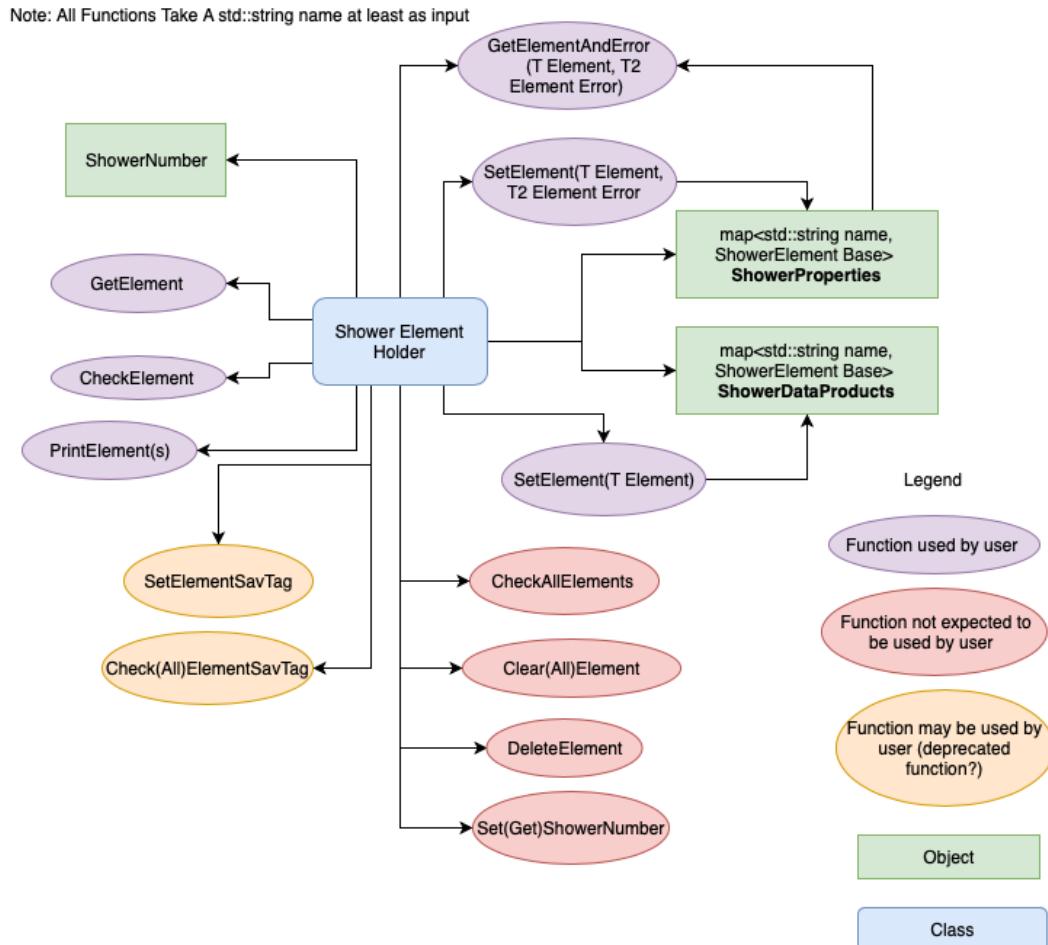


Fig. B.1 Structure of the `ShowerElementHolder` which the user directly interacts with.

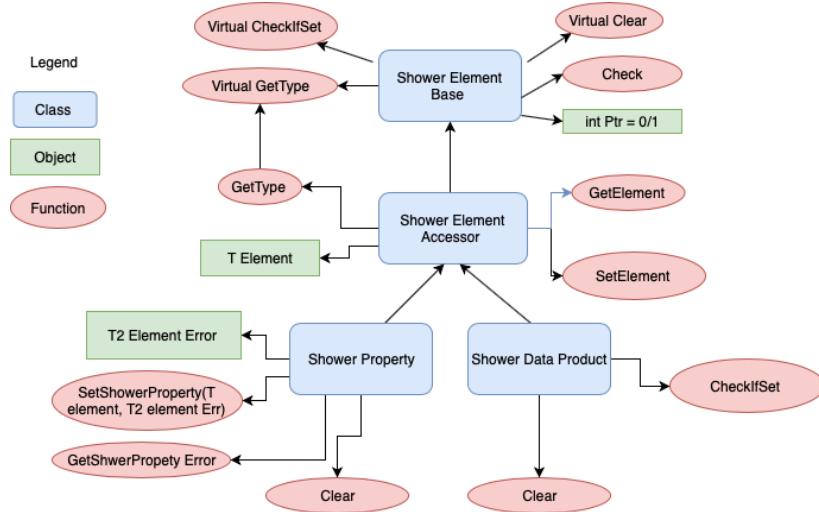


Fig. B.2 Structure of the `ShowerElementBase` which is held in the `ShowerElementHolder` and holds a specific data product the user has created.

A tool that calculates an element which is to be saved into the `art::Event` overrides the `InitialiseProducers` function. The user can employ the `InitialiseProduct<T>` function within `InitialiseProducers` to set up the required background art infrastructure to save the object to the `art::Event`, e.g. saving the `recob::Track` created from the initial track stub of the shower. To allow users to make associations (`art::Assns`) between data products, a function `AddAssociations` is available. This function allows downstream modules to match data products, e.g. the initial `recob::Track` stub and the `reco::Hits` used to create it. The class `ShowerProducedPtrHolder`, shown in Figure B.3, has been created to handle the products being stored. Data objects that are stored in the map are of a template class `ShowerDataProduct<std::vector<T>>`, whilst the associations are of the template class `ShowerDataProduct<T>`. The user never deals with `ShowerProducedPtrHolder` directly and a `Get` function, defined in the `IShowerTool` base class, is used to access the elements. At the end of the tool characterisation chain, the data products in the `ShowerElementHolder` which have been initialised in the `InitialiseProducers` functions are saved.

At each stage of the characterisation, it is also possible to create a 3D event display showing the current reconstructed objects. Figures 6.24 give an example of the event display. The event display exists in a TRACS algorithm, called `TRACAlg`, that also provides functions that are regularly used within the tools. There also exists a TRACS cheating suite which provides truth based results for the shower characteristics. This suite is useful for developing downstream tools and evaluating performance.

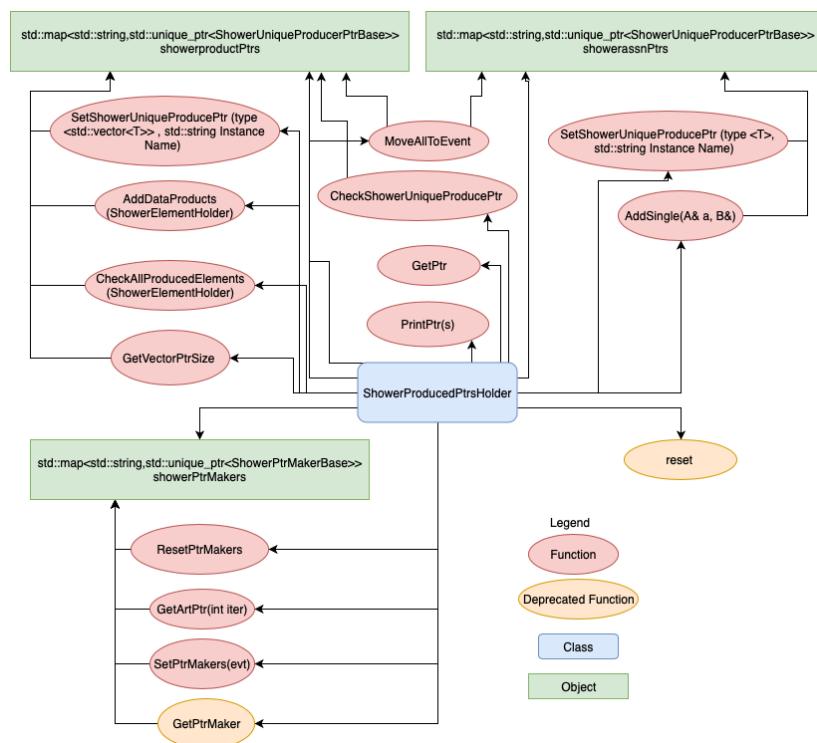


Fig. B.3 The structure of the `ShowerProducedPtrHolder` which handles the objects to be stored in the `art::Event`.

## Appendix C

### Differences Between GENIE version v2\_12 and v2\_8

In order to mimic the sample used in the SBN proposal [17], which was used to perform the  $\nu_e$  oscillation analysis, various weights were applied to account for difference between the GENIE version used for this thesis (v2\_12) and the one used for the proposal (v2\_8). The differences are discussed in Chapter 7 Section 7.1.2 and the following presents the weights that were applied to produce a proposal-era sample. The weights for a specific interaction type and neutrino energy is calculated by looking at the ratio in the event rates between the two versions. This weight was then applied to the event to scale the sample to the older simulation.

For the  $\nu_e$  selection, the modes that are important when considering the signal are the charged current events where the electron candidate is visible. Quasi-elastic CC events with zero final state charge pions (CC0Pi), are the most common signal event corresponding to approximately 62% (with MEC events removed) of the CC events. The version ratio is shown in Figure C.1. There is a 6.3% decrease in the total event rate of CC0Pi events from the BNB in the modern era (v2\_12) by comparison to the proposal era (v2\_8). This loss is relatively constant as a function of the neutrino energy.

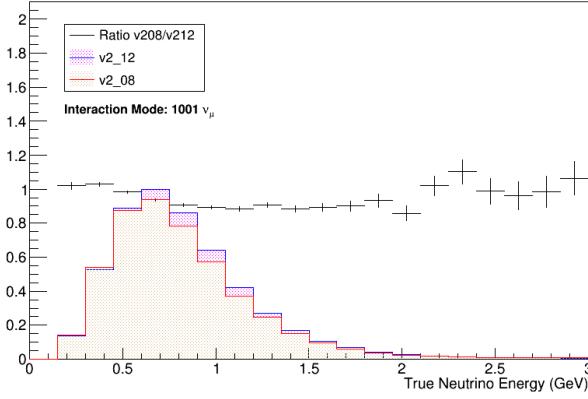
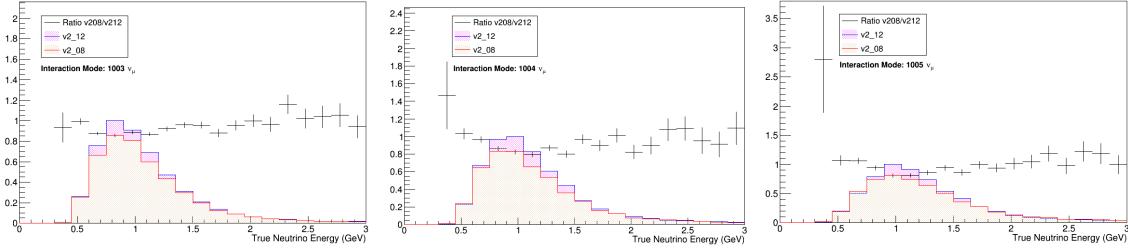


Fig. C.1 The interaction rates of CC0Pi events in GENIE v2\_8 and v2\_12. The errors on the plot are Poissonian.

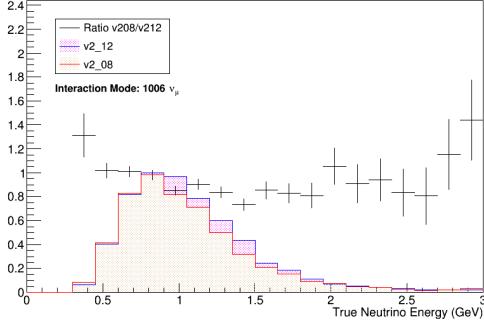
The ratios for resonant interactions where a charged pion is created are shown in Figures C.2a, C.2b and C.2c. The total resonant interactions correspond to 28.9% of the events and there is a 10.0% decrease in events in the proposal era compared to the modern era samples.



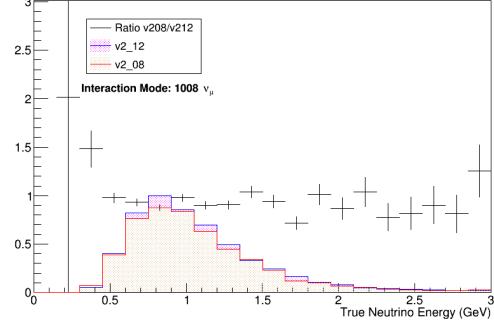
(a) CC resonant interactions on protons producing a final state positively charged pion.  
(b) CC resonant interactions on protons producing a final state neutral pion.  
(c) CC resonant interactions on neutrons producing a final state positively charged pion.

Fig. C.2 Figures showing the differences in rates of CC resonant interaction types from GENIE version v2\_12 and v2\_8. The errors arise from the Poissonian errors on the bin.

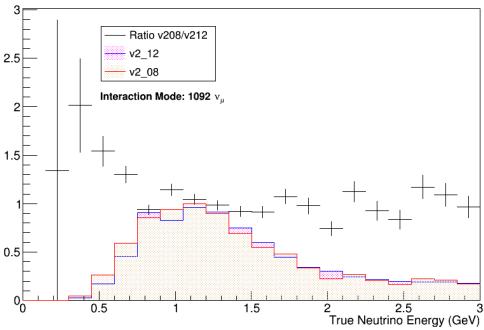
One of the main backgrounds for a  $\nu_e$  selection is photon showers arising from a final state neutral pion decay from neutral current interactions. Figures C.3a to C.3d correspond to NC interaction modes which can produce neutral pions in the final state. There is a total loss of 7.7% from the "modern" sample to the "proposal" sample for NC resonant events, with final state neutral pions. There is also a global 2.5% increase in Deep Inelastic Scattering (DIS) interactions, as seen in Figure C.3c, as well as an 11.3% increase in Coherent Scattering interactions, as seen in Figure C.3d. The error bars arise from the Poissonian errors on the bin and thus bins where there are low statistics have large errors. Thus there is larger errors on rarer interactions.



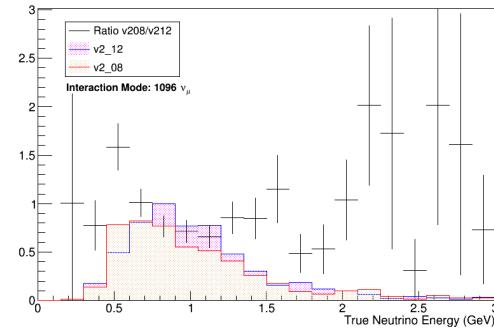
(a) NC resonant interactions on protons producing a final state neutral pion.



(b) NC resonant interactions on protons producing a final state neutral pion.



(c) NC deep inelastic scattering events.



(d) NC coherent scattering events.

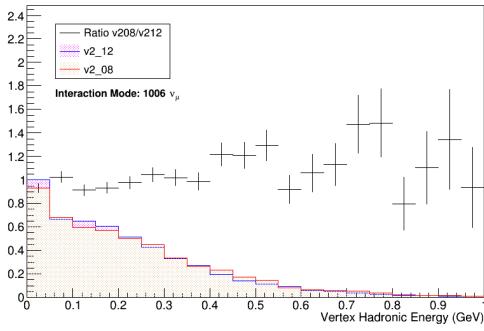
Fig. C.3 Figures showing the differences in rates of NC interactions between GENIE versions v2\_12 and v2\_8. Note that a neutral pion is not always created in DIS and coherent scattering events. The errors arise from the Poissonian errors on the bin.

The differences in the GENIE versions are considered for interactions with  $\nu_\mu$  particles. However, the weights shown are applied to both  $\nu_\mu$  and  $\nu_e$  interactions. Additional studies should be designed to account for differences in  $\nu_e$  interactions.

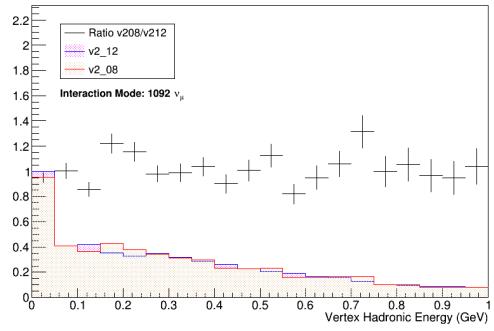
By applying the weights above the sample produced were then used in the truth-based proposal like selection as the "proposal" sample.

Additional studies were made to identify changes in the rate and energy of final state particles. Hadronic final state particles are used to identify the neutrino vertex. Identifying the vertex correctly is useful for removing the NC background. This is because photons can travel several centimetres in liquid argon without interacting. As the photons do not ionise the liquid argon, there can be a physical gap, known as the conversion distance, between the neutrino vertex and the shower vertex. This is not the case for the electron shower and hence the events can be removed if the conversion gap is visible. Figures C.4a and C.4b show the ratio between the total final state hadronic energy at the vertex of the different GENIE

versions, after the differences in the interaction rates described above are accounted for. The interaction rate corrections are applied for all other final state particle properties below as well. Figure C.4a corresponds to the NC resonant events where the neutrino interacts with a proton and a final state neutral pion is produced. Figure C.4b corresponds to events that underwent a NC DIS interaction. The ratio of the hadronic energy at the vertex is approximately consistent between the two versions.



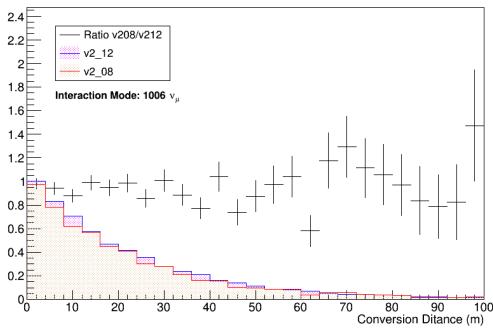
(a) NC resonant interactions on protons producing a final state positively charged pion.



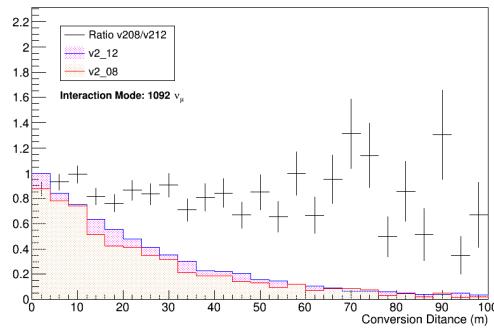
(b) NC DIS interactions.

Fig. C.4 Figures showing the differences in the hadronic energy of NC interactions between GENIE versions v2\_12 and v2\_8. The errors arise from the Poissonian errors on the bin.

The difference between the conversion distance of photons for NC resonant interactions, Figure C.5a, and NC DIS interaction, Figure C.5b, show the GENIE versions are also approximately consistent.



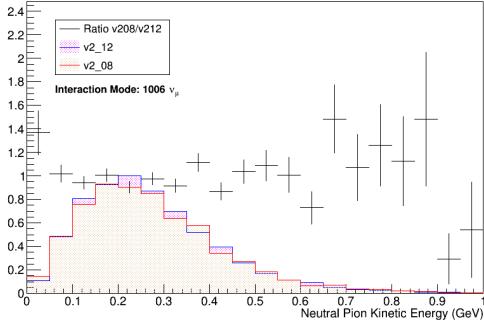
(a) CC resonant interactions on protons producing a final state positively charged pion.



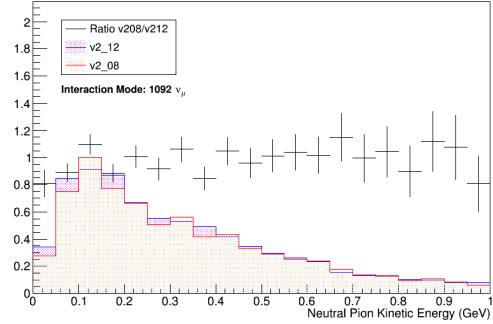
(b) NC DIS interactions.

Fig. C.5 Figures showing the differences in the conversion distance of the most energetic photon arising from vertex interactions of NC interactions between GENIE versions v2\_12 and v2\_8. The errors arise from the Poissonian errors on the bin.

Also, outgoing neutral pion energies for interaction of type 1006, Figure C.6a, and for interaction type 1092, Figure C.6b, were shown to be consistent between the two versions.



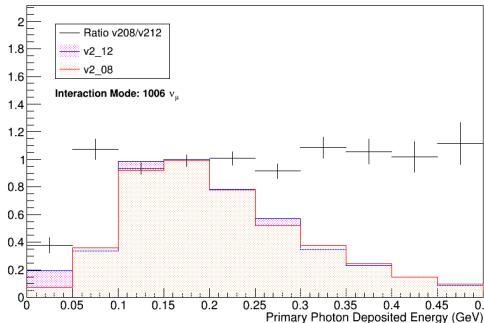
(a) CC resonant interactions on protons producing a final state positively charged pion.



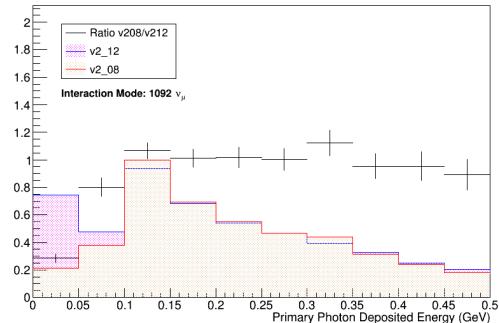
(b) NC DIS interactions.

Fig. C.6 Figures showing the differences in the energy of a final state neutral pion in NC interactions between GENIE versions v2\_12 and v2\_8. The errors arise from the Poissonian errors on the bin.

In addition, changes in the photon energy distributions were considered. Figures C.7a and C.8a, for interaction type 1006, and Figures C.7b and C.8b, for interaction type 1092, show the changes in the primary photon energy and the secondary photon energy respectively. The primary and secondary photons are the photons with the highest and second highest energy arising from the vertex respectively. Changes in the primary photon energy are important when considering changes in the background selection spectra. This is because primary photons are misidentified as electron arising from a CC interaction. Changes in the secondary shower energies affect the visibility of the second shower which is used to identify NC pion events. The Figures show the GENIE versions are roughly consistent over the photon energy



(a) CC resonant interactions on protons producing a final state positively charged pion.



(b) NC DIS interactions.

Fig. C.7 Figures showing the differences in the energy of the most energetic photon arising from the vertex in NC interactions between GENIE versions. The errors the Poissonian.

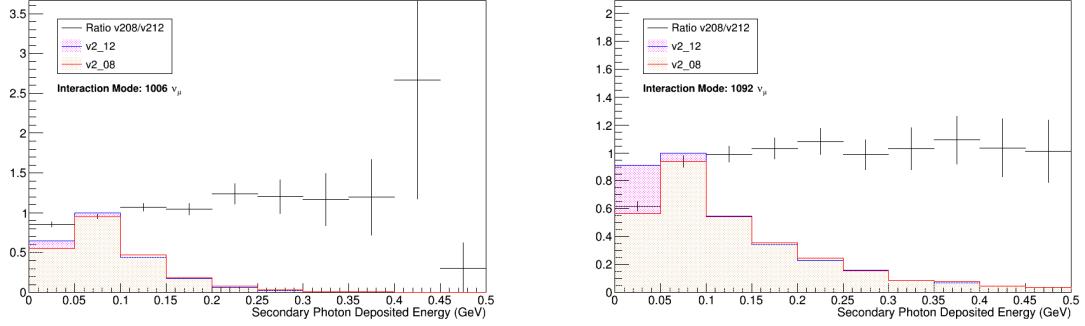


Fig. C.8 Figures showing the differences in the energy of the second most energetic photon arising from the vertex of NC interactions between GENIE versions v2\_12 and v2\_8. The errors arise from the Poissonian errors on the bin.

Other interaction modes were also considered in the analysis. For all important modes the properties of the final state particles are roughly consistent between the GENIE versions. Therefore, no additional weights were applied. A more in-depth analysis with dedicated productions of rare modes should be considered to verify these findings.

## Appendix D

# Including a SVM for PFO Characterisation

Although the events were reconstructed using the methods described in Chapter 6, a Support Vector Machine (SVM) [260] was added at the particle characterisation stage of the Pandora pattern recognition [27], discussed in Chapter 6, Section 6.1.3. This SVM was implemented to improve the characterisation of proton tracks and energetic showers.

The effects of including the SVM can be seen in Figure D.1 for proton identification and Figure D.2 for shower identification. The analysis was performed on the  $\nu_e$  events after requiring a PFParticle in the active volume and a reconstructed shower. The SVM alters the identification of clusters as showers or tracks. However, the SVM does not improve the clustering procedure or the overall reconstruction efficiency.

Figure D.1 shows the improvement in identifying protons as track objects in the oscillated neutrino events. Specifically, the Figure shows an efficiency of 80% is achieved at  $\sim 100$  MeV compared to the cut-based method at  $\sim 200$  MeV. This change is due to the correct identification of low energy protons as tracks rather than showers using the SVM. This increase in reconstruction efficiency improves the correctness of calorimetry and identification methods. Furthermore, this efficiency increase prevents events being removed due to selection cuts on the number of showers.

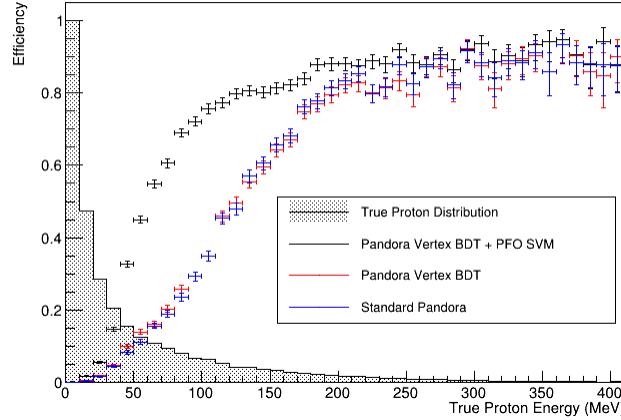


Fig. D.1 The proton reconstruction efficiency for the different Pandora pattern recognition reconstruction pathways, discussed in Chapter 6, Section 6.1.3, as a function of true energy deposited by the proton. The true deposited energy distribution is plotted (shaded black). The errors on the figures are the  $1\sigma$  Clopper-Pearson intervals [32].

Additionally, the SVM is more successful at correctly characterising the bulk of the showers. This is shown in Figure D.2, where the average completeness for the higher energy showers improves. This completeness gain improves the selection efficiency, e.g. on a shower energy cut. Furthermore, the reconstruction of the shower and neutrino energy is improved.

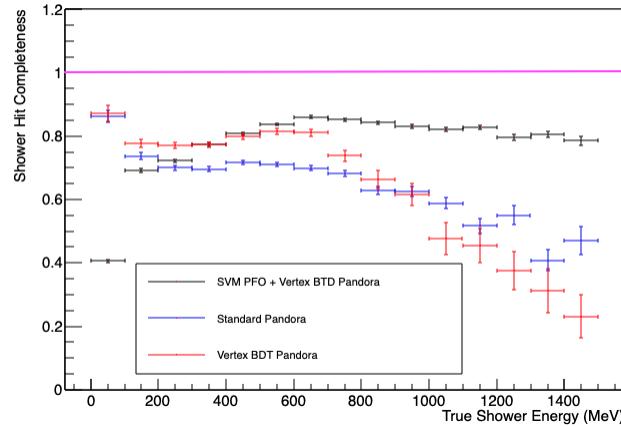


Fig. D.2 The mean shower hit completeness for the different Pandora pattern recognition reconstruction pathways, discussed in Chapter 6, Section 6.1.3, as a function of true shower energy. The errors in  $y$  are the standard error on the mean for values within the true energy range  $x - x_{err} \rightarrow x + x_{err}$  where  $x_{err}$  corresponds to the error bar bar on the  $x$  axis. The pink line corresponds to the ideal value.

## Appendix E

# Further Analysis on the Pre-selection Cuts

The first cut removed events where a reconstructed PFP neutrino did not exist in the drift window. The efficiency for selecting  $\nu_e$  Charge Current (CC) and Neutral Current (NC) and removing  $\nu_\mu$  CC and NC events is shown in Figures E.1. The Figures show that some NC events are not reconstructed because NC events do not necessarily have to deposit energy in the TPC.

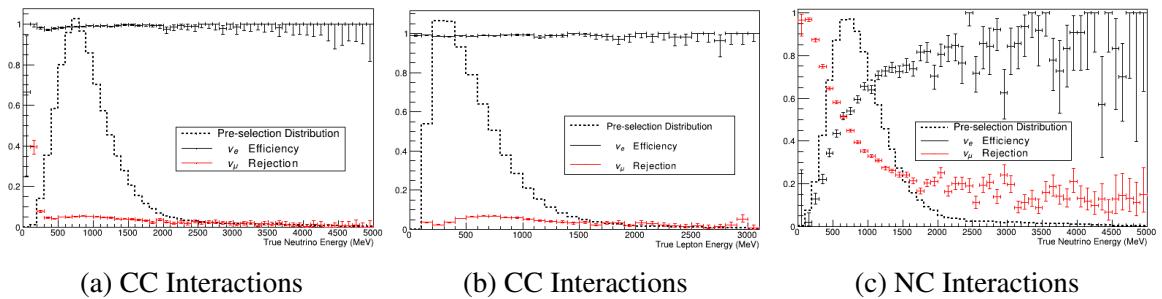


Fig. E.1 The  $\nu_e$  efficiency and  $\nu_\mu$  background rejection for the CC interactions (a) and NC interactions (c) when removing events with no reconstructed PFP neutrino. b) The CC interactions efficiency and background rejection as a function of the outgoing charged lepton energy. The underlying neutrino distribution is also plotted. The errors on the figures are the  $1\sigma$  Clopper-Pearson intervals [32].

NC events where final state particles are visible in the detector, such as resonant and deep inelastic scattering, are more likely to be reconstructed due to the presence of charge depositions. This is shown in Figure E.2c and E.2d where the reconstruction efficiency for the resonant mode and DIS (including both CC and NC) are not degraded, unlike the elastic NC shown in Figure E.2b. As the elastic NC events are the dominant contribution, the

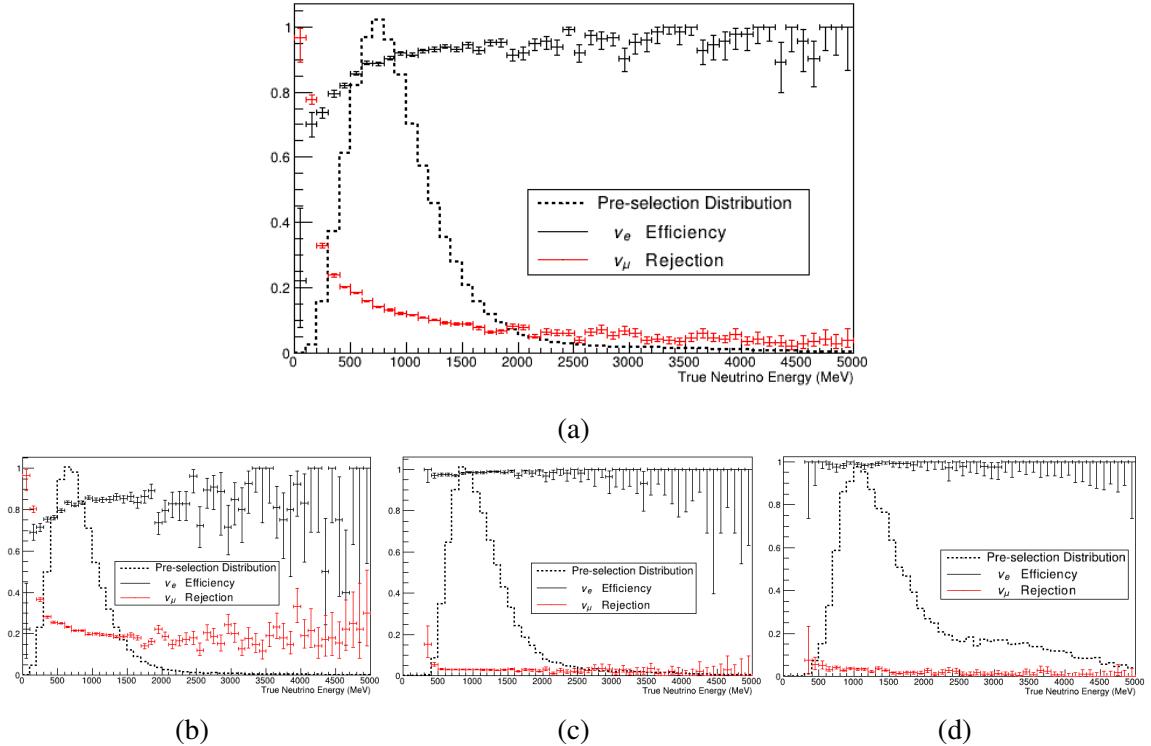


Fig. E.2 The total oscillated  $\nu_e$  efficiency and  $\nu_\mu$  background rejection when removing events with no PFP reconstructed neutrino. The efficiency is split between the different interaction modes, a) Elastic , b) Resonant and c) DIS, with the underlying initial neutrino energy distribution also plotted. The errors on the Figures are the  $1\sigma$  Clopper-Pearson intervals [32].

removal of elastic NC events is observable. This is demonstrated in Figure E.2a which shows a noticeable drop in the total  $\nu_e$  and  $\nu_\mu$  efficiency. As can be seen in Figures E.1a and E.1b, the efficiency for CC  $\nu_e$  is  $\sim 100\%$  for all neutrino energies and final state lepton energies.

The second pre-selection cut requires that the event contains at least one reconstructed shower in the drift window. This is a straight forward requirement of a  $\nu_e$  CC selection as the identifiable part of the interaction is the CC showering electron. Also performing the cut significantly removes Quasi-Elastic (QE)  $\nu_\mu$  interactions, which is demonstrated in Figure E.3b. It also removes some of the resonant and DIS interactions where a photon is not present, demonstrated in Figure E.3c and Figure E.3d respectively. The remaining resonant and DIS interactions with photons present decrease the background rejection in the higher energy regions of the total CC and NC spectra, where the interaction modes become dominant.

As can be seen in Figure E.3a, there is  $76.4^{+0.2}_{-0.2}\%$  total selection efficiency in  $\nu_e$  interactions. Figure E.4c shows the removal of NC interactions, which have already been significantly removed during the first PFP neutrino cut. Therefore, the NC events only

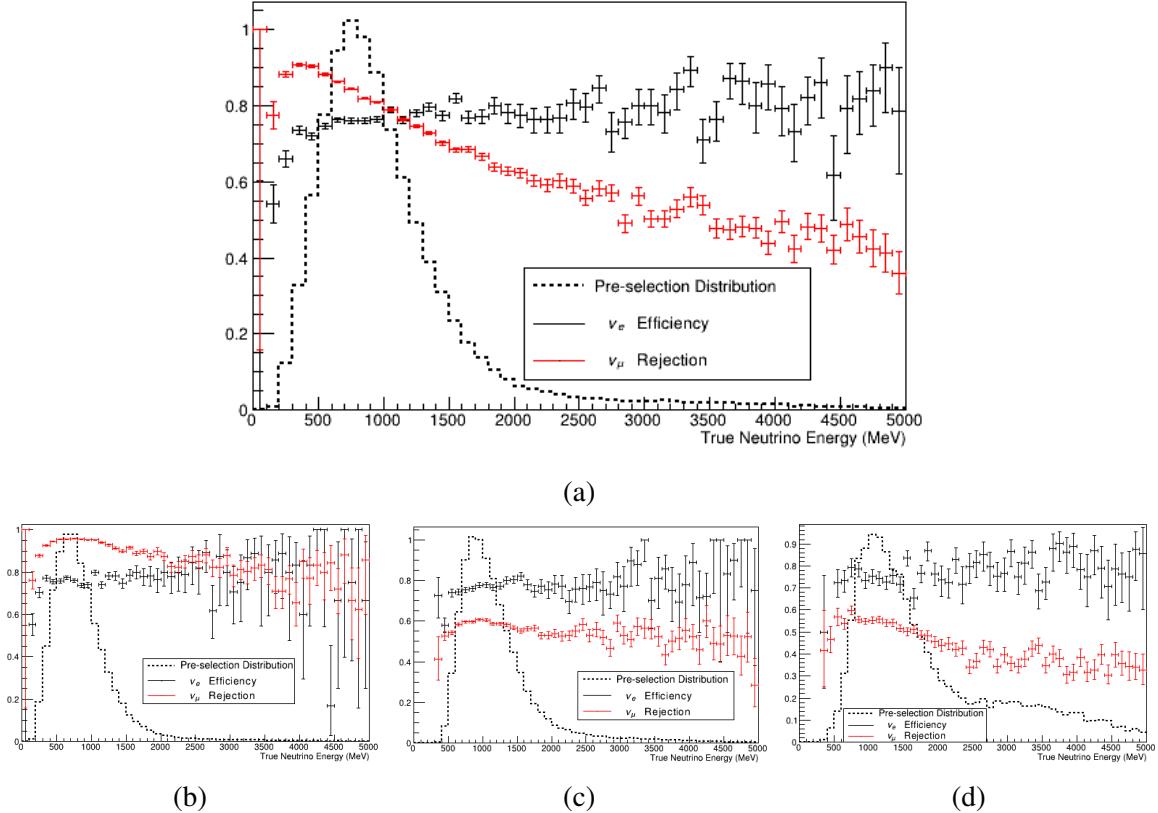


Fig. E.3 a) The total oscillated  $\nu_e$  efficiency and  $\nu_\mu$  background rejection when removing events with no reconstructed shower after the PFP neutrino cut. The efficiency is split between the different interaction modes a) QE, b) Resonant and c) DIS, with the underlying initial neutrino energy distribution also plotted. The errors on the figures are the  $1\sigma$  Clopper-Pearson intervals [32].

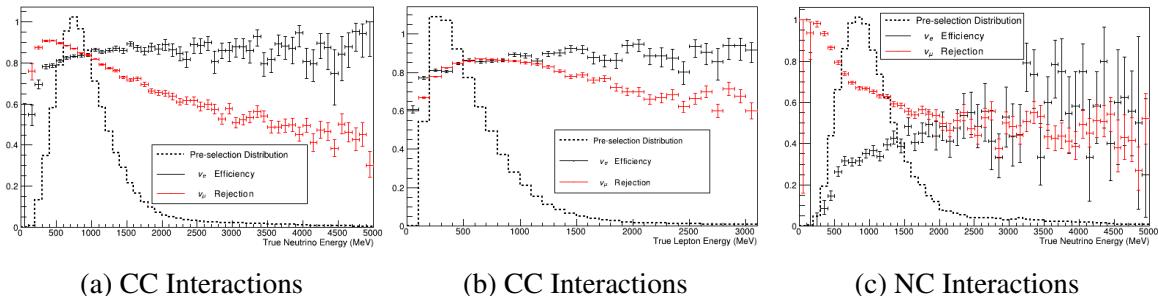


Fig. E.4 The  $\nu_e$  efficiency and  $\nu_\mu$  background rejection for the CC interactions and NC interactions for removing events with no reconstructed shower as a function of the true incoming neutrino energy (a) and as a function of the outgoing charged lepton energy (b). The underlying neutrino distribution is also plotted. The errors on the figures are the  $1\sigma$  Clopper-Pearson intervals [32].

contribute to a  $\sim 3\%$  removal of  $\nu_e$  events. Figure E.4a shows the CC contribution which accounts for the remaining 15% of the loss of  $\nu_e$  events. This is due to poor reconstruction and topological issues. However, some events would have been removed during a fiducial volume cut. As can be seen in Figures E.4a and E.4b, the signal efficiency increases as a function of the neutrino and lepton energy. The final  $\nu_e$  CC signal efficiency and the  $\nu_\mu$  background rejection. The resulting efficiency is  $83.5^{+0.2}_{-0.2}(\text{stat})\%$  and background rejection is  $79.2^{+0.2}_{-0.2}(\text{stat})\%$ .

## Appendix F

# Neutrino Energy Reconstruction and Correctness

Prior to the full selection, the neutrino energy was reconstructed. This was evaluated so that the relative difference between the reconstructed and true neutrinos for the signal  $\nu_e$  CC events could be corrected for.

Combining the sum of the energies of particles in the PFP hierarchy, which were reconstructed in the calorimetry, see Chapter 6, Section 6.1.7, was carried out in order to reconstruct the neutrino energy. As discussed in Chapter 6, Section 6.2.2, there is a deficit of 14.5% in the reconstructed shower energy. Therefore, shower-like particles are scaled by a factor of 1.145. Figure F.1a shows the energy resolution for showers from  $\nu_e$  events that pass the first two pre-selection cuts. The Figure shows a deficit of 17.1% (before scaling) in the reconstructed shower energy. It is assumed that the vertex sample provides a better description of the selected events. Note: the scale factor is not applied to the shower energy cuts described in the upcoming Sections. There is also a 16.1% energy deficit, as seen in Figure F.1b, in reconstructed energy of the track-like particles present in events that pass the first two pre-selection cuts for the  $\nu_\mu$  background. Therefore, the energy is scaled by a factor of 1.161 for all track-like particles. Applying the track and shower corrections results in Figure F.1c, which shows the fractional neutrino energy correctness for  $\nu_e$  CC QE events that pass the first two pre-selection cuts and are also in the fiducial volume. As can be seen in the Figure, there is an additional energy deficit of 14.13%. Hence the final neutrino energy is scaled by 1.143 to correct for the losses.

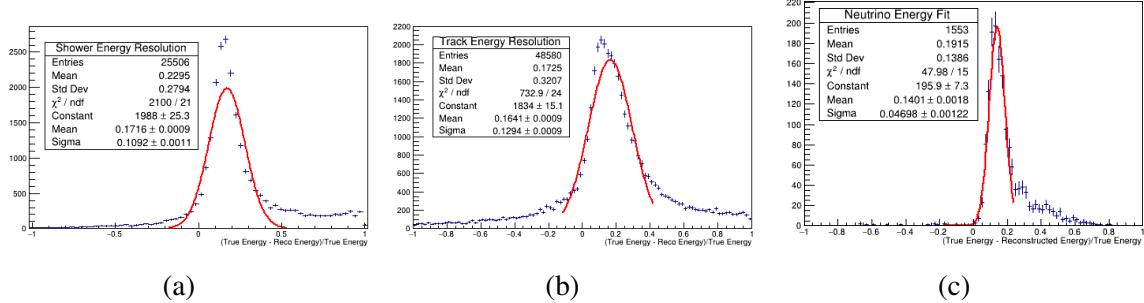


Fig. F.1 The energy reconstruction correctness for tracks (a), showers (b) and neutrinos (c). The neutrino reconstructed energy is calculated after corrections have been applied to the track and shower for only events above the reconstructed energy of 225 MeV. The  $\nu_\mu$  sample is used for the tracks and the  $\nu_e$  sample for the showers. All samples used underwent the no PFP neutrino and no reconstructed shower cuts. Errors on the graphs are statistical.

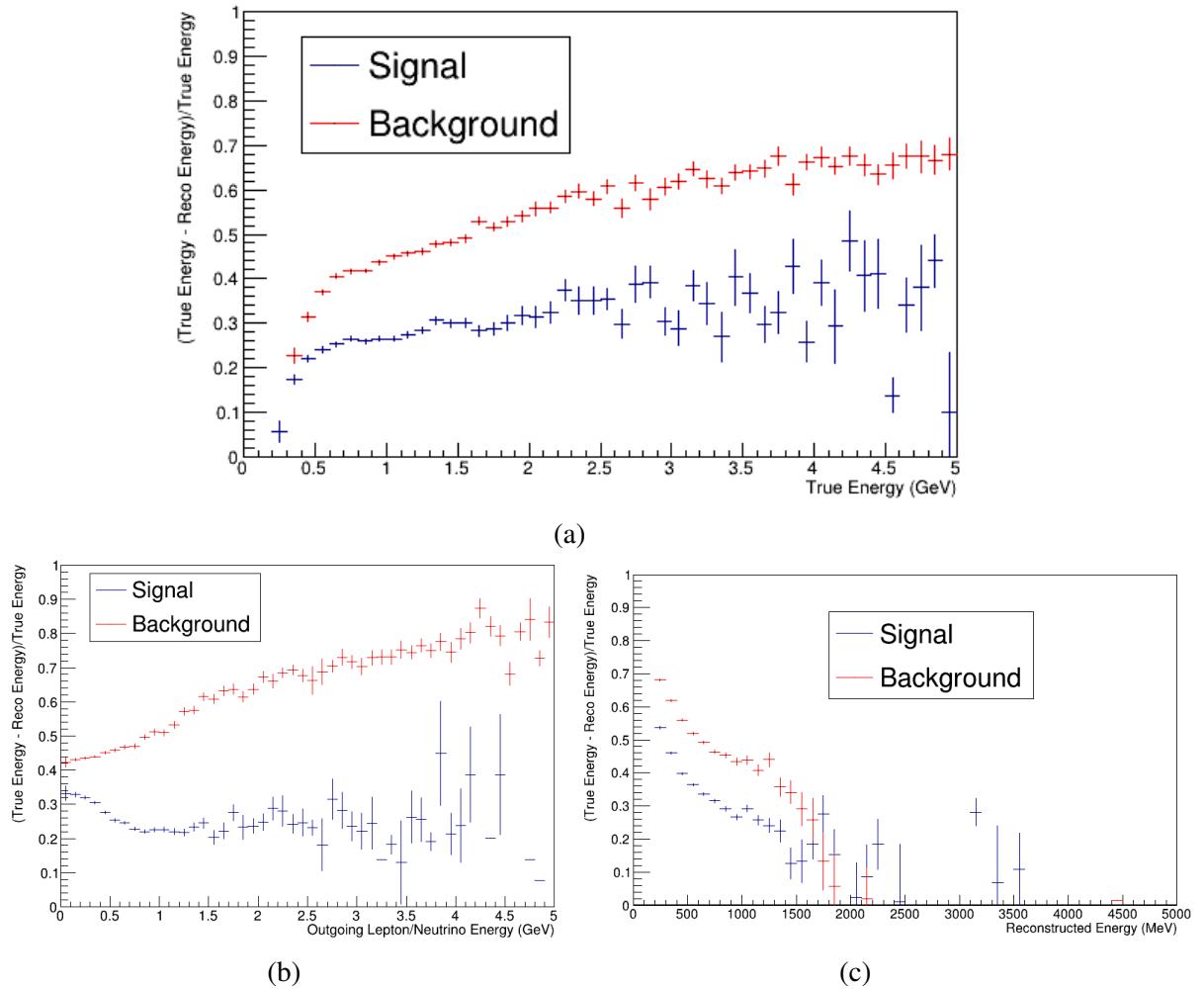
The neutrino energy correctness, before the 1.143 scaling and after the pre-selection cuts, is also shown as a function of the following variables [320] (see Figures F.2):

- Outgoing charged lepton (for CC interactions) or neutrino (for NC interactions) energy of the event.
- $\theta$  (theta) The angle between the neutrino and the outgoing charged lepton/neutrino.
- $Q^2$  The momentum transfer from the neutrino to the target particle ( $Q^2 = -q^2$  (the momentum of the propagator)).
- The transverse momentum of the incoming neutrino.
- The hadronic invariant mass  $W = \sqrt{M^2 + 2Mv - Q^2}$ , where  $M$  is the mass of the target particle and  $v = E_v - E_l$  is the energy transferred via the propagator.
- The Bjorken  $x = \frac{Q^2}{2Mv}$ .
- The inelasticity  $y = 1 - \frac{E_l}{E_v}$ .

Note that the neutrino energy has not been corrected in the plots, but the shower and track energies have. Also, a 225 MeV reconstruction energy cut has been applied to the Figures to remove events which have been poorly reconstructed. All the interaction kinematics, shown in Figures F.2f, F.2d, F.2e and F.2i, show that as more energy is given to the hadronic part of the interaction then more of the energy is lost in the reconstruction. This is to be expected as more energy can be lost through FSI nuclear effects, nuclear recoil and neutral particles (i.e. neutrons). In addition, more of the energy is lost in the background events compared to the

signal events. This is because the majority of the background events are NC and therefore, as can be seen in Figure F.2b, the larger the outgoing neutrino energy the more energy is not visible. Furthermore, as the neutrino energy increases, more complex interactions occur, such as resonant and DIS events. Such interactions create more particles which can be missed during the reconstruction. Therefore, the energy correctness decreases as a function of the neutrino energy.

The directionality of the lepton, which is shown in Figure F.2g, affects the reconstruction efficiency as events travelling into the wire planes are harder to reconstruct and the pattern recognition assumes events travel in the direction of the beam. Therefore, more energy in the signal events is lost due to poor reconstruction.



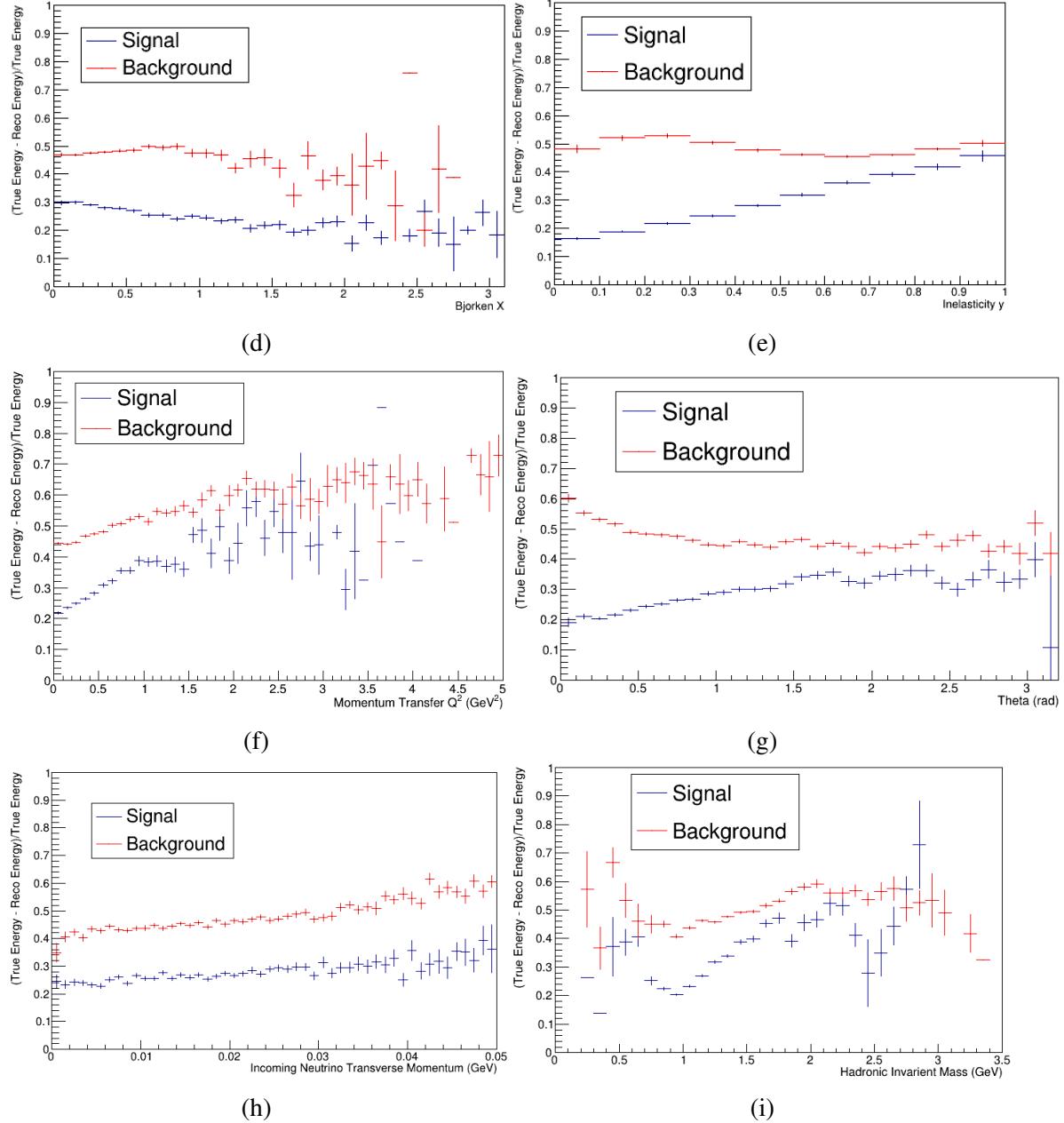


Fig. F.2 Neutrino energy correctness of CC  $\nu_e$  signal (blue) events and the  $\nu_\mu$  (red) background events as a function of the neutrino interaction variables: a) the true neutrino energy, b) the outgoing charged lepton/neutrino energy, c) the reconstructed neutrino energy, d) the Bjorken parameter  $x$ , e) the inelasticity  $y$ , f) the momentum transfer  $Q^2$ , g) the angle between the outgoing charged lepton/neutrino and the incoming neutrino, h) the incoming transverse momentum, and i) the hadronic invariant mass. The errors on the y-axis are the errors on the mean from the distribution of values between  $x - \delta x \rightarrow x + \delta x$ . All samples used underwent the no PFP neutrino and no reconstructed shower cuts and only the track and shower energies have been corrected for. A 225 MeV reconstructed energy cut has been applied to the Figures to remove poor reconstruction.

## Appendix G

### Discussion of a Truth-Based $dE/dx$ study

The author would like to give credit to Ed Tyley for the following results. Figures G.1 show the results from a truth-based  $dE/dx$  study for events from the BNB simulation and the vertex simulation. The true  $dE/dx$  is calculated from the energy deposited within 3 cm of the shower start position. The median  $dE/dx$  on a wire is then taken as the shower  $dE/dx$ . This method mimics the current standard shower reconstruction methods, discussed in Chapter 6, Section 6.2.2. Due to the change in energy, there is an increase of 2.9% from the vertex sample to the BNB sample in photon Compton scattering, which is indistinguishable from an electron shower  $dE/dx$ . Additionally, from the vertex events to the BNB events there is an increase of 8.9% in pair production events where the  $dE/dx$  is less than 3 MeV/cm. This is due to one of the electrons in the photon pair stopping within the initial track stub. The Bayesian truncating tool, discussed in Chapter 6, Section 6.2.2, has the potential to identify these cases. However the tool is limited if the lower energy electron only travels along a couple of wires. In the truth analysis, 5% of BNB photon events were Compton scatters, of which 93.8% were below 3 MeV/cm. 12.6% of pair production photons in the BNB events were below 3 MeV/cm resulting in a total of 83.3% background rejection at 3 cm. This value is the limit of the reconstruction using the current method. 95.4% of the electron events were selected. Therefore, a significant change to the  $dE/dx$  calculation is required to reach the 100% efficiency and 94% background rejection suggested in the proposal truth-based analysis [246].

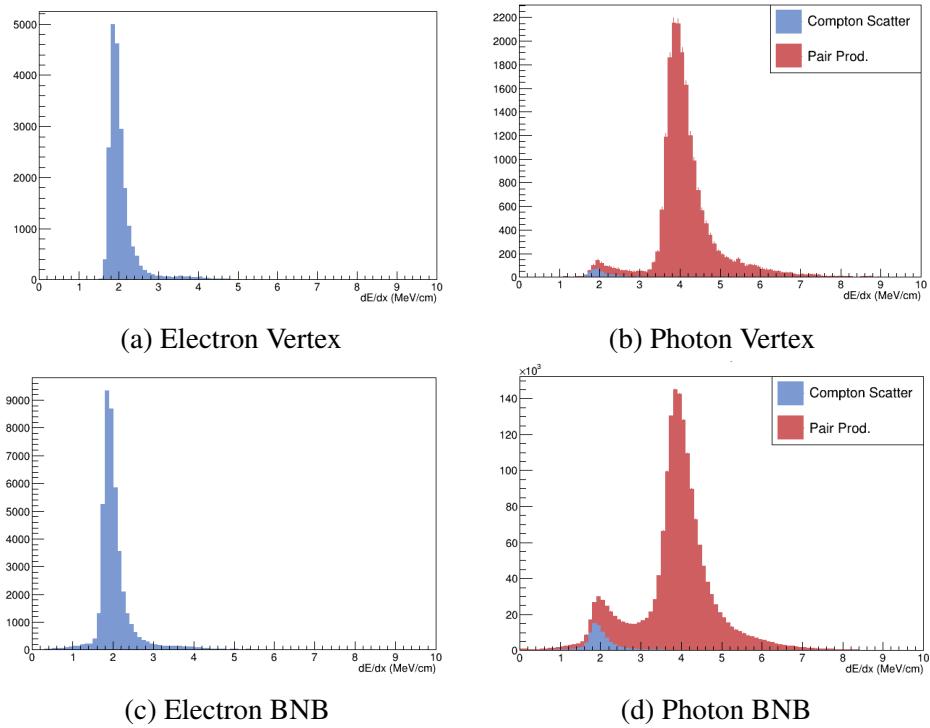


Fig. G.1 The truth-based  $dE/dx$  distributions from electrons (a) and photons (b) from the vertex simulated sample as well as electrons (c) and photons (d) from the BNB sample. The photon samples are stacked and split into the different end processes of the photon, Compton scattering (blue) and pair production (red). Credit: Ed Tyley.

## Appendix H

# $\nu_e$ CC Selection as a Function of Interaction Variables

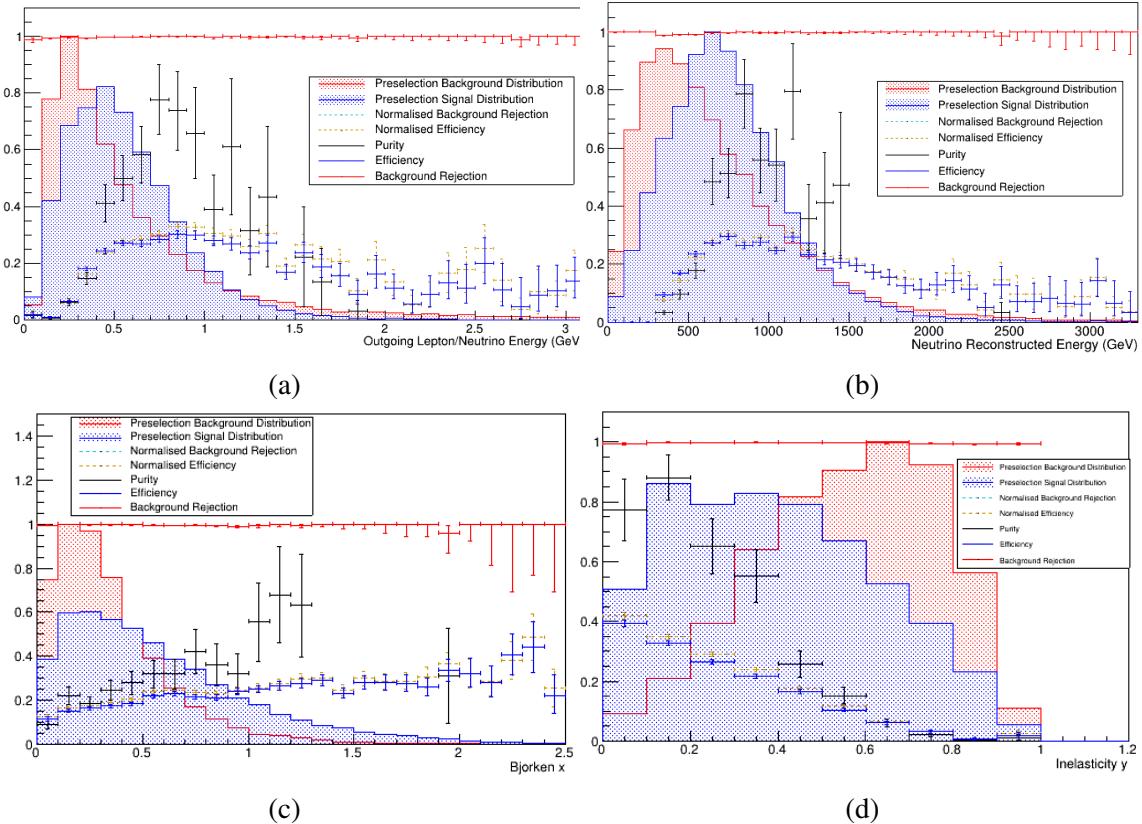
The following present the selection efficiencies of the proposal style selection (described in Chapter 7, Section 7.4.3), the new cut-based selection (described in Chapter 7, Section 7.4.4) and a Boosted Decision Tree (BDT) selection (described in Chapter 7, Section 7.4.5). All the selections were performed on the same simulated reconstructed sample which had undergone the pre-selection cuts described in Chapter 7, Section 7.4.2). The losses due the pre-selection cuts are not accounted for in the figures and the reader should follow Chapter 7 and Appendix E to understand the full selection efficiencies.

The efficiencies are presented as function of the following variables [320]:

- Outgoing charged lepton (for CC interactions) or neutrino (for NC interactions) energy of the event.
- $\theta$  (theta) The angle between the neutrino and the outgoing charged lepton/neutrino.
- $Q^2$  The momentum transfer from the neutrino to the target particle ( $Q^2 = -q^2$  (the momentum of the propagator)).
- The transverse momentum of the incoming neutrino.
- The hadronic invariant mass  $W = \sqrt{M^2 + 2Mv - Q^2}$ , where  $M$  is the mass of the target particle and  $v = E_v - E_l$  is the energy transferred via the propagator.
- The Bjorken  $x = \frac{Q^2}{2Mv}$ .
- The inelasticity  $y = 1 - \frac{E_l}{E_v}$ .

### Proposal Style Selection

The efficiency curves are depicted as a function of the true interaction variables in Figures H.1. These Figures show the efficiency curves when maximising the efficiency  $\times$  background rejection. It can be seen that the efficiency increases as a function of the outgoing charged lepton/neutrino energy. This is firstly due to the lepton becoming more visible in the CC signal events. Secondly, more of the energy is taken out of the event by the neutrino in the background NC cases. This effect is also visible in the inelasticity parameters:  $y$ , shown in Figure H.1d;  $Q^2$ , shown in Figure H.1e; and  $x$ , shown in Figure H.1c. For the CC background, the muon becomes more visible as the inelasticity decreases. This is because the muon has more energy and travels further. Hence, the event will be removed via the length cut. The background, therefore, is reduced for both elastic and inelastic events. The efficiency also decreases as a function of the hadronic invariant mass, see Figure H.1h. This is because as the hadronic invariant mass increases, more pions are created. This results in more photon showers (from  $\pi^0$  decays) and long charged pion tracks. The one shower cut and the track length cut then decrease the selection efficiency.



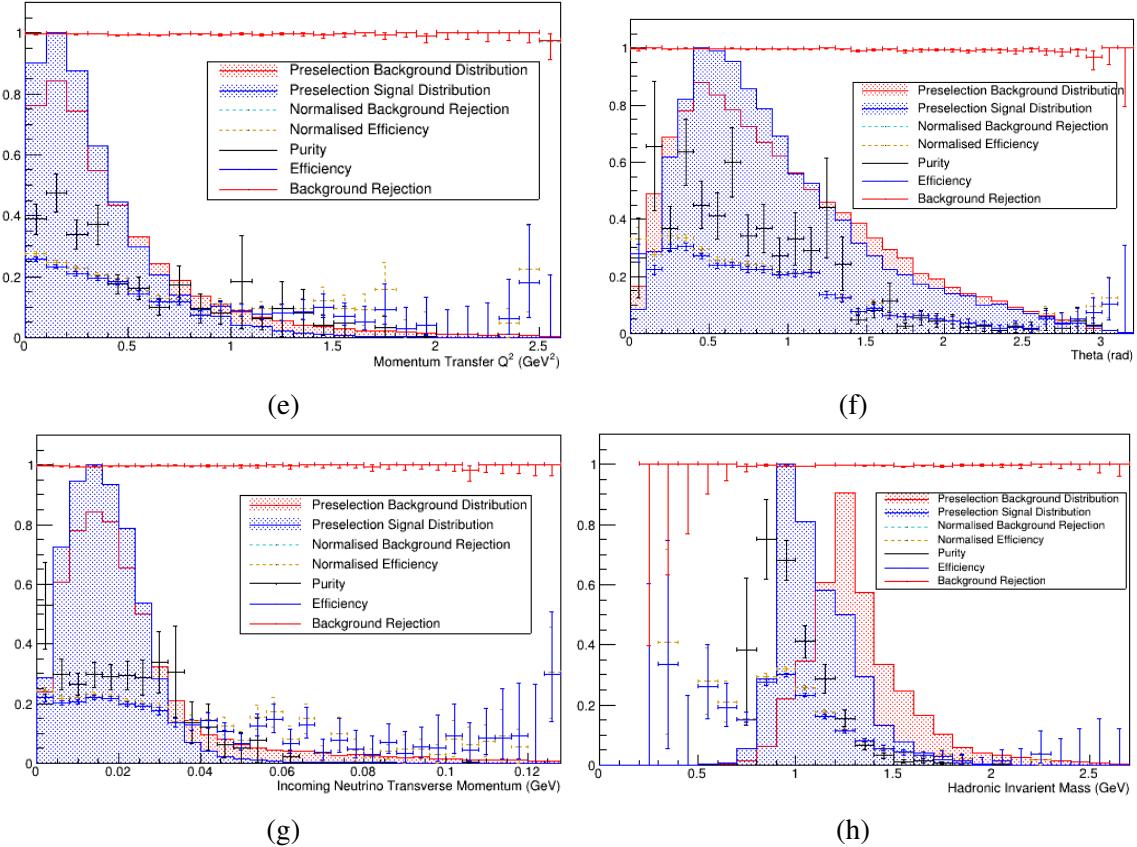
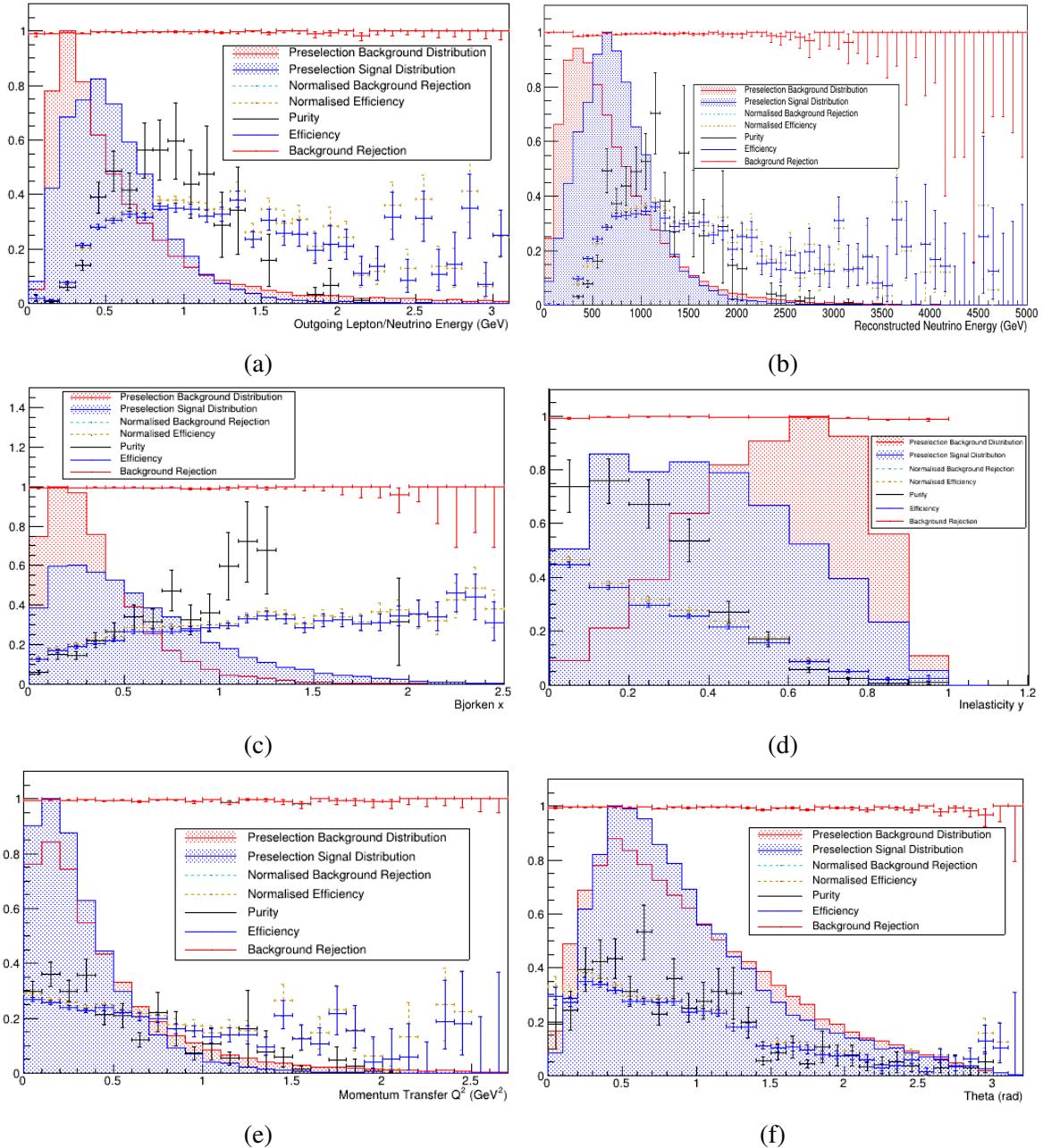


Fig. H.1 The efficiency of selecting CC  $\nu_e$  events (blue), the background rejection for removing  $\nu_\mu$  events (red) and the purity (black) of the proposal-like selection as a function of the neutrino interaction variables: a) the outgoing charged lepton/neutrino energy, b) the reconstructed neutrino energy, c) the Bjorken parameter  $x$ , d) the inelasticity  $y$ , e) the momentum transfer  $q^2$ , f) the angle between the outgoing charged lepton/neutrino and the incoming neutrino, g) the incoming transverse momentum, and h) the hadronic invariant mass. The underlying distributions before selection for the signal (red) and background (blue) are also plotted, along with the efficiency (dashed orange) and background rejection (dashed light blue) for the POT and oscillation normalised events. The efficiencies are calculated from the remaining events after the pre-selection cuts and the full selection. For the unnormalised events the errors are the  $1\sigma$  Clopper-Pearson intervals [32], whilst for the normalised distributions the normal approximation is used.

### New Cut-Based Selection

The efficiency curves for the new cut-based selection is shown in Figure H.2. The structure of the curves are explained using the arguments for the proposal style selection above. In this case the efficiency is marginally improved.



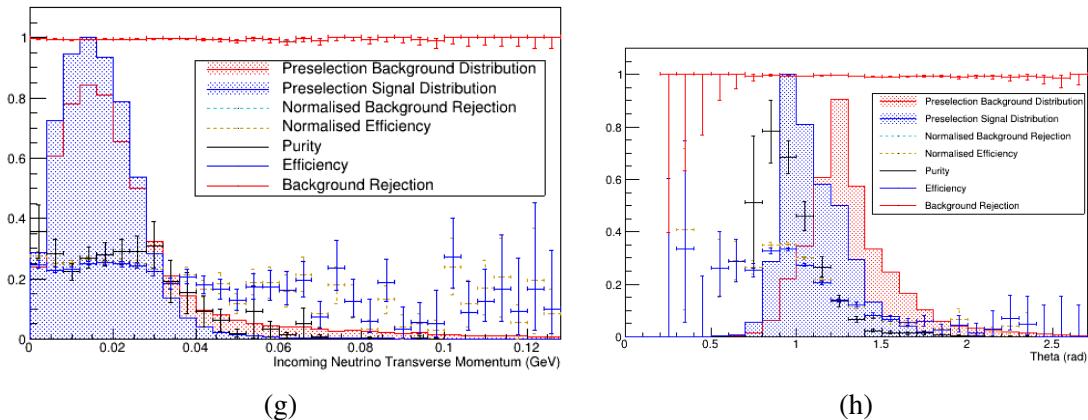
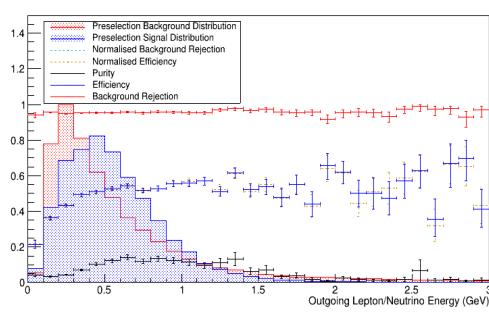


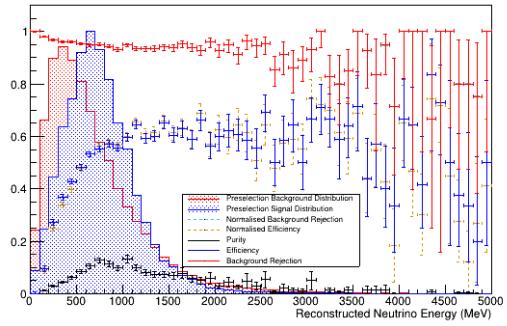
Fig. H.2 The selection efficiency of CC  $\nu_e$  events (blue), the background rejection for removing  $\nu_\mu$  events (red) and the purity (black) of the updated selection as a function of the neutrino interaction variables: a) the outgoing charged lepton/neutrino energy, b) the reconstructed neutrino energy, c) the Bjorken parameter  $x$ , d) the inelasticity  $y$ , e) the momentum transfer  $q^2$ , f) the angle between the outgoing charged lepton/neutrino and the incoming neutrino, g) the incoming transverse momentum, and h) the hadronic invariant mass. The underlying distributions before selection for the signal (red) and background (blue) are also plotted, along with the efficiency (dashed orange) and background rejection (dashed light blue) for the POT and oscillation normalised events. The efficiencies are calculated from the remaining events after the pre-selection cuts and the full selection. For the unnormalised events, the errors are the  $1\sigma$  Clopper-Pearson intervals [32] whilst for the normalised distributions the normal approximation is used.

## **BDT Selection**

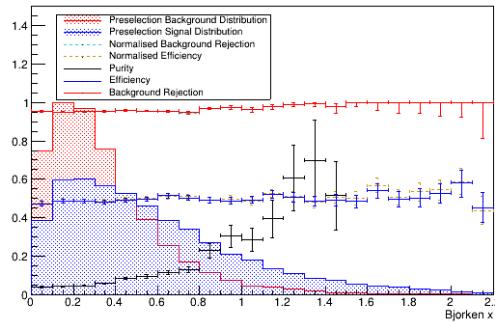
The efficiency curves for the BDT selection is shown in Figure H.3. The BDT reduces the dependencies of the selection on the neutrino parameters, notably in the higher energy range. There is only a notable drop in efficiency in the lower energy regions due to poor reconstruction and the 100 MeV cut-off. When performing a sensitivity analysis, the flattening of the efficiency as a function of the neutrino energy makes the BDT less susceptible to changes in the event rate due to underlying oscillation parameters. For example, the oscillation values could be such that the probability is significantly lower for neutrino energy around the beam peak ( $\sim 1$  GeV). The capabilities of the analysis would then rely on the oscillations for higher (or lower) neutrino energies. Thus it is useful to maintain a high efficiency across all neutrino energies.



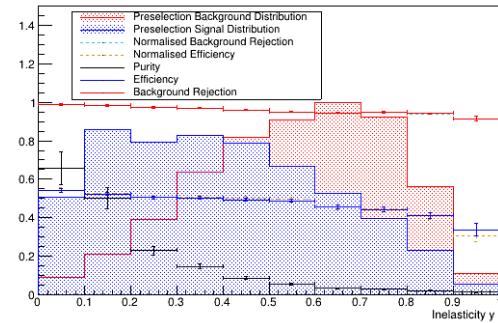
(a)



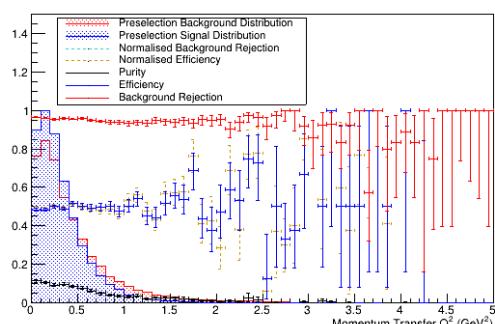
(b)



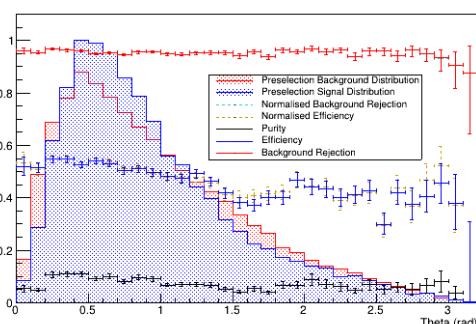
(c)



(d)



(e)



(f)

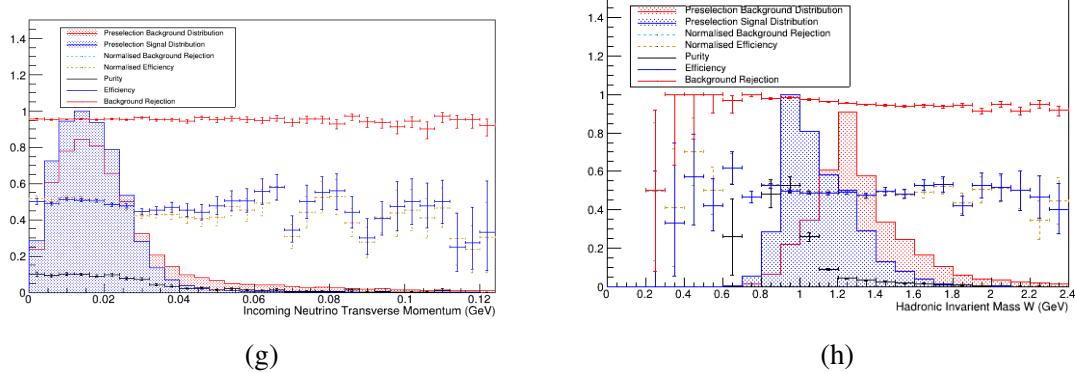


Fig. H.3 The efficiency of selecting CC  $\nu_e$  events (blue), the background rejection for removing  $\nu_\mu$  events (red) and the purity (black) of the standard BDT selection with a BDT cut off score at 0, a fiducial volume cut and a minimum shower energy cut of 100 MeV, as a function of the neutrino interaction variables: a) the outgoing charged lepton/neutrino energy, b) the reconstructed neutrino energy, c) the Bjorken parameter  $x$ , d) the inelasticity  $y$ , e) the momentum transfer  $q^2$ , f) the angle between the outgoing charged lepton/neutrino and the incoming neutrino, g) the incoming transverse momentum, and h) the hadronic invariant mass. The underlying distributions before selection for the signal (red) and background (blue) are also plotted, along with the efficiency (dashed orange) and background rejection (dashed light blue) for the POT and oscillation normalised events. For the unnormalised events, the errors are the  $1\sigma$  Clopper-Pearson intervals [32] whilst for the normalised distributions the normal approximation is used.



# Appendix I

## More Detail on the Additional Selection Metrics

This section discusses in more detail the metrics which were used to perform a new cut-based selection (described in Chapter 7, Section 7.4.4) and a Boosted Decision Tree (BDT) selection (described in Chapter 7, Section 7.4.5). The selections were then used in the oscillation analysis discussed in Chapter 8.

### Shower Residual Analysis Cut

This method calculates the number of showers within the reconstructed neutrino interaction. However, showers are ignored if the reconstructed energy was less than a specified threshold (10 MeV) and the showers are significantly far away from the primary shower. This distance is discussed more in Chapter 7 and is depicted in Figure 7.30.

Secondary showers were considered for analysis if the residual distance was greater than a fraction (known as the residual fraction  $f = 0.025$ ) of the base length of a cone, with a length corresponding to the position of the secondary shower, subtending the reconstructed opening angle of the shower. A 2 cm buffer was also added to allow for shower segments near the shower vertex to be removed from the analysis. Therefore, showers are considered for the analysis if the shower has an energy greater than the energy cut off and the residual cut,  $R$ , satisfies the following inequality:

$$R = |(\mathbf{S}_{sp} - \mathbf{S}_{sp} \cdot \mathbf{D}_p)\mathbf{D}_p| > f |\mathbf{S}_{sp} \cdot \mathbf{D}_p \tan(0.5\theta_p)| + 2, \quad (I.1)$$

where  $S_{sp} = S_s - S_p$ , for which  $S_p$  and  $S_s$  are the start position of the primary shower and the secondary shower respectively.  $D_p$  is the direction of the primary shower,  $\theta_p$  is the opening angle of the primary shower and  $f$  is the residual fraction threshold. The procedure allows

for segments to be further away from the core if they are further downstream of the shower start position.

A 2D grid-search for the energy cut as well as the optimal values of the residual fraction is shown in the Figures I.1. If there was at least one secondary shower using the parameters at the corresponding bin the event was removed. The maximum efficiency  $\times$  background rejection was found to be  $0.29 \pm 0.01$  at an energy cut off at 11.25 MeV and residual fraction of 0.019. This corresponds to an efficiency of  $78.8 \pm 0.7\%$  and a background rejection of  $37.7 \pm 0.5\%$ . The maximum efficiency  $\times$  purity was found to be  $0.008 \pm 0.000$  (efficiency of  $86.5 \pm 0.8\%$  and purity of  $0.9 \pm 0.0\%$ ) with an energy cut off of 23.75 MeV and a residual fraction of 2.109. A 10 MeV cut and a residual fraction of 0.025 was used, which results in a marginal difference from the optimal parameters described above. When events with no showers are removed, and allowing no additional showers from the residual analysis, the best cut off for efficiency  $\times$  background rejection results in an efficiency of  $77.4 \pm 1.0\%$  and a background rejection of  $38.9 \pm 0.3\%$ . Compared to the standard one shower cut in Section 7.4.3, an improvement of  $5.6 \pm 1.3\%$  in efficiency and a decrease of  $2.0 \pm 0.6\%$  in the background rejection is present. This results in an improvement of  $0.007 \pm 0.007$  in efficiency  $\times$  background rejection. Therefore, the residual analysis marginally improves the selection.

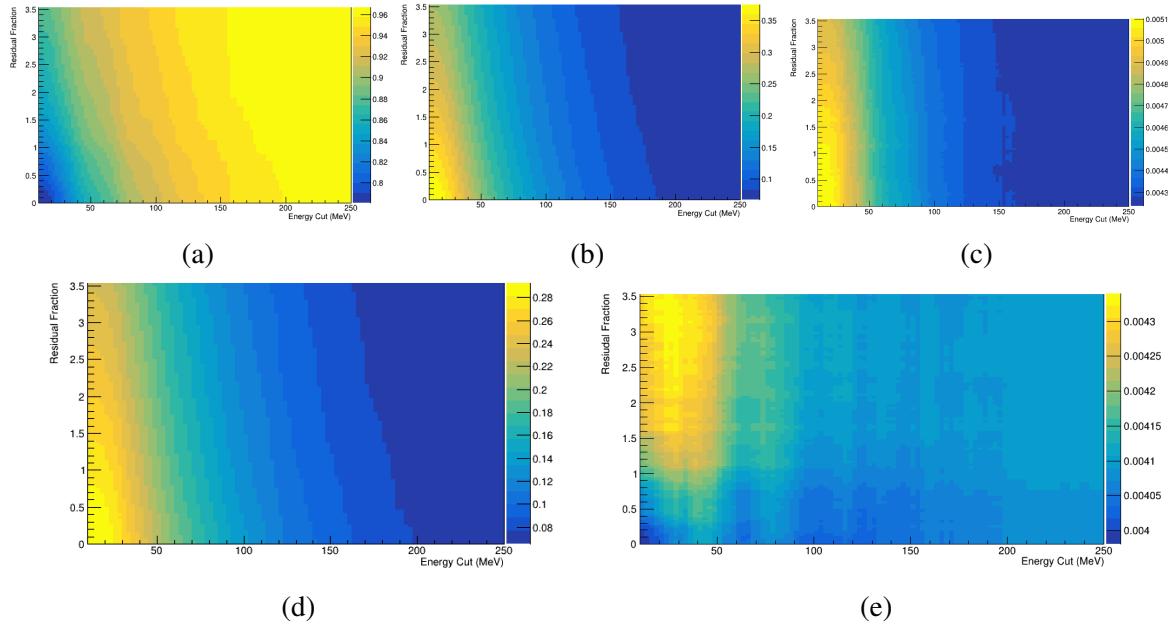


Fig. I.1 The 2D distributions of the efficiency (a), background rejection (b), purity (c), efficiency  $\times$  background rejection (d) and efficiency  $\times$  purity (e) for the normalised signal and background events when removing the events with a secondary shower, defined by residual analysis. The secondary shower is only considered for the analysis if the shower has a greater energy than the energy cut and the residual of the shower obeys the equality, Equation I.1, for the residual fraction in the plot.

### Maximum Track PIDA Cut

Track calorimetry information was also introduced to improve the selection using a PIDA analysis [271, 29], see Chapter 6, Section 6.1.8.

The PIDA score was combined with the maximum track length to enhance the separation. Figures I.2 show that a 2D analysis which removes events from the selection if the maximum track length is above 7.75 cm and the PIDA score is below 9.56 resulted in the best efficiency  $\times$  background rejection of  $0.49 \pm 0.01$  ( $78.3 \pm 1.0\%$  efficiency and a background rejection of  $62.3 \pm 0.7\%$ ). The parameters which maximise the efficiency  $\times$  purity are a maximum track length of 37.25 cm and a PIDA score of 9.94. This resulted in an efficiency  $\times$  purity of  $0.013 \pm 0.007$  with an efficiency of  $92.3 \pm 1.1\%$  and a purity of  $1.42 \pm 0.0\%$ . Combining the PIDA with the track length provides an improvement of  $0.03 \pm 0.01$  in efficiency  $\times$  background rejection ( $3.7 \pm 1.4\%$  loss in efficiency and  $5.7 \pm 1.0\%$  increase in background rejection) compared to the track length cut alone.

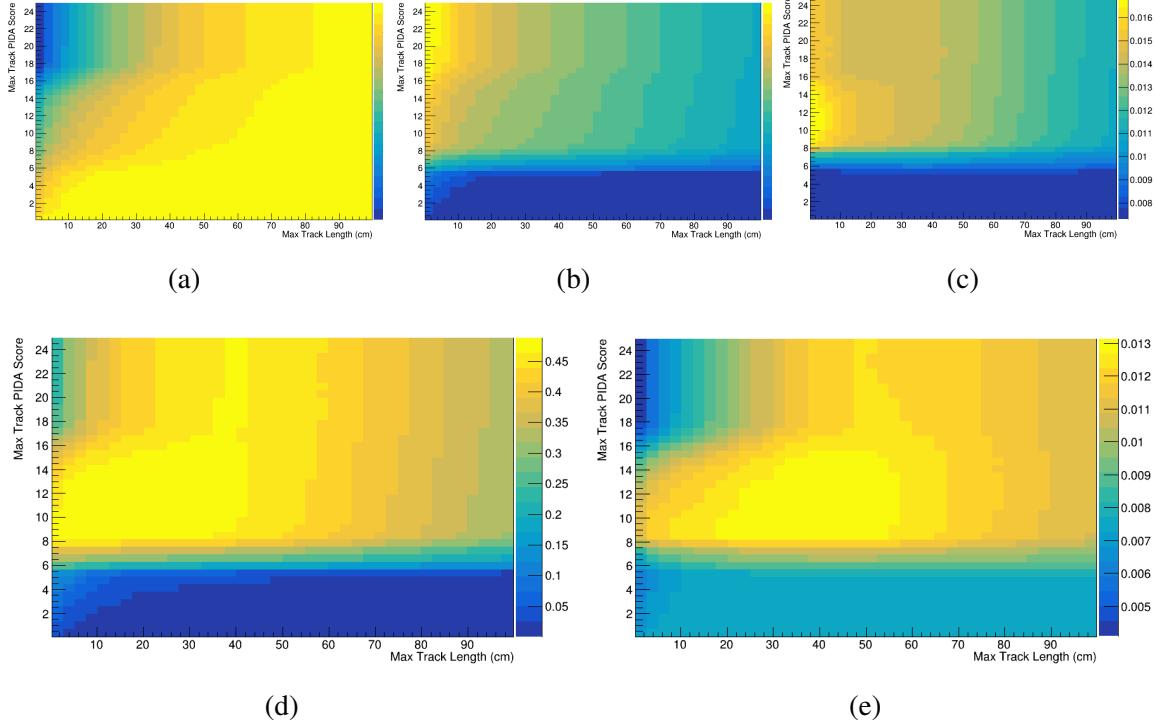


Fig. I.2 The 2D signal (a) and background (b) distributions of the length and the PIDA score of the longest track. Also shown are the 2D efficiency (b), background rejection (c), purity (d), efficiency  $\times$  background rejection, (e) and efficiency  $\times$  purity (f) for the normalised signal and background events when removing events where the length is greater than the length cut and the PIDA score is less than the PIDA cut off.

### Shower Length and Opening Angle Cuts

The length and the opening angle of a shower can be used to separate electron showers from photon showers. In the analysis, the largest shower is used for this.

The shower length can be parameterised in terms of the energy of the shower by

$$L = (\chi_0 \ln \frac{E}{E_c} + C_j), \quad (j = e/\gamma), \quad (\text{I.2})$$

where  $E_c$  is the critical energy of the material,  $\chi_0$  is the radiation length and  $C_e = -0.5/C_\gamma = 0.5$  [92, 321]. As the radiation length in liquid argon is 14 cm [171], the difference in length between photon and electron showers is, on average, 14 cm for the same shower energy. The separation between the  $\nu_e$  and  $\nu_\mu$  background is shown in Figure I.4. However, the dominating factor in this separation is due to the difference in the energy distributions of the signal and background, which is shown in Figure 7.28. The true separational power

is more prominent in the vertex sample, described in Chapter 6, where the photons and electrons are simulated with the same energy. Figure I.3a shows the vertex showers with a reconstructed energy greater than 200 MeV and still shows the separation. Figure I.3b shows the shower length as a function of the true shower energy and the logarithmic dependence is also visible. The Figure shows the separational power is independent of energy. It can be seen that the separational power is small with respect to the RMS on the distribution and thus the shower length is not a powerful separational metric for electron/photon showers of the same energy.

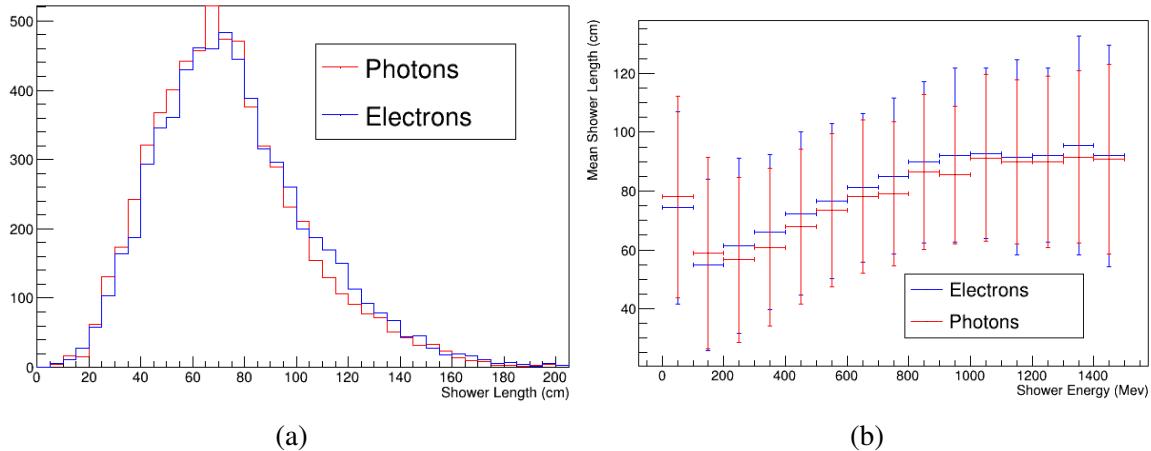


Fig. I.3 The largest shower length for electrons (blue) and photons (red) for all events where the reconstructed shower energy is greater than 200 MeV for the vertex sample. This is shown as a total distribution (a) and as a function of the true energy (b). The errors on the distribution are the standard deviation of the events between energies  $x - \delta x \rightarrow x + \delta x$ .

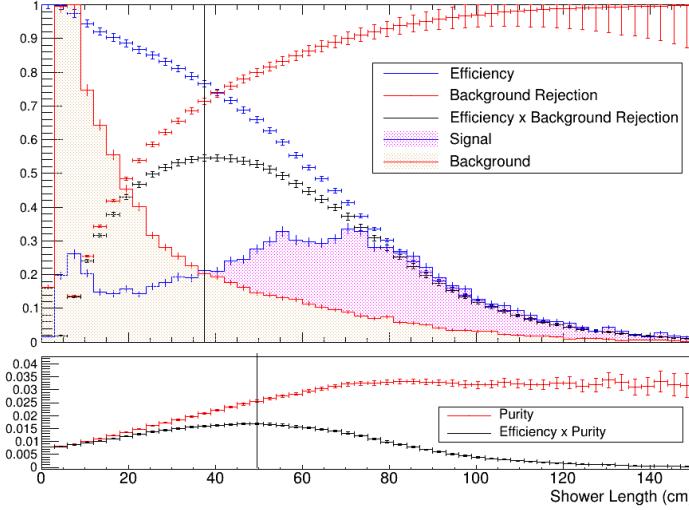


Fig. I.4 Largest shower length for signal (blue hatched) and background (red hatched) events. The efficiency (blue), background rejection (red), efficiency  $\times$  background rejection (black), purity (red) and efficiency  $\times$  purity (black) curves are also plotted. The errors on the curves are the cumulative statistical error.

The width of the shower is described by the Molière radius

$$R_M = \chi_0 \frac{E_s}{E_c}, \quad (\text{I.3})$$

where  $E_s \approx 21$  MeV [92, 322]. This is known as the multiple-scattering energy. 99% of the shower is contained within  $3.5 R_M$  ( $R_M = 10$  cm in LAr) [92]. The reconstructed opening angle is defined as  $\tan^{-1}(0.5 \text{ Width}/\text{Length})$ . As the width is related to the Molière radius, the opening angle is a function of the shower energy and the critical energy. Figure I.6 shows the separation between the signal and background for the BNB sample. Figures I.5 show the separation between electrons and photons for the vertex sample where the reconstructed events have an energy greater than 200 MeV for the opening angle. Just as the shower length case, the vertex sample Figures show the separation is primarily due to the energy of the shower.

The best cuts to maximise the efficiency  $\times$  background rejection or efficiency  $\times$  purity for the shower length and opening angle are shown in Table 7.6.

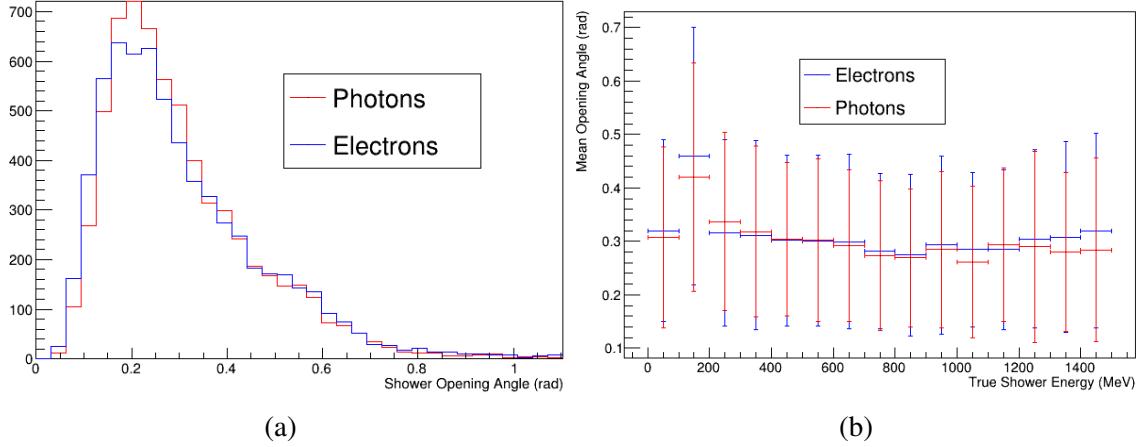


Fig. I.5 The largest shower opening angle for electrons (blue) and photons (red) for all events where the reconstructed shower energy is greater than 200 MeV for the vertex sample. This is shown as a total distribution (a) and as a function of the true energy (b). The errors on the distribution are the standard deviation of the events between energies  $x - \delta x \rightarrow x + \delta x$ .

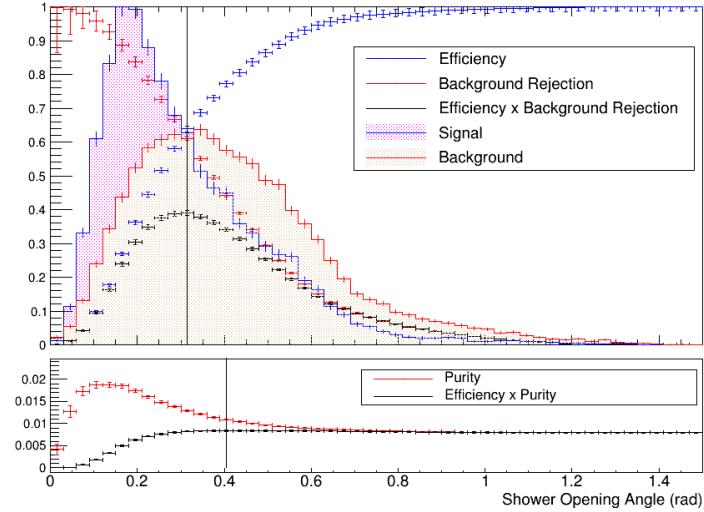


Fig. I.6 The largest shower opening angle for signal (blue hatched) and background (red hatched) events. The efficiency (blue), background rejection (red), efficiency  $\times$  background rejection (black), purity (red) and efficiency  $\times$  purity (black) curves are also plotted. The errors on the curves are the cumulative statistical error.

### Shower Relative Energy Density Cut

To further separate electrons and photons, the relative energy density profile as a function of the distance was undertaken. When viewing a photon with energy  $E$ , which has undergone pair production, as two electron-like showers overlaid with an energy  $\sim E/2$ , it is a reasonable

assumption that the length of the shower is shorter than a single electron shower with energy  $E$ . Furthermore, the density at the start of the shower would be greater. Therefore, when compared to a single shower, the density would be lower further away from the vertex and higher closer to the vertex.

In order to identify these differences, showers were split into ten segments from the start of the shower to the longitudinal length of the shower as defined by the reconstruction. Spacepoints not contained within a cone are removed. The cone is defined by the reconstructed shower opening angle and the projected length of the spacepoint from the shower start position. Such a cone is depicted in Figure 7.30. Segments with less than ten spacepoints are not considered in the analysis.

The relative energy density of the segment is then calculated by taking the reconstructed energy deposited within the segment and dividing the volume of the segment subtended by the cone. A  $\chi^2$  minimisation fit of the distance from the centre of the segment against the relative energy density to the function  $ax^{-b}$  is then performed, where  $a$  and  $b$  are fit parameters. Fits to the average values for the electron and photon vertex distributions are shown in Figure I.7, showing discriminatory power. When considering the events independently, limits on the parameter  $a$  are set to  $0 \rightarrow 1$  and on  $b$  are set to  $1 \rightarrow 2$ . These parameters are set so that the fit is constrained around the average results obtained from the vertex sample. The relative energy density gradient  $a$  and the power  $b$  for the  $\nu_e$  CC and  $\nu_\mu$  background are shown in Figures I.8a and I.8b respectively. The best cuts to maximise the efficiency  $\times$  background rejection or efficiency  $\times$  purity are shown in Table 7.6.

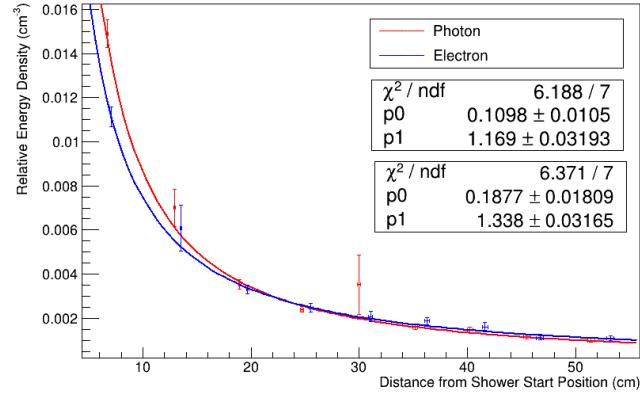


Fig. I.7 The average relative energy density of the shower segments as a function of the distance from the start position of the largest shower in the vertex sample events. The points represent the average value for one of the ten segments analysed. The errors correspond to the error on the mean for the values within the segment. The blue points are from the electron sample and the blue line is the best fit for the electrons. The red points are from the photon sample and the red line is the best fit. A  $\chi^2$  minimisation is performed to do the fit.

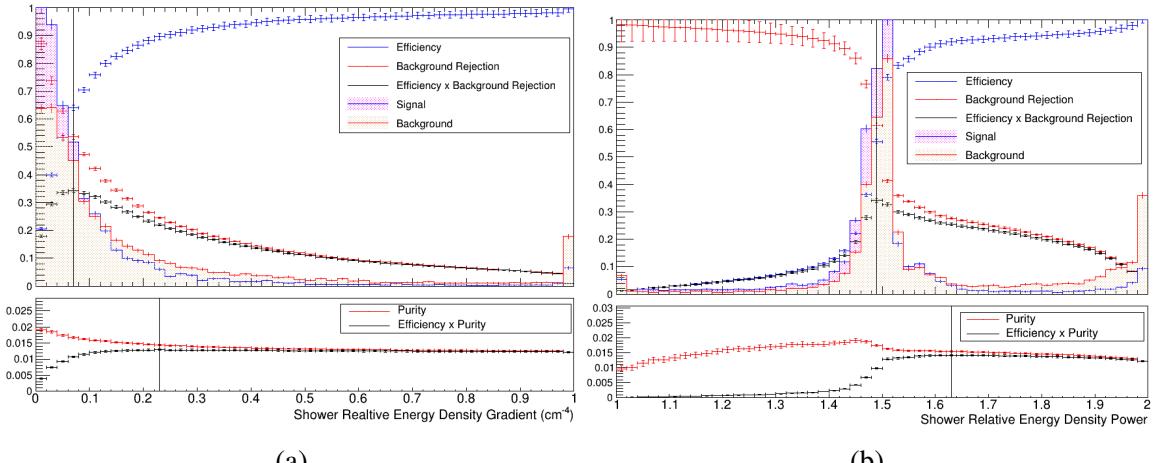


Fig. I.8 The relative fit energy density gradient fit parameter (a) and the power fit parameter (b) for the largest shower for the signal (blue hatched) and background (red hatched) events. The efficiency (blue), background rejection (red), efficiency  $\times$  background rejection (black), purity (red) and efficiency  $\times$  purity (black) curves are also plotted. The errors of the curves are the cumulative statistical error.

### Shower Track Stub Width and Length Cuts

As was seen in Chapter 6, Section 6.2.2, the true length and width of the track stub of the shower separates the photons and electrons. This separation is due to the background photon  $e^-e^+$  pairs having a higher probability of interacting significantly compared to a signal  $e^-$ .

There is also the possibility that one of the pair stops. Therefore, the electron track length is longer than a photon track stub.

The track stub length is defined as the distance from the first trajectory point to the last trajectory point of the reconstructed track, see Chapter 6, Section 6.1.6 for further detail. The track stub width is defined as the average perpendicular distance of the spacepoints to the direction of the shower. Figures I.9a,I.9b, I.10a and I.10b show the separation between the photons and electrons for the vertex sample with a reconstructed energy greater than 200 MeV. As can be seen in the Figures, for the vertex case where the reconstruction performance increases, there is separational power using the track length and width. The distributions for the BNB signal and background samples after the pre-selection cuts are shown in Figures I.11a and I.11b for the length and width respectively. Like the other shower metrics discussed in this Section, the separational power is low. Therefore the metrics have been implemented in a BDT such that the small separational power can be extracted without a significant loss in the selection efficiency. This process is described in the next Section. The best cuts to maximise the efficiency  $\times$  background rejection or efficiency  $\times$  purity is shown in Table 7.6.

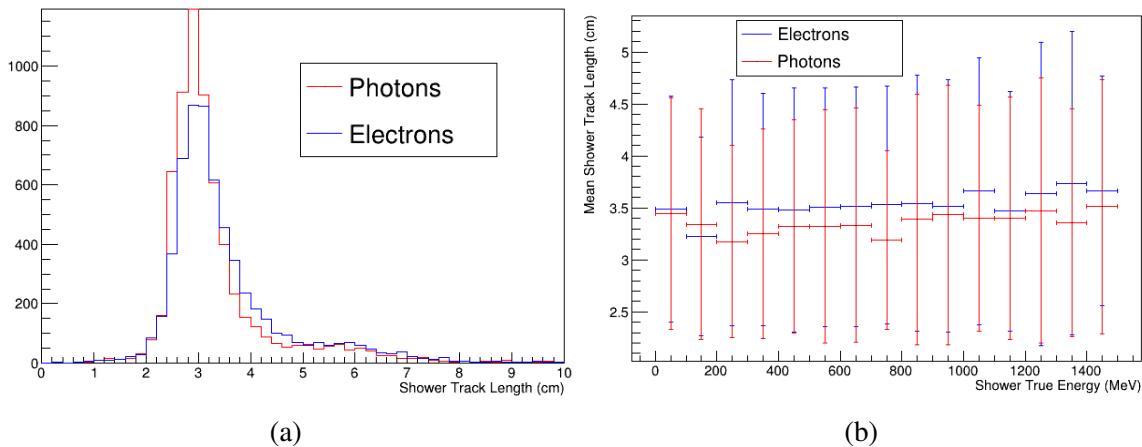


Fig. I.9 The track stub length of the largest shower for electrons (blue) and photons (red) for the vertex sample events where the reconstructed shower energy is greater than 200 MeV. The total distribution is shown in (a) and as a function of the true energy in (b). The errors on the distribution are the standard deviation of the events between energies  $x - \delta x \rightarrow x + \delta x$ .

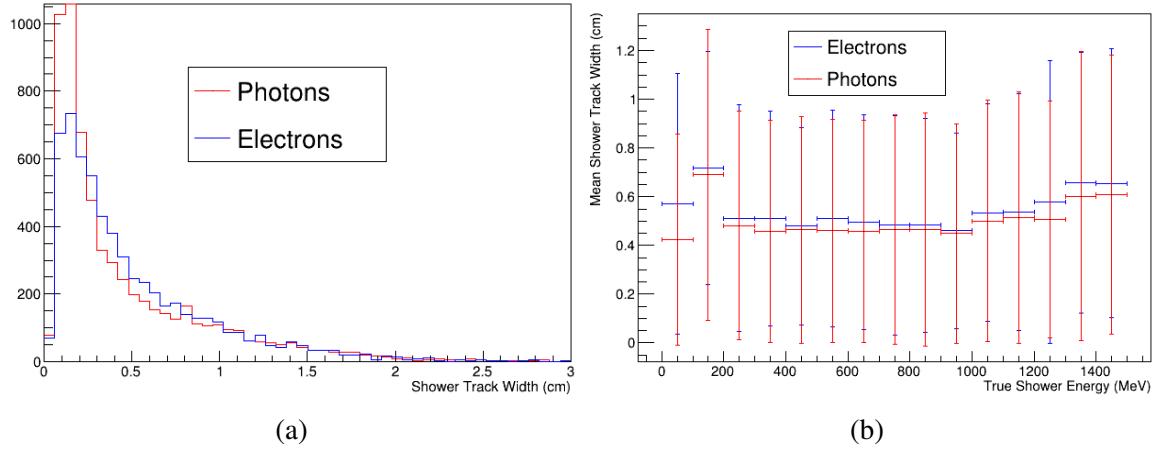


Fig. I.10 The track stub width of the largest shower for electrons (blue) and photons (red) for events where the reconstructed shower energy is greater than 200 MeV. a) Shows the total distribution and b) shows the events as a function of the true energy for the vertex sample. The errors on the distribution are the standard deviation of the events between energies  $x - \delta x \rightarrow x + \delta x$ .

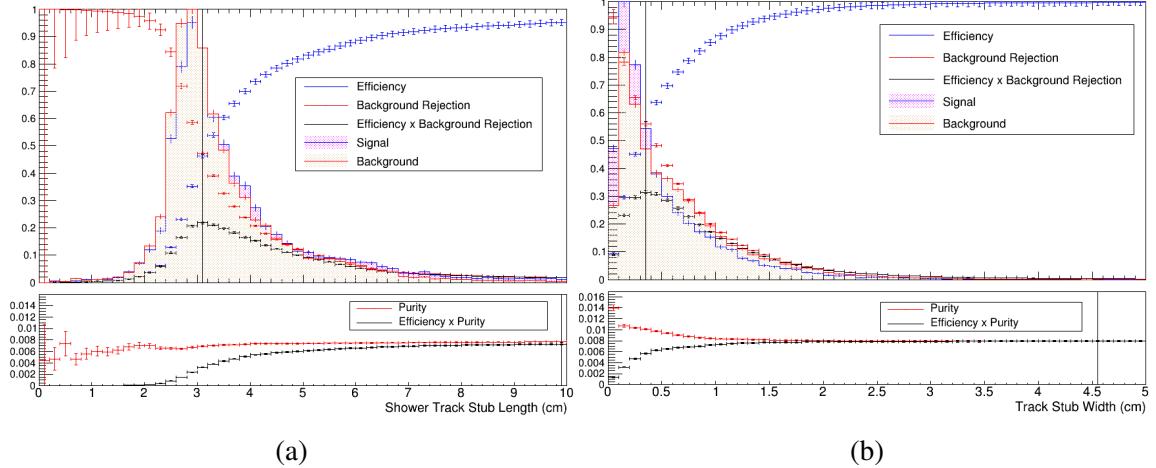


Fig. I.11 The track stub length (a) and width (b) for largest shower in the events for BNB oscillated CC  $\nu_e$  signal (blue hatched) and BNB  $\nu_\mu$  background (red hatched). The efficiency (blue), background rejection (red), efficiency  $\times$  background rejection (black), purity (red) and efficiency  $\times$  purity (black) curves are also plotted. The errors on the curves are the cumulative statistical error.

### Number of Reconstructed Neutrinos Cut

It is also possible to separate events by the number of reconstructed neutrinos. For example, if there is a significant distance between photons from decaying  $\pi^0$  events and the Pandora slicing fails, the event could be separated into two neutrinos with starting positions

corresponding to the start of the shower. Figure I.12 shows the number of reconstructed neutrinos for signal events and background events. Although there is a difference, multiple neutrino interactions, known as pile up, are possible, particularly in SBND and thus the cut should be used with caution. The probability of at least two neutrinos in the detector per spill is relatively low (calculated by D.Brailsford) at approximately 0.17% compared to one neutrino at 6.77 %. This results in a probability of having two neutrinos in the detector of 2.52% when there is neutrino activity in the TPC. This value is significantly less than the reconstructed value of  $\sim 9.5\%$  of background events with two or more reconstructed neutrinos. The efficiencies of the cut are shown in Table 7.6.

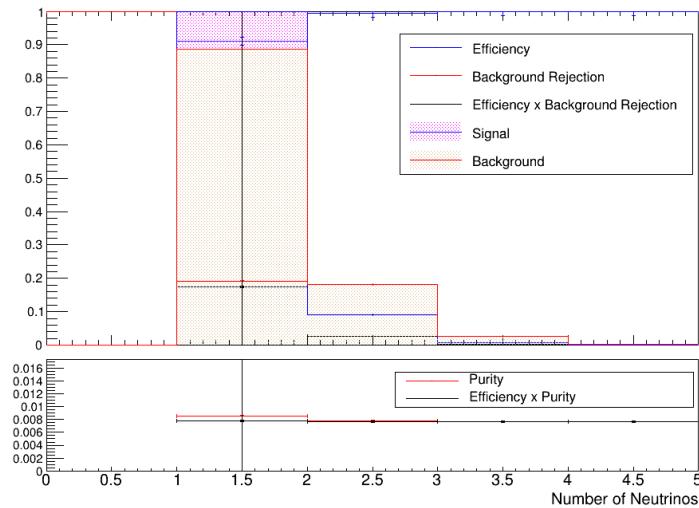


Fig. I.12 The number of neutrinos in the event for BNB oscillated CC  $\nu_e$  signal (blue hatched) and BNB  $\nu_\mu$  background (red hatched) events. The efficiency (blue), background rejection (red), efficiency  $\times$  background rejection (black), purity (red) and efficiency  $\times$  purity (black) curves are also plotted. The errors on the curves are the cumulative statistical error.

# Appendix J

## $\nu_e$ Selection Spectra

The following are resultant spectra when performing the various selections in Chapter 7 using the samples discussed in Chapter 7, Section 7.1.1. The spectra are discussed in Chapter 8 Section 8.2

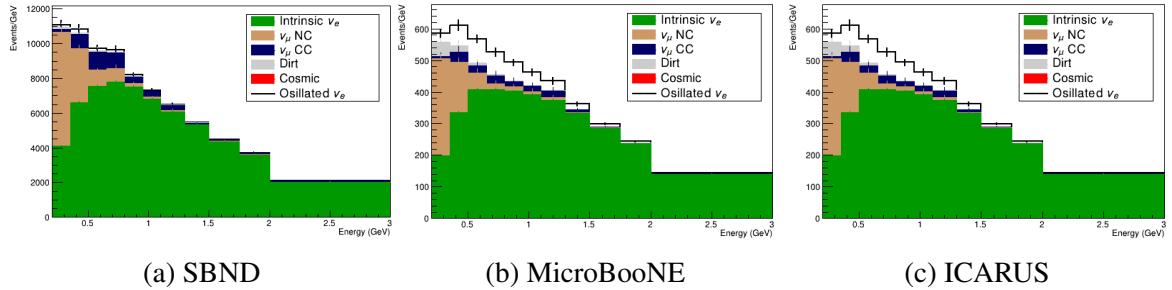


Fig. J.1 The spectra from the three detectors for the truth based selection. The selection is performed on the proposal-like sample described in Chapter 7, Section 7.1.2.

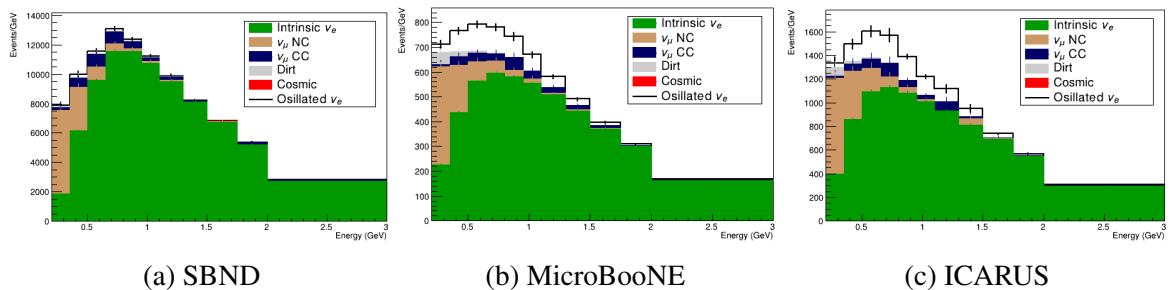


Fig. J.2 The spectra from the three detectors for the truth-based selection. The selection is performed on the modern-like sample described in Chapter 7, Section 7.1.1. The errors are statistical.

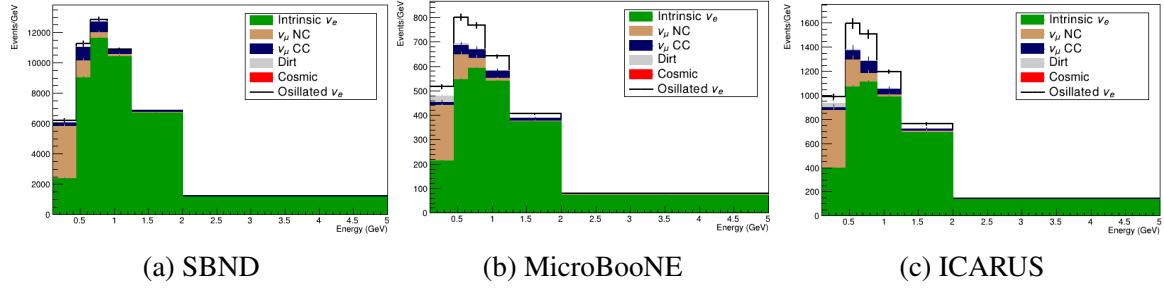


Fig. J.3 The spectra from the three detectors for the truth-based selection with the new binning. The selection is performed on the modern-like sample described in Chapter 7, Section 7.1.1. The errors are statistical.

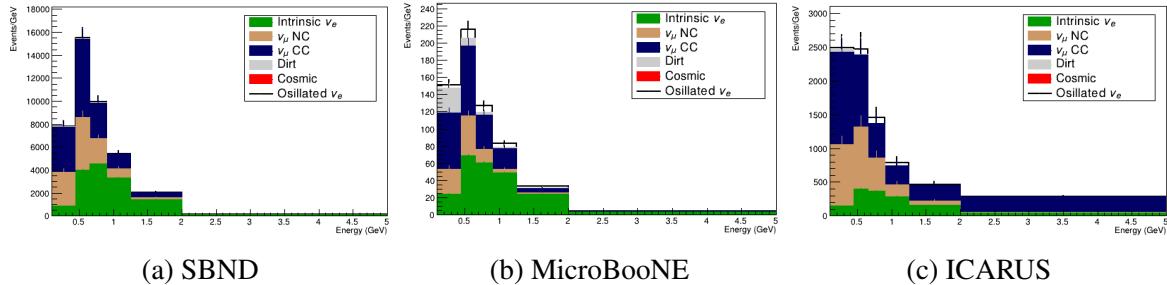


Fig. J.4 The spectra from the three detectors for the proposal-like selection with the parameters discussed in Chapter 7, Section 7.4.3 set to maximise the efficiency  $\times$  background rejection. The spectra are created using the reconstruction information of the Time Projection Chamber (TPC) signal and backgrounds and the truth-based dirt and cosmic background. The errors are statistical.

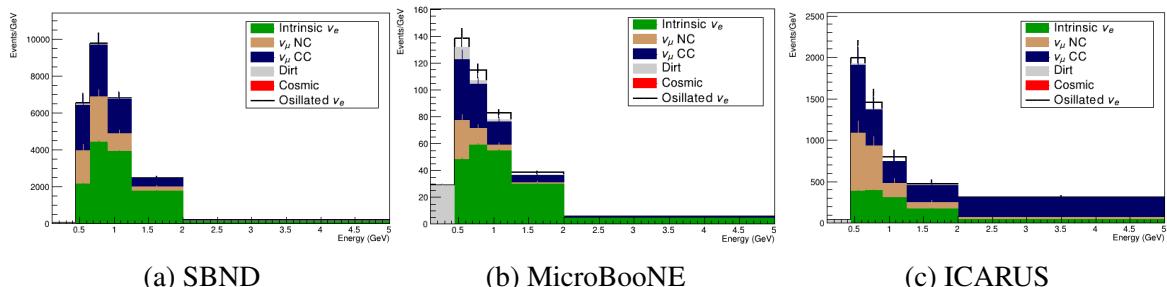


Fig. J.5 The spectra from the three detectors for the proposal-like selection with the parameters discussed in Chapter 7, Section 7.4.3 set to maximise the efficiency  $\times$  purity. The spectra are created using the reconstruction information of the TPC signal and backgrounds and the truth-based dirt and cosmic background. The errors are statistical.

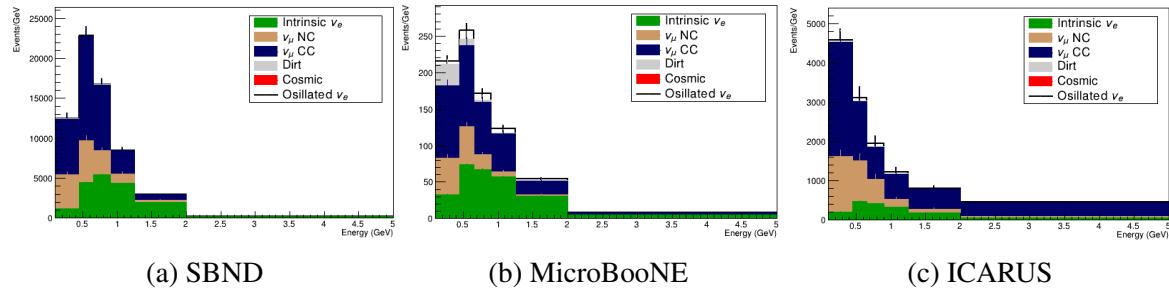


Fig. J.6 The spectra from the three detectors for the proposal-like selection with the parameters discussed in Chapter 7, Section 7.4.3 set to the proposal cut values. The spectra are created using the reconstruction information of the TPC signal and backgrounds and the truth-based dirt and cosmic background. The errors are statistical.

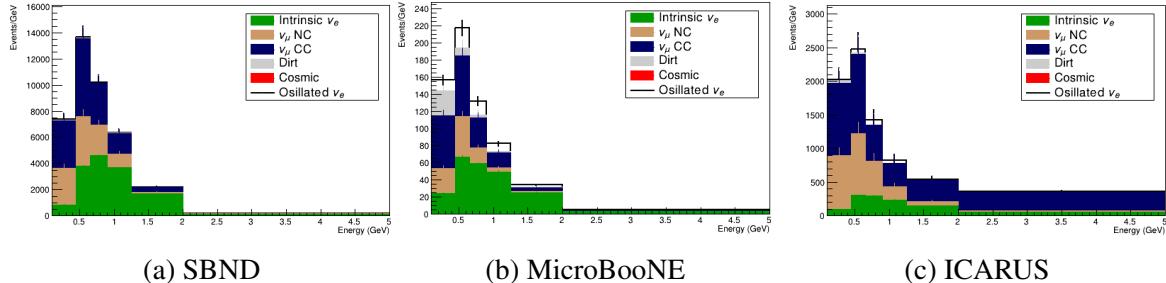


Fig. J.7 The spectra from the three detectors for the new selection with the parameters discussed in Chapter 7, Section 7.4.4 set to maximise the efficiency  $\times$  background rejection. The spectra are created using the reconstruction information of the TPC signal and backgrounds and the truth-based dirt and cosmic background. The errors are statistical.

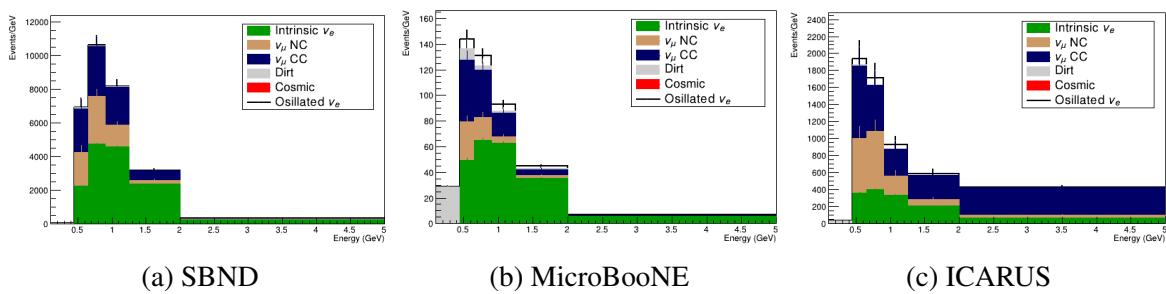


Fig. J.8 The spectra from the three detectors for the new selection with the parameters discussed in Chapter 7, Section 7.4.4 set to maximise the efficiency  $\times$  purity. The spectra are created using the reconstruction information of the TPC signal and backgrounds and the truth-based dirt and cosmic background. The errors are statistical.

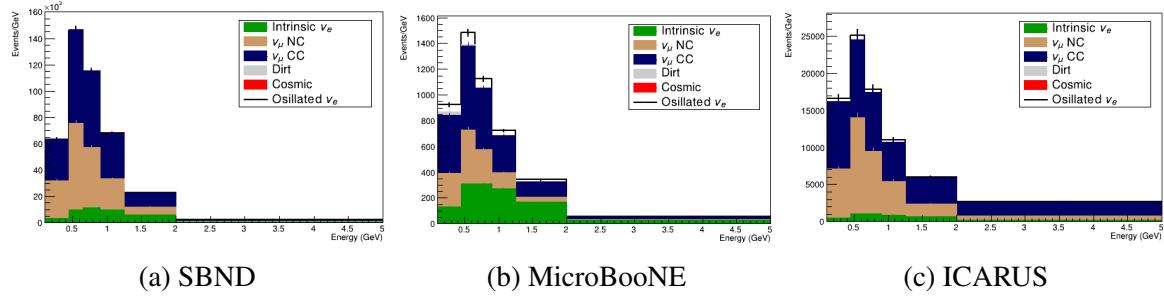


Fig. J.9 The spectra from the three detectors for the Boosted Decision Tree (BDT) selection with the parameters discussed in Chapter 7, Section 7.4.5 for a BDT score of 0.2. The BDT where the signal and background have been normalised to 1:1 is used. The spectra are created using the reconstruction information of the TPC signal and backgrounds and the truth-based dirt and cosmic background. The errors are statistical.

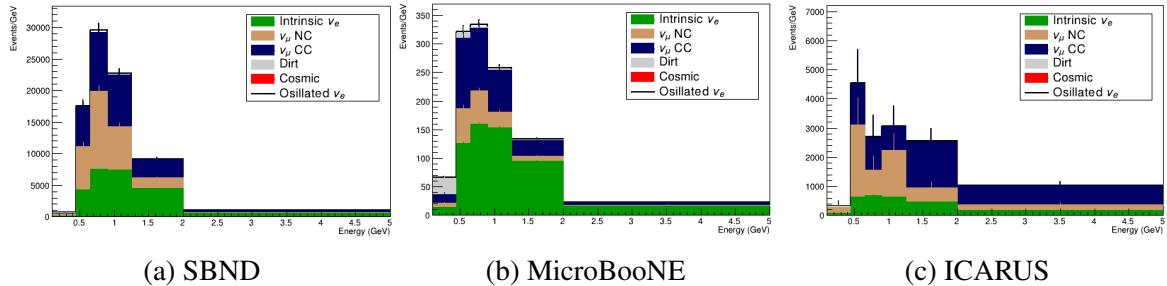


Fig. J.10 The spectra from the three detectors for the BDT selection with the parameters discussed in Chapter 7, Section 7.4.5 for a BDT score of -0.99. This is the BDT where the signal and background have been normalised with the Protons On Target (POT) and oscillation parameters described in Chapter 7. The spectra are created using the reconstruction information of the TPC signal and backgrounds and the truth-based dirt and cosmic background. The errors are statistical.

## Appendix K

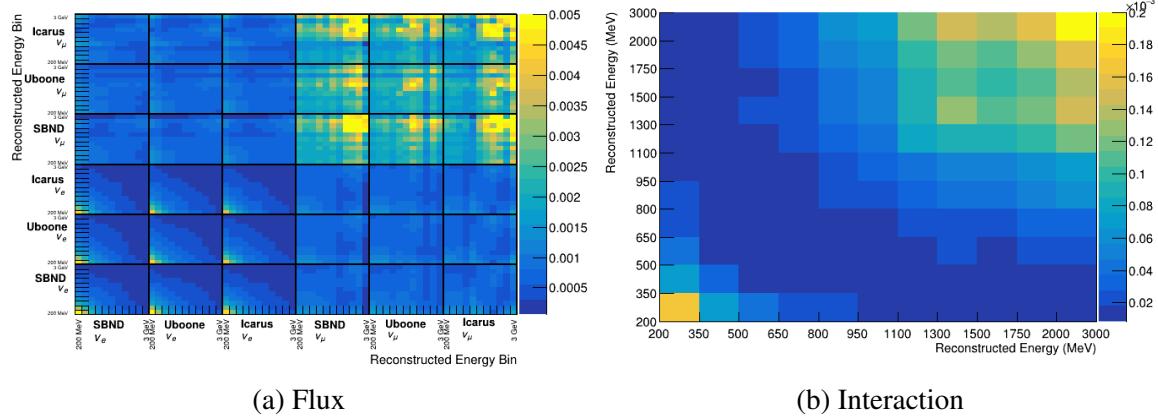
# Systematical Covarience Matrices for the $\nu_e$ Analysis

The following are the covariance matrices that are created as a result of applying the systematic treatment discussed in Chapter 8, Section 8.3 to the selected spectra presented in Appendix J.

As there are no MEC events within the proposal selection, the uncertainty is reduced significantly at the  $\sim 1$  GeV energy bin. The higher energy correlation in the interaction matrices is due to the uncertainty on the resonance events.

Differences occur due to the energy cut-off of the selections that maximise efficiency  $\times$  purity. The cut-off results in having only dirt events in the first energy bins for the interaction matrices and no events in the flux matrices.

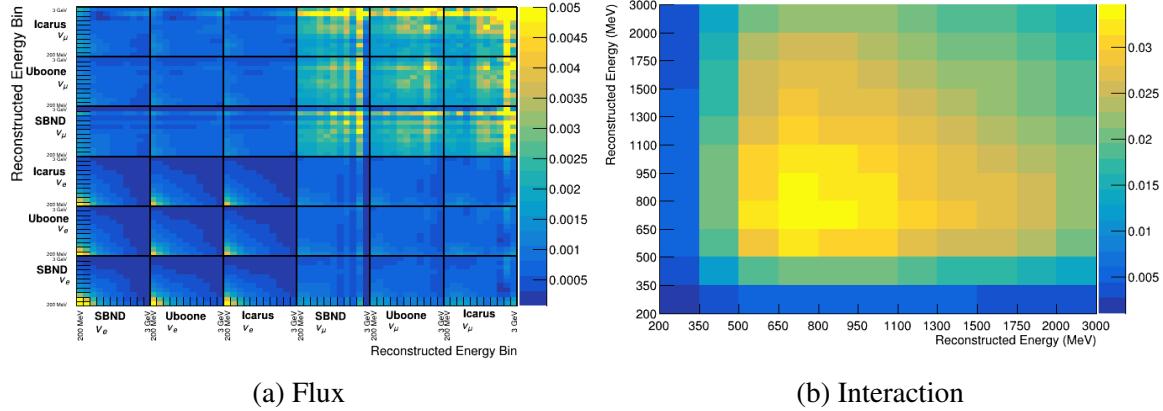
There are also noticeable correlations in the lower energy bins of the interaction matrices and the higher energy bins for the BDT selections that are not visible in truth-based selections. This could be due to a lower energy cut-off; however, further investigation into this is required. Also, there is a lower overall correlation between the bins of the BDT selections than the traditional cut-based methods.



(a) Flux

(b) Interaction

Fig. K.1 The fractional covariance matrices for the flux (a) and interaction (b) for the truth-based proposal selection, described in Chapter 7, Section 7.1.2. The matrices are created from spectra, Figure J.1.



(a) Flux

(b) Interaction

Fig. K.2 The fractional covariance matrices for the flux (a) and interaction (b) for the truth-based modern selection discussed in Chapter 7, Section 7.1.1. The matrices are created from spectra, Figure J.2.

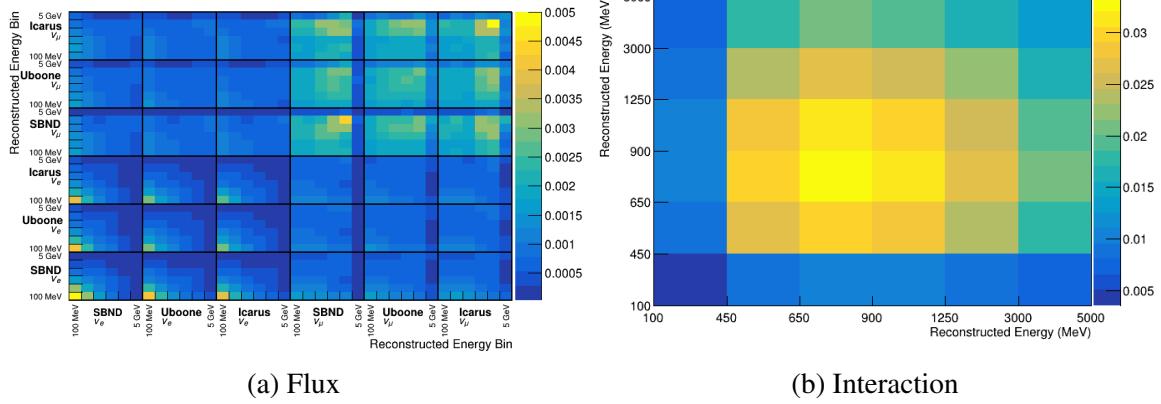


Fig. K.3 The fractional covariance matrices for the flux (a) and interaction (b) for the truth-based modern selection discussed in Chapter 7, Section 7.1.1. The matrices are created from spectra, Figure J.3. The new binning scheme has been used.

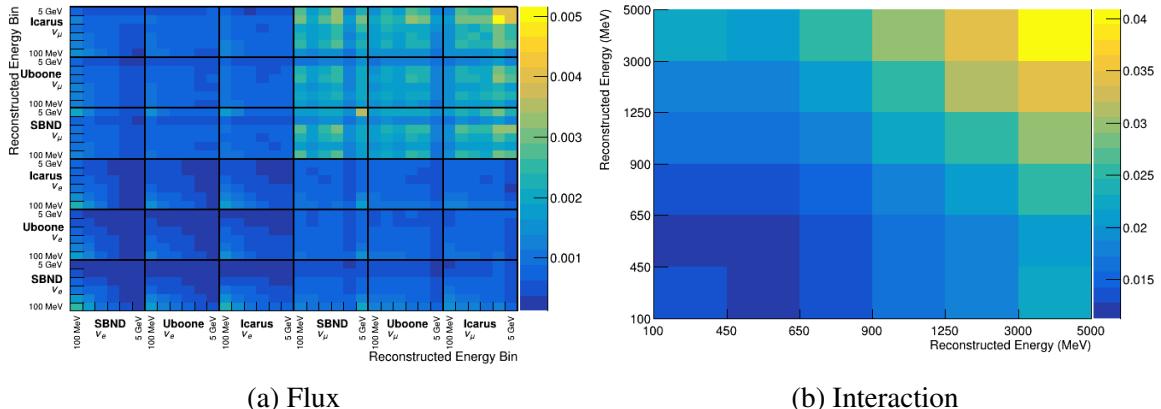


Fig. K.4 The fractional covariance matrices for the flux (a) and interaction (b) for the reconstruction selection discussed in Chapter 7, Section 7.4.3. The matrices are created from spectra, Figure J.4, which uses the proposal selection, maximising the efficiency  $\times$  background rejection. The new binning scheme has been used.

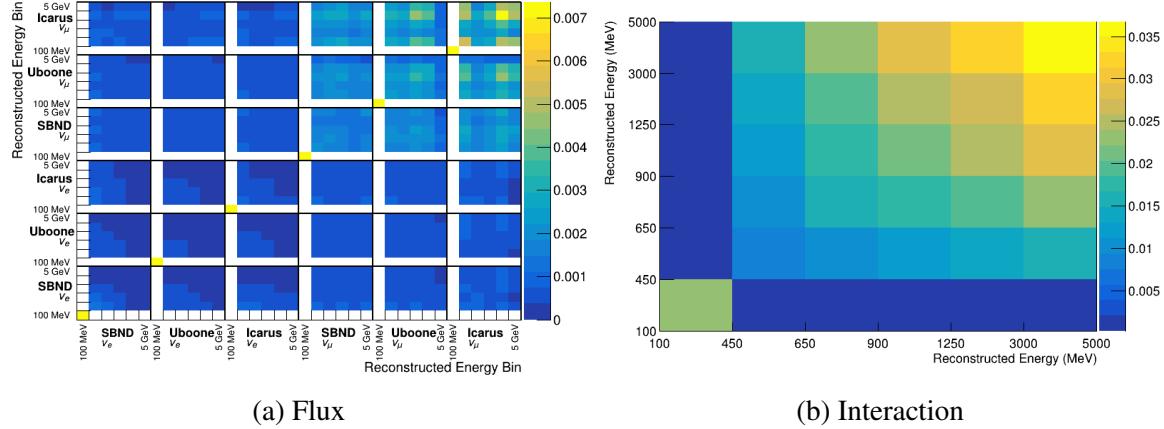


Fig. K.5 The fractional covariance matrices for the flux (a) and interaction (b) for the reconstruction selection discussed in Chapter 7, Section 7.4.3. The matrices are created from spectra, Figure J.5, which uses the proposal selection, maximising the efficiency  $\times$  purity. The new binning scheme has been used.

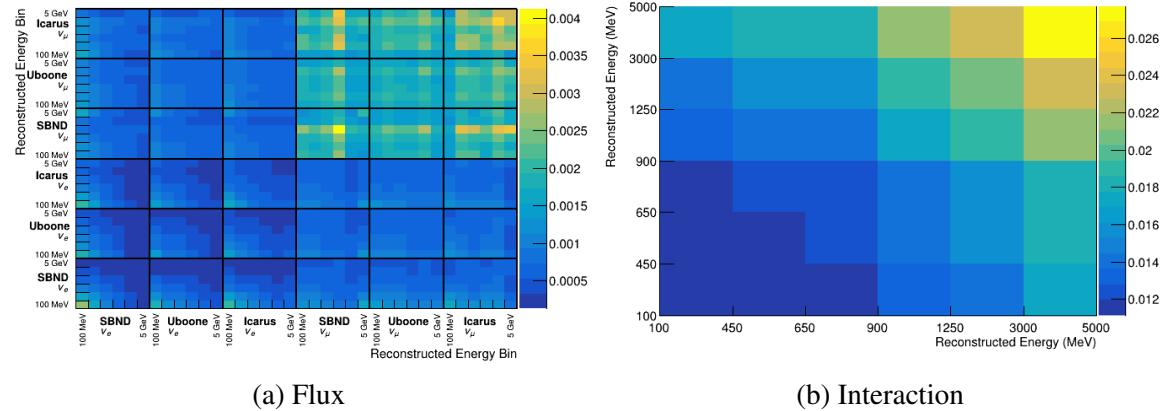


Fig. K.6 The fractional covariance matrices for the flux (a) and interaction (b) for the reconstruction selection discussed in Chapter 7, Section 7.4.3. The matrices are created from spectra, Figure J.6, which uses the proposal selection using the proposal cut values. The new binning scheme has been used.

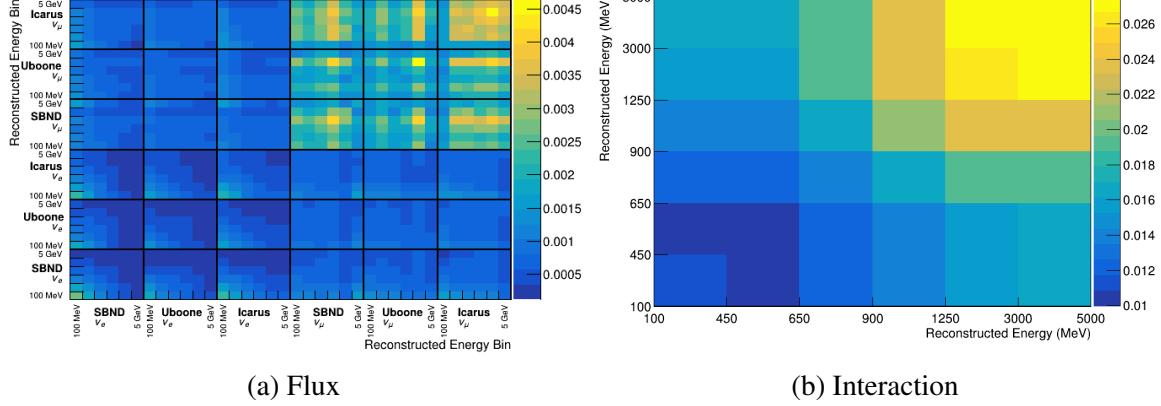


Fig. K.7 The fractional covariance matrices for the flux (a) and interaction (b) for the reconstruction selection discussed in Chapter 7, Section 7.4.4. The matrices are created from spectra, Figure J.7, which uses the new selection, maximising the efficiency  $\times$  background rejection. The new binning scheme has been used.

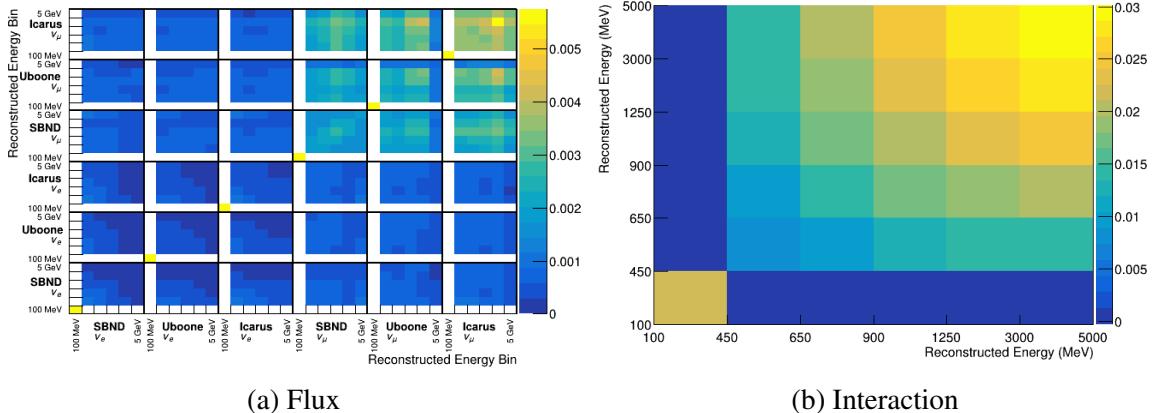


Fig. K.8 The fractional covariance matrices for the flux (a) and interaction (b) for the reconstruction selection discussed in Chapter 7, Section 7.4.4. The matrices are created from spectra, Figure J.8, which uses the new selection, maximising the efficiency  $\times$  purity. The new binning scheme has been used.

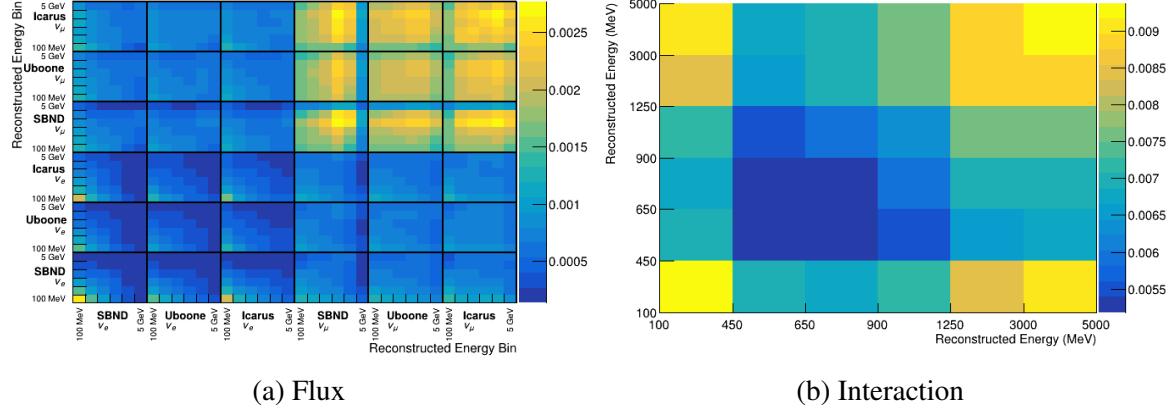


Fig. K.9 The fractional covariance matrices for the flux (a) and interaction (b) for the reconstruction selection discussed in Chapter 7, Section 7.4.5. The matrices are created from spectra, Figure J.9, which uses the BDT selection. The BDT where the signal and background have been normalised to 1:1 is used with a BDT score threshold of 0.2. The new binning scheme has been used.

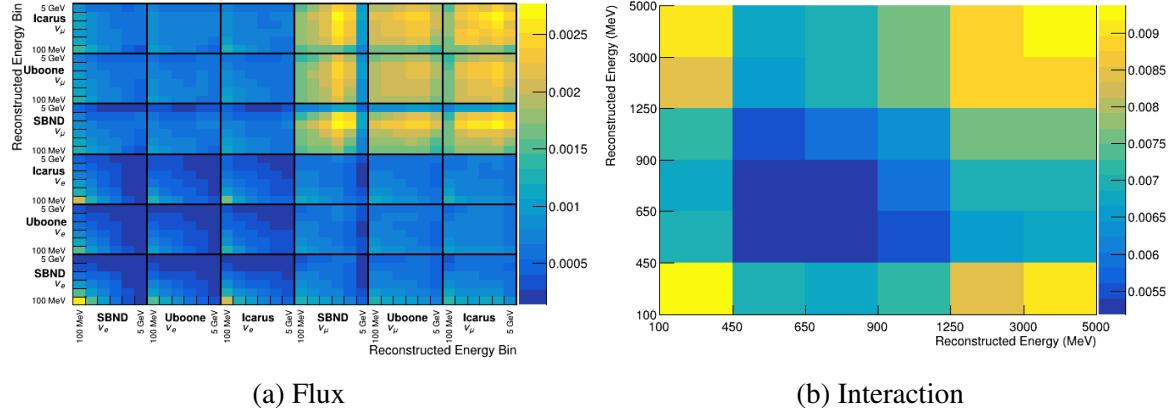


Fig. K.10 The fractional covariance matrices for the flux (a) and interaction (b) for the reconstruction selection discussed in Chapter 7, Section 7.4.5. The matrices are created from spectra, Figure J.10 which uses the BDT selection. The BDT where the signal and background have been normalised with the POT and oscillation parameters described in Chapter 7 is used. A BDT score threshold of -0.99 was used. The new binning scheme has been used.