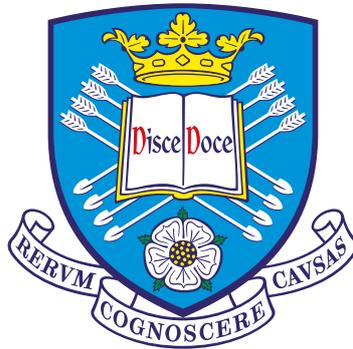


Simulations and Data analysis for the 35 ton Liquid Argon detector as a prototype for the DUNE experiment



Thomas Karl Warburton

Supervisor: Vitaly Kudryavtsev

Department of Physics and Astronomy
University of Sheffield

This dissertation is submitted for the degree of
Doctor of Philosophy

November 2017

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

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this thesis are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This thesis is my own work, and where work has been done in collaboration with others the best attempts have been made to indicate this. Any work which was performed by another individual has been expressed as such, with these figures and tables being labelled with a source.

The information contained in Chapters 1 and 2 are high level summaries of the theory necessary to support this thesis, and the various facets of the DUNE experiment. These summaries are based on a number of articles, that are referenced as such. The LArSoft software which DUNE uses is a multi-experiment project, and many people have contributed to its development, including myself.

Chapter 3 details the camera system which was installed in the 35 ton detector. The selection and basic design of the camera system was principally done by Dr. Nicola McConkey (University of Sheffield), and Mr Matthew Thiesse (University of Sheffield), with only very minor input from the author. All work performed at Fermilab was done in conjunction with Michael Wallbank (University of Sheffield), and would have been impossible without the assistance of Dr. Linda Bagby (Fermi National Laboratory).

Chapter 4 details the simulations performed on the 35 ton detector. All work presented here is my own, though it builds on large pieces of work developed by others, such as the reconstruction of both TPC and photon detector data, as well as the premise of the particle identification method which is proposed. I must thank Dr. Tingjun Yang (Fermi National Laboratory) and Dr. Tom Junk (Fermi National Laboratory) for their helpful insights in developing the methods shown here.

Chapter 5 details all aspects of the 35 ton data which were recorded. The method of data factorisation is my own work, as is the inclusion of a Wiener filter to the data, other work is

referenced to the working group presentations where they were proposed. The final section on electron diffusion is my own work, though I must thank Dr. Michelle Stancari (Fermi National Laboratory) and Dr. Dominic Brailsford (University of Lancaster) for their helpful discussions.

Chapter 6 details the far detector simulations which were performed. Section 6.1 details simulations performed for the LBNE surface detector, where the author's work concentrated on the implementation of the accurate detector geometry and surface profile. Simulations involving the simple geometry, and simple surface profile were performed by Dr. Martin Richardson (University of Sheffield) as part of his PhD work. All work was shown in his thesis, with according references, and all tables have been taken from his thesis, as these were based on an improved analysis to that used when I was working on the project. Section 6.3 details simulations which were performed in establishing the muon-induced background to nucleon decay events in a Liquid Argon Time Projection Chamber (LArTPC) at depth. The generator used in these simulations is accordingly referenced, and the generation of the muon sample which is used was done by Dr. Matthew Robinson (University of Sheffield). The analysis presented is unique to the author, though thanks must be given to all members of the cosmogenic working group for their helpful insights.

Thomas Karl Warburton
November 2017

Acknowledgements

First of all, I would like to express my deepest thanks to my supervisor, Dr. Vitaly Kudryavtsev. I could not have produced this work without your ever watchful guidance and support. It has been a truly wonderful experience working with you for the last three and a half years.

A huge thanks must also go to Dr. Michelle Stancari, Dr. Tingjun Yang, and Dr. Tom Junk, who supervised me during my attachment at Fermilab. I could not have asked for more friendly, knowledgeable and passionate people to be in collaboration with and if I can become the half the physicists that any of you are, I would be immensely proud. I would also like to thank every single member of the dedicated team of people who worked on the 35 ton, you all contributed to my development as a physicist. A special mention must of course go to Michael Wallbank, whom I have spent the majority of my PhD living with and sitting next to, and whose laugh will forever be etched in my psyche. Concerning Fermilab, it is impossible to list all of the amazing (and truly interesting) people that I met there, though every single one of them made my stay at Fermilab two of the best years of my life. I look forward to the next four years being as memorable as the last two!

It would be impossible to not thank everyone in the Particle Physics and Particle Astrophysics group of Sheffield, whose sometimes constant distractions only made to make living Sheffield more enjoyable. Special thanks must go to those people that I shared an office with though, who in no particular order are, Martin, Andy, Matt, Dave, Anthony, Mike, Nicola, Patrick, Celeste and Liam.

Finally, it goes without saying that none of this would have been possible without the fantastic support that I receive from my entire family. I wouldn't swap their feigning of interest and sarcastic comments about Sheldon for anything. It is immeasurably valuable to have a best friend for a sister, a bed of support for a mother, a rock for a nan, and a Pete for a "Pete". None of this would have been possible without each and every one of you.

Abstract

The Deep Underground Neutrino Experiment (DUNE) is a next-generation neutrino experiment which will be built at the Sanford Underground Research Facility (SURF), and will receive a wide-band neutrino beam from Fermilab, 1300 km away. At this baseline DUNE will be able to study many of the properties of neutrino mixing, including the neutrino mass hierarchy and the value of the CP-violating complex phase (δ_{CP}). DUNE will utilise Liquid Argon (LAr) Time Projection Chamber (TPC) (LArTPC) technology, and the Far Detector (FD) will consist of four modules, each containing 17.1 kt of LAr with a fiducial mass of around 10 kt. Each of these FD modules represents around an order of magnitude increase in size, when compared to existing LArTPC experiments.

The 35 ton detector is the first DUNE prototype for the single (LAr) phase design of the FD. There were two running periods, one from November 2013 to February 2014, and a second from November 2015 to March 2016. During the second running period, a system of TPCs was installed, and cosmic-ray data were collected. A method of particle identification was developed using simulations, though this was not applied to the data due to the higher than expected noise level. A new method of determining the interaction time of a track, using the effects of longitudinal diffusion, was developed using the cosmic-ray data. A camera system was also installed in the detector for monitoring purposes, and to look for high voltage breakdowns.

Simulations concerning the muon-induced background rate to nucleon decay are performed, following the incorporation of the MUon Simulations UNderground (MUSUN) generator into the DUNE software framework. A series of cuts which are based on Monte Carlo truth information is developed, designed to reject simulated background events, whilst preserving simulated signal events in the $n \rightarrow K^+ + e^-$ decay channel. No background events are seen to survive the application of these cuts in a sample of 2×10^9 muons, representing 401.6 years of detector live time. This corresponds to an annual background rate of $< 0.44 \text{ events} \cdot \text{Mt}^{-1} \cdot \text{year}^{-1}$ at 90% confidence, using a fiducial mass of 13.8 kt.

Table of contents

List of figures	x
List of tables	xvii
Nomenclature	xix
1 Introduction and Theory	1
1.1 Neutrino physics	2
1.1.1 The history of neutrino oscillations	3
1.1.2 The theory of neutrino oscillations	5
1.1.3 Current experimental limits, and unanswered questions	12
1.2 Grand Unified Theories and nucleon decay	15
1.2.1 Overview of grand unified theories	15
1.2.2 Backgrounds to nucleon decay	17
2 The Deep Underground Neutrino Experiment	20
2.1 Brief experimental overview	20
2.2 The single phase detector design	22
2.3 The physics capabilities of DUNE	27
2.3.1 Neutrino physics	27
2.3.2 Nucleon decay	34
2.3.3 Additional physics opportunities	37
2.4 Path to building DUNE - The 35 ton prototype	38
2.5 The DUNE software	43
3 The 35 ton camera system	53
3.1 The selection and characterisation of cameras	54
3.2 The design of the camera system	55
3.3 Performance in the 35 ton	59

4	Simulations of the 35 ton prototype	63
4.1	Determination of interaction times	63
4.2	Calibrating calorimetric constants	67
4.3	Discerning reconstruction efficiencies	69
4.4	Performing particle identification	78
5	The 35 ton data sample	99
5.1	Organisation of the data structure	99
5.2	Reformatting the data to the offline structure	103
5.3	Observations on data quality and noise mitigation	104
5.4	Performance of reconstruction algorithms	109
5.5	Measuring interaction times using electron diffusion	118
5.5.1	Determining interaction times in 35 ton data	121
5.5.2	Determining interaction times in a low-noise detector using Monte Carlo, and differences with data	131
5.5.3	Impact of changing detector properties using Monte Carlo samples	138
5.5.4	The limitations of and future improvements to the method of interaction time determination using diffusion	150
6	Simulations of the DUNE Far Detector	152
6.1	Simulations of the LBNE surface detector	152
6.1.1	Classifying signal and background events	155
6.1.2	Description of cuts used	156
6.1.3	Generating particles for background studies	158
6.1.4	Results from background simulations	160
6.1.5	Summary of simulations for the LBNE surface detector	165
6.2	MUSUN in LArSoft	166
6.3	Nucleon decay channels in DUNE	168
6.3.1	Cosmogenic background to the $n \rightarrow K^+ + e^-$ decay channel	171
6.3.2	Signal events in the $n \rightarrow K^+ + e^-$ decay channel	178
6.3.3	Kinematics of nucleon decay	191
6.3.4	Energy constraints on the cosmogenic background to the $n \rightarrow K^+ + e^-$ decay channel	192
6.3.5	Future improvements to nucleon decay studies and conclusions	201
7	Concluding remarks	204

References	207
-------------------	------------

Appendix A Supporting figures to Monte Carlo studies concerning determining interaction times using the effects of diffusion	216
---	------------

List of figures

1.1	A schematic representation of the two mass hierarchies, showing the fractional components of flavour states in each mass state	13
1.2	Schematic of a muon event mimicking a nucleon decay	18
2.1	A representation of the DUNE experimental setup	22
2.2	A cross section of the DUNE single phase detector design, showing 8 TPCs	24
2.3	An illustration of the wire wrapping in the DUNE single phase design . . .	26
2.4	The ν_e appearance probability at 1300 km as a function of neutrino energy, for a range of values of δ_{CP}	28
2.5	The significance with which DUNE will be able to determine the neutrino mass hierarchy, for all values of δ_{CP}	29
2.6	The significance with which DUNE will be able to determine the value of δ_{CP} , for all values of δ_{CP}	30
2.7	The resolution with which DUNE will be able to determine the value of δ_{CP} , for increasing beam exposures	31
2.8	The significance to which the octant of θ_{23} can be determined, for different values of θ_{23}	32
2.9	The resolution with which DUNE will be able to determine the value of $\sin^2 \theta_{23}$ for increasing beam exposures	32
2.10	The resolution with which DUNE will be able to determine the value of $\sin^2 \theta_{13}$ for increasing beam exposures	33
2.11	The resolution with which DUNE will be able to determine the value of Δm_{31}^2 for increasing beam exposures	34
2.12	A comparison of current, and future, nucleon decay lifetime limits, compared with the ranges predicted by Grand Unified Theories.	36
2.13	A kaon event which was observed in the ICARUS T600 detector, in the CNGS data	37

2.14	An overview of the DUNE prototyping schedule, including complementary experiments	39
2.15	A schematic representation of the 35 ton prototype detector during the Phase II run	41
2.16	A representation of the cosmic ray counter locations in the 35 ton	42
2.17	The LArSoft coordinate system as it is represented in the 35 ton detector	45
2.18	A selection of reconstructed hits from simulated energy depositions	47
2.19	Performing disambiguation with different wire pitches.	49
3.1	The components which made up a camera module used in the 35 ton camera system	56
3.2	Two camera modules which are mounted in the 35 ton prototype detector	57
3.3	The system diagram for the 35 ton camera system	58
3.4	The calibration images for the 8 cameras used in the 35 ton camera system	60
3.5	The signal degradation over time for two cameras in the 35 ton camera system	61
4.1	Matching tracks and flashes in the 35 ton using positions in the yz plane	65
4.2	The central x position of a reconstructed track versus the number of detected photoelectrons	66
4.3	The predicted x positions of flashes using the relationship between photoelectron and drift distance	66
4.4	The number of events as a function of the difference between Monte Carlo and photon detector times	68
4.5	The calibration of the calorimetric constants in the 35 ton	70
4.6	The reconstruction efficiencies for simulated events as a function of the track length in the detector from Monte Carlo truth.	73
4.7	The reconstruction efficiencies for simulated events as a function of the deposited energy from Monte Carlo truth.	74
4.8	The reconstruction efficiencies for simulated events as a function of the θ track angle from Monte Carlo truth track.	74
4.9	The reconstruction efficiencies for simulated events as a function of the ϕ track angle from Monte Carlo truth track.	75
4.10	The reconstruction efficiencies for simulated events as a function of the θ and ϕ track angles from Monte Carlo truth track.	76
4.11	The mean energy loss per unit track length of different particle masses in different materials	79

4.12	Stopping power for different particle masses as a function of residual range in liquid argon	81
4.13	The distribution of PIDA values, calculated using Monte Carlo truth, for different particle masses	82
4.14	The reconstruction efficiencies for protons in a sample generated using CRY.	83
4.15	The reconstruction efficiencies for the simulated isolated muon and proton samples in the 35 ton detector.	85
4.16	The calculated PIDA values for the simulated isolated proton and muon samples in the 35 ton detector.	87
4.17	The $\frac{dE}{dx}$ versus residual range plot for the simulated isolated proton and muon samples in the 35 ton detector.	88
4.18	The calculated PIDA values for the simulated proton enriched sample in the 35 ton detector	92
4.19	The calculated PIDA values, as a function of the reconstructed track length, for the simulated proton enriched sample in the 35 ton detector	94
4.20	The calculated PIDA values, as a function of the reconstructed track length, for muons in the simulated proton enriched sample in the 35 ton detector	95
4.21	The calculated PIDA values for the simulated proton enriched sample in the 35 ton detector, after a cut on the minimum reconstructed track length is applied	96
4.22	The calculated PIDA values for the simulated proton enriched sample in the 35 ton detector, after tracks from particles which do not stop in the detector are removed	98
5.1	The 35 ton data sample	100
5.2	The 35 ton data structure	102
5.3	Dropped TPC data in the 35 ton	105
5.4	Recovering stuck ADC codes in the 35 ton	106
5.5	Removing coherent noise in the 35 ton	107
5.6	Applying Wiener filters to the 35 ton data	108
5.7	The effect of noise removal algorithms in the 35 ton data	110
5.8	dQ/dx in the 35 ton as a function of drift time	112
5.9	The numbering scheme for the east - west counters in the 35 ton	113
5.10	The dot product of the track and vector joining the centres of the coincidence counters in the yz plane	114
5.11	The alignment of reconstructed tracks with the vectors joining the centres of the coincidence counters	115

5.12	Reconstruction efficiencies of through going tracks in the 35 ton data	117
5.13	Schematic showing the process of diffusion	119
5.14	A simulated event display showing multiple tracks and flashes in the 35 ton detector	120
5.15	The effect of adding a noise baseline to a hit	122
5.16	The distributions of the <i>RMS</i> and <i>RMS/Charge</i> values for tracks with a counter difference of 4 in the 35 ton data	123
5.17	The drift distance dependence of diffusion in the 35 ton dataset for coincidences with a counter difference of 4	124
5.18	The angular dependence of diffusion in the 35 ton dataset for hits within 10 cm of the APAs	125
5.19	The normalised distribution of hit <i>Charge</i> in the 35 ton dataset	126
5.20	The difference between the predicted and reconstructed hit times in the 35 ton dataset	127
5.21	The accuracy of the hit <i>RMS</i> method in the 35 ton dataset	129
5.22	The accuracy of the hit <i>RMS/Charge</i> method in the 35 ton dataset	130
5.23	The distributions of hit <i>RMS</i> and <i>RMS/Charge</i> for tracks with a counter difference of 4 in the 35 ton dataset, and a low noise 35 ton detector	132
5.24	The drift distance dependence of diffusion in the 35 ton dataset and Monte Carlo for coincidences with a counter difference of 4	133
5.25	The angular dependence of diffusion in the 35 ton dataset and Monte Carlo for hits within 10 cm of the APAs	134
5.26	The normalised distribution of hit charge in the 35 ton dataset and a Monte Carlo sample	135
5.27	Comparing the accuracy of the hit <i>RMS</i> method in the 35 ton dataset and a Monte Carlo simulation	136
5.28	Comparing the accuracy of the hit <i>RMS</i> method in the 35 ton dataset and a Monte Carlo simulation	137
5.29	Comparing the accuracy of the hit <i>RMS</i> method, as the electronic noise changes	141
5.30	Comparing the accuracy of the hit <i>RMS</i> method, as the electronic noise level changes	142
5.31	Comparing the accuracy of the hit <i>RMS</i> method, as the electron lifetime changes	143
5.32	Comparing the accuracy of the hit <i>RMS</i> method, as the electron lifetime changes	144
5.33	Comparing the accuracy of the hit <i>RMS</i> method, as the electric field changes	145

5.34	Comparing the accuracy of the hit <i>RMS</i> method, as the electric field changes	146
5.35	Comparing the accuracy of the hit <i>RMS</i> method, as the constant of longitudinal diffusion changes	147
5.36	Comparing the accuracy of the hit <i>RMS</i> method, as the constant of longitudinal diffusion changes	148
6.1	The <i>complex</i> detector geometry used in the LBNE surface detector simulations	154
6.2	A diagram of the PoCA, and the angle w.r.t the beam calculations in the LBNE surface detector simulations	157
6.3	The correlation between the surface profile, and the distribution of azimuthal angles at the DUNE far detector site	167
6.4	The distributions of muon parameters for a sample of 10^7 muons generated by MUSUN in LArSoft	169
6.5	The initial positions of 10^4 muons generated by MUSUN around a DUNE 10 kt module	170
6.6	A simulated $n \rightarrow K^+ + e^-$ decay which occurred in a gap between TPCs . .	175
6.7	The energy distribution of background events surviving the application of sequential cuts.	177
6.8	The energy distribution of background events per MeV of deposited energy surviving the application of sequential cuts	177
6.9	A simulated $n \rightarrow K^+ + e^-$ decay which occurred near the edge of the detector volume	180
6.10	A simulated $n \rightarrow K^+ + e^-$ decay which is fully contained in a single TPC volume	181
6.11	The energy distribution of signal events surviving the application of sequential cuts in the $n \rightarrow K^+ + e^-$ channel	182
6.12	The energy distribution of signal events per MeV of deposited energy surviving the application of sequential cuts	182
6.13	The number of events as a function of the energy deposited by the kaon decay products.	184
6.14	A simulated $n \rightarrow K^+ + e^-$ decay where the kaon did not deposit any energy in the active volume	185
6.15	The number of events, as a function of the energy deposited within 2 cm, 5 cm, and 10 cm of the detector edges.	187
6.16	The separation of the kaon and the electron produced in the simulated $n \rightarrow K^+ + e^-$ decays	189

6.17	The energy distributions per MeV of deposited energy for the signal and cosmic background events, surviving the application of sequential cuts after the fiducial cut is modified	190
6.18	The energy directly deposited by kaons versus the energy directly deposited by electrons, in the simulated nucleon decay and cosmic background samples	195
6.19	The energy directly deposited by kaons, plus the energy directly deposited by electrons versus the energy deposited near the kaon and electron vertex, in the simulated nucleon decay and cosmic background samples	196
6.20	The energy directly deposited by kaons, plus the energy deposited by the kaon decay products versus the energy depositions which do not fit any of the other criteria, in the simulated nucleon decay and cosmic background samples	197
6.21	The energy directly deposited by kaons, plus the energy deposited by the kaon decay products, plus the energy directly deposited by electrons, plus the energy deposited near the kaon and electron vertex versus the energy depositions which do not fit any of the other criteria, in the simulated nucleon decay and cosmic background samples	198
A.1	The distributions of the hit <i>RMS</i> and hit <i>RMS/Charge</i> values for tracks with a counter difference of 4, for different values of the electronics noise	217
A.2	The most probable values of hit <i>RMS</i> as a function of drift distance, for tracks associated with a coincidence that had a counter difference of 4, for different values of the electronics noise	217
A.3	The angular dependence of hits within 10 cm of the APAs, for different values of the electronics noise	218
A.4	The normalised hit charge distribution for different values of the electronics noise	218
A.5	The distributions of the hit <i>RMS</i> and hit <i>RMS/Charge</i> values for tracks with a counter difference of 4, for different values of the electron lifetime	219
A.6	The most probable values of hit <i>RMS</i> as a function of drift distance, for tracks associated with a coincidence that had a counter difference of 4, for different values of the electron lifetime	219
A.7	The angular dependence of hits within 10 cm of the APAs, for different values of the electron lifetime	220
A.8	The normalised hit charge distribution for different values of the electron lifetime	220

A.9	The distributions of the hit <i>RMS</i> and hit <i>RMS/Charge</i> values for tracks with a counter difference of 4, for different values of the electric field	221
A.10	The most probable values of hit <i>RMS</i> as a function of drift distance, for tracks associated with a coincidence that had a counter difference of 4, for different values of the electric field	221
A.11	The angular dependence of hits within 10 cm of the APAs, for different values of the electric field	222
A.12	The normalised hit charge distribution for different values of the electric field	222
A.13	The distributions of the hit <i>RMS</i> and hit <i>RMS/Charge</i> values for tracks with a counter difference of 4, for different values of the constant of longitudinal diffusion	223
A.14	The most probable values of hit <i>RMS</i> as a function of drift distance, for tracks associated with a coincidence that had a counter difference of 4, for different values of the constant of longitudinal diffusion	223
A.15	The angular dependence of hits within 10 cm of the APAs, for different values of the constant of longitudinal diffusion	224
A.16	The normalised hit charge distribution for different values of the constant of longitudinal diffusion	224

List of tables

1.1	Three-flavour oscillation parameters from a fit to global data after the NOW-2016 and ICHEP-2016 conferences	13
2.1	The parameters of the wire planes in the DUNE FD single phase design . .	25
2.2	The beam exposure in units in units of kt·MW·years, as a function of time, assuming a staged DUNE construction	28
2.3	Nucleon decay limits in DUNE and Super-Kamiokande, in some favoured decay channels	36
4.1	Stopping power parameterisation for various particle types in liquid argon .	80
4.2	The properties of initial particles simulated in the muon and proton samples. The angle θ_{xz} , is defined as the angle that a vector makes in the xz plane with the x axis. The angle θ_{yz} , is defined as the angle that a vector make in the yz plane with the z axis.	85
4.3	A summary of the PIDA values calculated for the simulated isolated proton sample, as sequential cuts are applied	89
4.4	A summary of the PIDA values calculated for the simulated isolated muon sample, as sequential cuts are applied	90
4.5	A summary of the PIDA values calculated for the primary particles in the simulated isolated proton sample, as sequential cuts are applied	90
5.1	The angles which tracks, with given counter differences, have relative to the APA frames	116
6.1	The minimum energies of simulated particles, when determining the cosmogenic background for the LBNE surface detector	159
6.2	The normalised background rate per calendar year for the <i>simple</i> detector geometry, using the flat surface profile	160

6.3	The normalised background rate per calendar year, for events where a primary muon enters the active volume of the detector, for the <i>complex</i> geometry and accurate surface profile	161
6.4	The normalised background rate per calendar year, for events where a primary muon misses the active volume of the detector, for the <i>complex</i> geometry and accurate surface profile	163
6.5	The normalised background rate per calendar year, for proton induced events, for the <i>complex</i> geometry	164
6.6	The normalised background rate per calendar year, for neutron induced events, for the <i>complex</i> geometry	164
6.7	The normalised background rate per calendar year for the <i>simple</i> detector geometry and flat surface profile, and for the <i>complex</i> geometry and accurate surface profile	165
6.8	Muon flux parameters as calculated with MUSIC/MUSUN.	167
6.9	The number of events which could mimic a $n \rightarrow K^+ + e^-$ decay, when cuts are applied in isolation	178
6.10	The most common decay modes of charged kaons, and their probabilities	183

Nomenclature

Roman Symbols

tick Unit of time equal to 500 ns

Acronyms / Abbreviations

art *analysis reconstruction framework*

ADC Analogue to Digital Converter

APA Anode Plane Assembly

CCD Charge-Coupled Device

CDR Conceptual Design Report

CMOS Complementary Metal-Oxide Semiconductor

CPA Cathode Plane Assembly

CP Charge-Parity

CRC Cosmic Ray Counter

CRY Cosmic RaY shower library

DUNE Deep Underground Neutrino Experiment

DVR Digital Video Recorder

FD Far Detector

Fermilab Fermi National Laboratory

FGT Fine-Grained Tracker

FNAL	Fermi National Laboratory
GENIE	Generates Events for Neutrino Interaction Experiments
GUT	Grand Unified Theory
HV	High Voltage
LAr	Liquid Argon
LArSoft	Liquid Argon Software
LArTPC	Liquid Argon Time Projection Chamber
LBNE	Long Baseline Neutrino Experiment
LBNO	Long Baseline Neutrino Observatory
LSP	Lightest Supersymmetric Particle
MIP	Minimum Ionising Particle
MPV	Most Probable Value
MSSM	Minimally Supersymmetric Standard Model
MUSIC	MUon Simulation Code
MUSUN	MUon Simulations UNderground
ND	Near Detector
PCB	Printed Circuit Board
PD	Photon Detector
PID	Particle IDentification A
PID	Particle IDentification
PIP-II	Proton Improvement Plan II
PoCA	Point of Closest Approach
RCE	Reconfigurable Computing Element
ROI	Region Of Interest

SiPM Silicon PhotoMultiplier

SK Super-Kamiokande

SM Standard Model

SSP SiPM Signal Processor

SURF Sanford Underground Research Facility

SUSY SuperSYmmetry

TPC Time Projection Chamber

Chapter 1

Introduction and Theory

The “Standard Model of Particle Physics” (SM) is a set of theories which has been widely tested and has been found to accurately predict the interactions of fundamental particles. These tests have come in many forms throughout the 20th and 21st centuries, and include the detection of all of the quarks and leptons which it predicts, as well as measurements of the properties of these particles. The recent discovery of the Higgs boson [1, 2] “completed” the SM, as this was the last particle which it predicted to be observed. However, despite its many successes, the SM does not represent the “final” theory of fundamental particle physics, should one exist. This is because there are many questions made by recent experimental observations which the SM is unable to address, some of these will be briefly discussed below.

Firstly, though the SM accurately predicts the interactions made by the electromagnetic, weak nuclear, and strong nuclear forces, it makes no mention of gravity. This is a major flaw of the SM as gravity is one of the driving forces in the formation of astronomical objects such as planets, stars, and galaxies. With the recent detection of gravitational waves [3] this issue has again be brought into focus. Secondly, the rotational velocities of galaxies is measured to be far greater than the predicted value, hinting at the presence of a significant amount of matter which we are unable to detect. The SM makes no prediction as to what this “dark-matter” is comprised of. Thirdly, measurements of distant supernovae appear to show that the expansion of the universe is accelerating, not decelerating as would be expected, this implies the presence of some form of unknown energy source. Again, the SM makes no prediction as to what this unknown energy source, or “dark-energy” is. A further point of consternation with the SM is that it is not “elegant,” as it has as many as 19 free parameters, which appear to be unrelated to each other. There are also unresolved questions regarding the particles which are predicted by the SM, such as, why charge is quantised, why there are exactly 3 families of quarks and leptons, and why they have the observed hierarchy of

masses. The SM also does not predict the matter-antimatter asymmetry which is observed in the universe today. Finally, the neutrinos predicted by the SM are massless, however, numerous measurements of neutrino oscillations show that neutrinos have a non-zero rest mass. A rigorous discussion of neutrino oscillations is presented in Section 1.1.

Extensive efforts have been made to resolve many of these issues with the SM, in the form of a so-called Grand Unified Theory (GUT). These theories propose that the electroweak and strong nuclear forces belong to an overarching symmetry group. The unification of these forces is predicted to occur at extremely high energies, far beyond the reach of current experiments. As a result, many of the experimental signatures which these GUTs predict are difficult to measure. However, many GUTs predict that the proton, a stable particle in the SM, should decay with a lifetime of around 10^{30-36} years, though some models predict much longer lifetimes. Some of the gauge groups which are invoked by GUTs are briefly discussed in Section 1.2.1, with reference to the proton lifetimes which they predict. A discussion of the backgrounds to proton decay searches is presented in Section 1.2.2.

The Deep Underground Neutrino Experiment (DUNE) is a next generation experiment to be built at the Sanford Underground Research Facility (SURF), which aims to measure many of the properties of neutrinos, as well as to search for nucleon decays. The experimental setup, physics capabilities, and prototyping schedule for DUNE are outlined in Chapter 2. A camera system which was installed in the DUNE 35 ton prototype detector is described in Chapter 3. Chapter 4 describes simulations which were made in preparation for data taking of the 35 ton prototype, and concludes with a description of how Particle IDentification (PID) could be performed in the 35 ton data. An overview of the data gathered by the 35 ton prototype is shown in Chapter 5, and a novel method of interaction time determination using the effects of diffusion is presented. Following this, Chapter 6 concerns simulations of the cosmogenic backgrounds seen in a large Liquid Argon Time Projection Chamber (LArTPC) at SURF. These simulations are first presented with respect to a surface detector measuring neutrino oscillations, and then to a detector at depth searching for nucleon decay events. Finally, Chapter 7 contains some final remarks and observations.

1.1 Neutrino physics

The study of neutrinos offers a chance to probe the limitations of the SM, as the neutrinos predicted by the SM are massless and do not oscillate. However, numerous measurements have

shown that neutrino oscillations occur, and that at least two of the neutrino mass eigenstates have non-zero mass. Notably, the 2015 Nobel prize in physics was given to T. Kajita and A. McDonald for “the discovery of neutrino oscillations, which shows that neutrinos have mass,” based on their work on Super-Kamiokande (SK) [4] and SNO [5] respectively. This means that through studying neutrino oscillations, it is possible to begin to get a handle on physics beyond the SM. The history of the discovery of neutrino oscillations, which culminated in this Nobel prize, is briefly outlined in Section 1.1.1. Following this, the formalism by which neutrino oscillations occur is presented in Section 1.1.2. Finally, the current state of neutrino physics, including the current best fit values for the various mixing parameters, is summarised in Section 1.1.3.

1.1.1 The history of neutrino oscillations

Neutrinos were first proposed to explain the continuous energy spectrum of the electrons produced in β decay, as due to kinematic constraints, it could not be explained by a two body decay. To this end, Pauli proposed the idea of a neutral particle, with mass less than that of the electron, which would not be observed in the reaction [6]. Pauli called this particle a “neutron.” Upon the discovery that the “neutron” was in fact of a similar mass to the proton, and that the nucleus was a bound state of protons and neutrons, Fermi proposed a more complete theory of β decay in 1934. In this theory, Fermi proposed that the light, neutral particle that was initially proposed by Pauli did exist, and was emitted from the nucleus in the reaction. Fermi called this particle a “neutrino,” meaning “little neutral one” in Italian. He also proposed that its mass could be measured by looking at the end point of the β spectrum [7]. The first experiments designed to measure the neutrino mass in this way set an upper mass limit of 500 eV [8, 9], which was improved to 250 eV in the 1950’s [10]. After becoming evident that the neutrino mass was so much less than that of the electron, the idea that neutrinos were massless gained traction.

The first direct observation of neutrinos was in 1956 [11], and in 1962 conclusive proof emerged that the electron and muon neutrinos were distinct particles [12]. The experiment which found this, did so by observing that it was far more likely that the neutrinos produced in the decay of pions would interact to create muons, as opposed to electrons. This meant that there had to be two flavours of neutrinos, as if there was only a single flavour of neutrino, then they should produce equal numbers of electrons and muons when they interact. However, soon after this an experiment by Ray Davis in 1968 at the Homestake Mine gave rise to the “Solar neutrino problem” [13]. The Homestake experiment used a

380 m³ tank of tetrachloroethene (C₂Cl₄), 4850 ft below ground, to look for the electron neutrinos produced by the sun. It did this by measuring the yield of Ar³⁷ which was produced by inverse beta decay reactions on the chlorine atoms. The neutrino flux could be determined by extracting the Ar³⁷ produced in the reaction, and then counting the number of radioactive decays it produced, as Ar³⁷ is unstable. The neutrino flux which was measured was roughly $\frac{1}{3}$ of the predicted flux from solar models. The experiment ran for over 20 years, with the measured ν_e flux being unchanged at roughly $\frac{1}{3}$ of the predicted solar flux [14].

The long standing observation that the solar ν_e flux was significantly lower than predicted meant that, either there was some mechanism by which the electron neutrinos were evading detection, or that the solar model was incorrect. It was plausible that the solar model was incorrect, however the scale of the difference in the observed and predicted ν_e fluxes proved difficult to resolve. As a result, the idea of neutrino oscillation grew momentum, drawing on a prediction made by Pontecorvo as far back as 1957 [15]. The Kamiokande-II experiment measured high energy solar neutrinos, and in 1989 measured an energy dependant deficit in the solar ν_e flux [16]. When studying the atmospheric neutrino flux, Kamiokande-II found an angular dependent deficit in the expected muon neutrino flux, though the electron neutrino flux was consistent with predictions. It was found that this deficit was consistent with oscillations of $\nu_\mu \leftrightarrow \nu_\tau$ [4]. The tau lepton, the third flavour of leptons, had been observed in the 1970s [17], though the ν_τ neutrino was not directly measured until 2000 [18].

The SNO experiment, which was a Cherenkov experiment that used 1,000 tons of heavy water (D₂O) surrounded by 9,600 PhotoMultiplier Tubes (PMTs), provided conclusive proof of neutrino oscillations in 2001. SNO was able to do this because the use of heavy water allowed it to be sensitive to both the Neutral Current (NC) and Charged Current (CC) interactions of solar neutrinos. Neutral current interactions are sensitive to neutrinos of all flavours, and require measuring the gamma ray produced when the emitted neutron is captured by either a deuteron or hydrogen nucleus. However, only electron neutrinos can undergo CC interactions, as solar neutrinos do not have enough energy to produce muon or tau leptons. SNO measured a charge current interaction rate which was consistent with earlier experiments [19], i.e. a deficit in the predicted solar flux. However, it found that the neutral current interaction rate matched the predicted solar flux [5]. This demonstrated that a significant part of the ν_e flux from the sun, had oscillated into ν_μ and ν_τ as they travelled to Earth. As described above, this flux of oscillated ν_μ and ν_τ could not interact via CC interactions, due to the high mass of the associated leptons relative to the neutrino energy,

however, they were able to interact via NC interactions.

These highlighted results, as well as many other accompanying results, form the basis of our current understanding of neutrino oscillations.

1.1.2 The theory of neutrino oscillations

Neutrino oscillations are described by the PMNS matrix, which is named after work initially done by Pontecorvo [15], and then later extended by Maki, Nakagawa and Sakata [20]. The PMNS matrix describes neutrino mixing in the context of the three known flavour states ν_e , ν_μ and ν_τ being related to three neutrino mass states ν_1 , ν_2 and ν_3 . This formalism then has the form shown in Equation 1.1.

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = U_{PMNS} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix} \quad (1.1)$$

The matrix labelled U_{PMNS} in Equation 1.1, is then expressed by Equations 1.2, 1.3 and 1.4. In these equations it has been assumed that U_{PMNS} is unitary, and that there are three angles (θ_{12} , θ_{13} and θ_{23}) plus a CP-violating phase δ , which explain the mixing between the mass and flavour eigenstates. The notation $s_{\alpha\beta}$ and $c_{\alpha\beta}$ has been used to denote $\sin \theta_{\alpha\beta}$ and $\cos \theta_{\alpha\beta}$ respectively.

$$U_{PMNS} = \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu1} & U_{\mu2} & U_{\mu3} \\ U_{\tau1} & U_{\tau2} & U_{\tau3} \end{pmatrix} \quad (1.2)$$

$$= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}s_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix} \quad (1.3)$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (1.4)$$

Equation 1.2 shows how each element in the U_{PMNS} relates the flavour states to the mass states, whilst Equation 1.3 shows the full mixing formalism for Dirac neutrinos. Finally,

Equation 1.4 separates the full formalism into three 3×3 matrices which each contain one of the three mixing angles.

Should neutrinos be Majorana particles, then the U_{PMNS} matrices should be multiplied by $\text{diag}(e^{i\alpha_1/2}, e^{i\alpha_2/2}, 1)$. The question of whether neutrinos are Majorana or Dirac particles does raise important questions for neutrino physics. This is because should neutrinos be Majorana particles, their masses could be generated via a Majorana mass term. There are next generation experiments such as SNO+ [21] and SuperNEMO [22], which will search for neutrinoless double beta decay as a means to test whether neutrinos are Majorana particles. SNO+ will replace the heavy water used in SNO with a liquid scintillator that is doped with ^{130}Te , which is a double beta decaying material. After 5 years of data taking SNO+ will be able to set a lower limit on the half-life of ^{130}Te of 9×10^{25} years at 90% confidence level, which will correspond to a limit on the effective Majorana mass of between 55 - 133 meV. SuperNEMO will aim to measure neutrinoless double beta decays from several isotopes, and will aim to establish half-lives of over 10^{26} years. However, the two Majorana phases do not affect neutrino oscillations, and so will not be covered further in this discussion of neutrino oscillations.

This results in the neutrino mixing matrix being constrained by four independent parameters:

- Three mixing angles (θ_{13} , θ_{12} , θ_{23}).
- The CP-violating phase (δ).

When neutrinos are produced and detected, we observe neutrinos of distinct flavour states, and not the distinct mass states. Therefore, a discussion of how mixing occurs will be presented in terms of an initial neutrino composed of a distinct flavour state, and multiple mass states. In this case a neutrino ν_α will be produced, of flavour α , which is a linear superposition of the three mass eigenstates ν_j , such that $j = 1, 2, 3$ (Equation 1.5).

$$|\nu_\alpha\rangle = \sum_j U_{\alpha j}^* |\nu_j\rangle \quad (1.5)$$

where $U_{\alpha j}^*$ is one of the elements in Equation 1.2. As this neutrino propagates, the mass eigenstates will evolve according to the time-dependant Schrödinger equation, such that after time t , each mass eigenstate will have the form shown in Equation 1.6.

$$|\nu_j(t)\rangle = e^{-i(E_j t - \vec{p}_j \cdot \vec{x}_j)} |\nu_j(0)\rangle \quad (1.6)$$

where assuming that the neutrino is ultra-relativistic, and setting $\hbar = c = 1$:

$$t \approx L \quad (1.7)$$

$$E = \sqrt{p^2 + m^2} = p \times \sqrt{1 + \frac{m^2}{p^2}} \approx p + \frac{m^2}{2p} \approx p + \frac{m^2}{2E} \quad (1.8)$$

$$E_j \cdot t - \vec{p}_j \cdot \vec{x}_j \approx \vec{p}_j L + \frac{m_j^2}{2E} L - \vec{p}_j L = \frac{m_j^2}{2E} L \quad (1.9)$$

where a Taylor series expansion of $\sqrt{1+x^2}$ has been used in Equation 1.8. Substituting Equation 1.9 into Equation 1.6 gives Equation 1.10.

$$|\nu_j(t)\rangle = e^{-im_j^2 L/2E} |\nu_j(0)\rangle \quad (1.10)$$

This then gives the time evolution of the original neutrino flavour state as Equation 1.11.

$$|\nu_\alpha(t)\rangle = \sum_j U_{\alpha j}^* e^{-im_j^2 L/2E} |\nu_j(0)\rangle \quad (1.11)$$

From this, it can be seen that the mass states propagate with different phases, and so should the neutrino be detected at a later time it would exist as a superposition of different flavour states. This results in there being a non-zero possibility that the flavour of the neutrino which is detected, β , is not the same as the flavour with which the neutrino was produced with, α . The amplitude with which this occurs is given by Equation 1.12.

$$\begin{aligned} A(\nu_\alpha \rightarrow \nu_\beta) &= \langle \nu_\beta | \nu_\alpha(t) \rangle \\ &= \sum_k \sum_j \langle \nu_j | U_{\beta j} U_{\alpha k}^* e^{-im_k^2 L/2E} | \nu_k \rangle \\ &= \sum_k U_{\alpha k}^* U_{\beta k} e^{-im_k^2 L/2E} \end{aligned} \quad (1.12)$$

Equation 1.12, can then be used to get the probability for the original neutrino ν_α to oscillate to a different flavour ν_β :

$$P(\nu_\alpha \rightarrow \nu_\beta) = |\langle \nu_\beta | \nu_\alpha(t) \rangle|^2 = \left| \sum_k U_{\alpha k}^* U_{\beta k} e^{-im_k^2 L/2E} \right|^2 \quad (1.13)$$

$$= \sum_k U_{\alpha k}^* U_{\beta k} e^{-im_k^2 L/2E} \sum_j U_{\alpha j}^* U_{\beta j} e^{-im_j^2 L/2E} = \sum_k \sum_j U_{\alpha k}^* U_{\beta k} U_{\alpha j}^* U_{\beta j} \exp\left(-i \frac{(m_k^2 - m_j^2)L}{2E}\right) \quad (1.14)$$

The $(m_k^2 - m_j^2)$ term in Equation 1.14 is often written as Δm_{kj}^2 , and will be written as such for the remainder of the discussion of neutrino mixing.

In the interests of simplicity, an explicit calculation of the neutrino oscillation probability will be given assuming that there are only two neutrino flavour and mass states. The reason for this is that then there is only one mixing angle (θ), and no complex phase. We are also free to choose the simplest mixing matrix, such that the mixing matrix becomes Equation 1.15.

$$\begin{pmatrix} \nu_\alpha \\ \nu_\beta \end{pmatrix} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \end{pmatrix} \quad (1.15)$$

The probability of a neutrino oscillating from initial flavour ν_α to flavour ν_β is then given by Equation 1.16, which starts from Equation 1.13.

$$\begin{aligned}
P(\nu_\alpha \rightarrow \nu_\beta) &= \left| (U_{\alpha 1} U_{\beta 1} e^{-im_1^2 L/2E}) + (U_{\alpha 2} U_{\beta 2} e^{-im_2^2 L/2E}) \right|^2 \\
&= \left| (\cos \theta)(-\sin \theta) e^{-im_1^2 L/2E} + (\sin \theta)(\cos \theta) e^{-im_2^2 L/2E} \right|^2 \\
&= 2 \cos^2 \theta \sin^2 \theta - \cos^2 \theta \sin^2 \theta \left[e^{-i \frac{(m_1^2 - m_2^2)L}{2E}} + e^{-i \frac{(m_2^2 - m_1^2)L}{2E}} \right] \\
&\text{using } \cos(\phi_1 - \phi_2) = \left(e^{i(\phi_1 - \phi_2)} + e^{-i(\phi_1 - \phi_2)} \right) / 2 \\
&= 2 \cos^2 \theta \sin^2 \theta - \cos^2 \theta \sin^2 \theta \left[2 \cos \left(\frac{(m_1^2 - m_2^2)L}{2E} \right) \right] \\
&= 2 \cos^2 \theta \sin^2 \theta \left[1 - \cos \left(\frac{\Delta m_{12}^2 L}{2E} \right) \right] \\
&\text{using } \cos \theta \sin \theta = \frac{1}{2} \sin(2\theta) \text{ and } 2 \sin^2(\theta) = 1 - \cos(2\theta) \\
&= \sin^2 2\theta \sin^2 \left(\frac{\Delta m_{12}^2 L}{4E} \right) \tag{1.16}
\end{aligned}$$

Measuring L in km, E in GeV, and using proper values and units of \hbar, c

$$= \sin^2 2\theta \sin^2 \left(1.27 \Delta m_{12}^2 \frac{L}{E} \right) \tag{1.17}$$

The presence of the Δm_{kj}^2 terms in Equations 1.14 and 1.16 is the reason why neutrino oscillation implies that at least two neutrinos are massive (at least one neutrino in the 2 neutrino case). This is because if the Δm_{kj}^2 term is 0, then the probability of oscillation is 0, and so oscillations would not occur.

It can often be assumed that there are only two flavours of neutrinos, as recent experimental data shows that the mass splittings of the three neutrino flavours are separated by many orders of magnitude. The current best fit experimental values are shown in Section 1.1.3, though it is sufficient to say $\Delta m_{23}^2 \approx \Delta m_{13}^2 \gg \Delta m_{12}^2$. For example, in the case of atmospheric neutrinos, oscillations are largely due to $\nu_\mu \rightarrow \nu_\tau$. When explaining the observed deficit of upward going ν_μ the L/E of oscillations must be considered, as the other terms in Equation 1.17 are all constants. When considering an initial ν_μ of energy 1 GeV, the oscillation probability will be small for down-going neutrinos ($L = 10$ km), however for upwards-going muons ($L = 10^4$ km) the oscillation probability is much larger, at around 50%. This is what is seen in experimental data.

When measuring neutrino oscillations using neutrinos produced by accelerators, one also has to consider the most optimal L/E . These experiments measure either the disappearance, or appearance of neutrino flavour states. MINOS [23] and MINOS+ [24] are examples of disappearance experiments, they measured the ratio of ν_μ neutrinos at their magnetised near and far detectors, both of which were made of thick iron plates interspersed with scintillator bars. However, appearance experiments such as DUNE primarily measure the appearance of ν_e from an initially pure ν_μ beam.

There are of course three flavours of neutrinos, not two as assumed above. Introducing an extra neutrino mass and flavour state makes the oscillation probabilities more complicated, due to the addition of an extra mixing angle and a complex phase. The oscillation probability for the three flavour neutrino case is shown in Equation 1.18.

$$P(\nu_\alpha \rightarrow \nu_\beta) = \delta_{\alpha\beta} - 4 \sum_{i<j} \text{Re}(U_{\alpha i}^* U_{\alpha j} U_{\beta i} U_{\beta j}) \sin^2 \left(\frac{\Delta m_{ij}^2 L}{4E} \right) + 2 \sum_{i<j} \text{Im}(U_{\alpha i}^* U_{\alpha j} U_{\beta i} U_{\beta j}) \sin \left(2 \frac{\Delta m_{ij}^2 L}{4E} \right) \quad (1.18)$$

Equation 1.18 considers oscillations in a vacuum, however DUNE will measure $\nu_\mu \rightarrow \nu_e$ oscillations after the neutrinos have passed through a large amount of matter. Measuring neutrinos which have passed through matter adds further complications to the observed oscillation probability. The full oscillation probability for $\nu_\mu \rightarrow \nu_e$ oscillations, after they have passed through matter of constant density, is shown in Equation 1.19 [25].

$$P(\nu_\mu \rightarrow \nu_e) \simeq \sin^2 \theta_{23} \sin^2 2\theta_{13} \frac{\sin^2(\Delta_{31} - aL)}{(\Delta_{31} - aL)^2} \Delta_{31}^2 + \sin 2\theta_{23} \sin 2\theta_{13} \sin 2\theta_{12} \frac{\sin(\Delta_{31} - aL)}{(\Delta_{31} - aL)} \Delta_{31} \frac{\sin(aL)}{(aL)} \Delta_{12} \cos(\Delta_{31} + \delta_{CP}) + \cos^2 \theta_{23} \sin^2 2\theta_{12} \frac{\sin^2(aL)}{(aL)^2} \Delta_{12}^2 \quad (1.19)$$

where $\Delta_{ij} = \Delta m_{ij}^2 L / 4E$, δ_{CP} is the CP-violating phase, and $a = G_F N_e / \sqrt{2}$ with G_F being the Fermi constant and N_e being the number density of electrons in the Earth. From Equation 1.19 it can be seen that the probability of neutrino oscillations is governed by the following parameters:

- Three mixing angles (θ_{13} , θ_{12} , θ_{23}).
- Three mass squared differences (Δm_{12}^2 , Δm_{13}^2 , Δm_{23}^2).

- The CP-violating phase (δ).
- The distance the neutrino travels (L).
- The energy of the neutrino (E).
- The magnitude of the matter effects (a).

Of these parameters, the mixing angles, mass squared differences, and δ_{CP} are fixed and can only be measured. The distance travelled by the neutrino and the neutrino energy vary, though may be chosen by the experiment, in the case of accelerator or reactor based experiments. The factor a involved in the matter effect cannot be controlled by the experiment, and arises naturally when neutrinos travel through matter. The CP-violating phase (δ_{CP}), and the matter effect (a), have been introduced in Equation 1.19 and will be briefly discussed below.

CP violation occurs when a particle and its anti-particle behave differently. This has been observed in the quark sector, where the CP violating phase δ_{CP}^{CKM} has been measured to be approximately 70° [26]. Despite this large CP-violating phase (90° represents a maximal CP violating phase), CP-violation in the quark sector is relatively small due to the small mixing present in the CKM matrix [27, 28]. The CKM matrix describes quark mixing, in the same way that the PMNS matrix describes neutrino mixing. This relative lack of CP-violation makes it difficult to explain the observed matter-antimatter asymmetry in the Universe, and so it is widely hoped that measuring a large leptogenic CP-violation may help to explain this. A large amount of CP-violation is possible in the neutrino sector due to the large mixing angles, meaning that the effects of CP-violation will not be suppressed as much as they are in the quark sector.

It can be seen from Equation 1.19 that in order for δ_{CP} to be measured, all three mixing angles (θ_{ij}) must be nonzero. This has now been confirmed by experiments, as will be shown in Section 1.1.3. CP-violation in neutrinos can be observed by measuring the differences in the oscillations of $\nu_\mu \rightarrow \nu_e$ and $\bar{\nu}_\mu \rightarrow \bar{\nu}_e$. The asymmetry between neutrino and antineutrino oscillations is given by Equation 1.20.

$$A_{CP} = \frac{P(\nu_\mu \rightarrow \nu_e) - P(\bar{\nu}_\mu \rightarrow \bar{\nu}_e)}{P(\nu_\mu \rightarrow \nu_e) + P(\bar{\nu}_\mu \rightarrow \bar{\nu}_e)} \quad (1.20)$$

However, the observed asymmetry between $\nu_\mu \rightarrow \nu_e$ and $\bar{\nu}_\mu \rightarrow \bar{\nu}_e$ oscillations will also be affected by matter effects, and so they must also be accurately understood before the value of δ_{CP} can be calculated.

The Mikheyev-Smirnov-Wolfenstein (MSW) effect [29, 30], or matter effect, causes the effective mass of neutrinos to increase due to coherent scattering. The increase in the effective mass of each neutrino is different though, and means that the effective mass splittings in matter become different from those in a vacuum. This leads to the possibility of measuring the absolute mass differences between different mass states, whilst the oscillation probabilities presented earlier were only sensitive to the squared mass differences. The increase in the effective mass of the ν_e is largest, as it is able to undergo charged current interactions with electrons in matter. This is not the case for $\bar{\nu}_e$ as there are few positrons in matter. This leads to the MSW effect being able to mimic the effects of CP violation.

The MSW effect is observed in solar neutrinos, where the ν_e which are produced in the core oscillate to ν_μ and ν_τ as they travel through dense regions of matter in the sun, causing the oscillation probability to increase. By comparing measurements made by experiments looking at solar and reactor neutrinos, it can be seen that the m_2 neutrino mass state is heavier than the m_1 mass state. This is because if the m_1 neutrino mass state was heavier, the oscillation probability would decrease, not increase [31]. In this discussion, the m_1 neutrino mass state has been defined to be the mass state which contains the largest component of ν_e .

However, it has not yet been determined whether the m_3 mass state is lighter or heavier than the $m_{1,2}$ mass states. This results in there being two potential mass hierarchies, a normal mass hierarchy where the mass states are ordered such that $m_1 < m_2 < m_3$, or an inverted mass hierarchy where the mass states are ordered such that $m_3 < m_1 < m_2$. The normal and inverted hierarchies are sometimes also referred to as normal and inverted orderings. Figure 1.1 is a schematic representation of the mass hierarchies, showing the fractional component of flavour states in each mass state.

1.1.3 Current experimental limits, and unanswered questions

There is a large amount of experimental data supporting the oscillation paradigm which has been outlined in Section 1.1.2. This is summarised in Table 1.1 [33], which combines measurements made by many neutrino experiments to produce global fits for the neutrino mixing parameters.

As can be seen from Table 1.1, the three mixing angles, and the mass squared differences are known to quite high precision. This could lead one to believe that the physics of neutrino oscillation is completely understood, however there are still many questions which remain

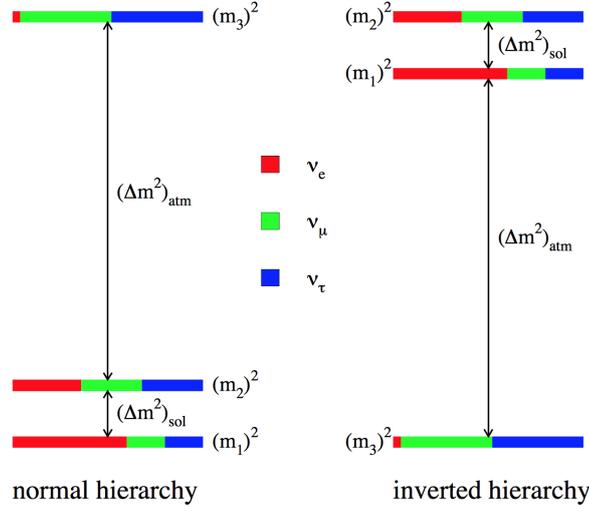


Fig. 1.1 A schematic representation of the two mass hierarchies, showing the fractional components of flavour states in each mass state. Left: the normal hierarchy, where the masses are ordered $m_1 < m_2 < m_3$. Right: the inverted mass hierarchy, where the masses are ordered $m_3 < m_1 < m_2$. The fractional components of ν_e , ν_μ and ν_τ are shown in red, green and blue respectively. Figure taken from [32].

Table 1.1 Three-flavour oscillation parameters from a fit to global data after the NOW-2016 and ICHEP-2016 conferences. The best fit points (bfp) in the 1st (2nd) column are obtained assuming normal (inverted) ordering, whereas in the 3rd column no ordering is assumed. Note that $\Delta m_{3\ell}^2 \equiv \Delta m_{31}^2 > 0$ for normal ordering, and $\Delta m_{3\ell}^2 \equiv \Delta m_{32}^2 < 0$ for inverted ordering. Table is taken in full from [33].

	Normal Ordering (best fit)		Inverted Ordering ($\Delta\chi^2 = 0.83$)		Any Ordering 3σ range
	bfp $\pm 1\sigma$	3σ range	bfp $\pm 1\sigma$	3σ range	
$\sin^2 \theta_{12}$	$0.306^{+0.012}_{-0.012}$	$0.271 \rightarrow 0.345$	$0.306^{+0.012}_{-0.012}$	$0.271 \rightarrow 0.345$	$0.271 \rightarrow 0.345$
$\theta_{12} (^\circ)$	$33.56^{+0.77}_{-0.75}$	$31.38 \rightarrow 35.99$	$33.56^{+0.77}_{-0.75}$	$31.38 \rightarrow 35.99$	$31.38 \rightarrow 35.99$
$\sin^2 \theta_{23}$	$0.441^{+0.027}_{-0.021}$	$0.385 \rightarrow 0.635$	$0.587^{+0.020}_{-0.024}$	$0.393 \rightarrow 0.640$	$0.385 \rightarrow 0.638$
$\theta_{23} (^\circ)$	$41.6^{+1.5}_{-1.2}$	$38.4 \rightarrow 52.8$	$50.0^{+1.1}_{-1.4}$	$38.8 \rightarrow 53.1$	$38.4 \rightarrow 53.0$
$\sin^2 \theta_{13}$	$0.02166^{+0.00075}_{-0.00075}$	$0.01934 \rightarrow 0.02392$	$0.02179^{+0.00076}_{-0.00076}$	$0.01953 \rightarrow 0.02408$	$0.01934 \rightarrow 0.02397$
$\theta_{13} (^\circ)$	$8.46^{+0.15}_{-0.15}$	$7.99 \rightarrow 8.90$	$8.49^{+0.15}_{-0.15}$	$8.03 \rightarrow 8.93$	$7.99 \rightarrow 8.91$
$\delta_{\text{CP}} (^\circ)$	261^{+51}_{-59}	$0 \rightarrow 360$	277^{+40}_{-46}	$145 \rightarrow 391$	$0 \rightarrow 360$
$\Delta m_{21}^2 \times 10^5 (\text{eV}^2)$	$7.50^{+0.19}_{-0.17}$	$7.03 \rightarrow 8.09$	$7.50^{+0.19}_{-0.17}$	$7.03 \rightarrow 8.09$	$7.03 \rightarrow 8.09$
$\Delta m_{3\ell}^2 \times 10^3 (\text{eV}^2)$	$+2.524^{+0.039}_{-0.040}$	$+2.407 \rightarrow +2.643$	$-2.514^{+0.038}_{-0.041}$	$-2.635 \rightarrow -2.399$	$[+2.407 \rightarrow +2.643]$ $[-2.629 \rightarrow -2.405]$

unanswered. Some of these have already been discussed earlier, and in no particular order are;

1. **What is the octant of θ_{23} ?** As can be seen from Table 1.1, θ_{23} is very close to 45° , representing maximal mixing. However, whether it is more or less than 45° remains to be seen.
2. **What is the value of δ_{CP} ?** The value of δ_{CP} is largely unconstrained to 3σ , and as discussed in Section 1.1.2, the observation of CP violation in neutrinos may help to explain the matter-antimatter asymmetry in the universe.
3. **What is the neutrino mass hierarchy?** As discussed in Section 1.1.2, though it is known that the m_1 mass state is lighter than the m_2 mass state, it is not known whether these are lighter or heavier than the ν_3 mass state. This is an essential step in determining the value of δ_{CP} as the differences between interactions involving neutrinos and antineutrinos can mimic the effect of δ_{CP} . It would also show that any beyond the standard model theories which predicted the incorrect mass hierarchy were wrong.
4. **What is the absolute neutrino mass scale?** The absolute mass of the neutrino is unknown, though cosmological constraints from Planck CMB data, such as the matter distribution in the universe and galaxy clustering, place an upper limit of $\sum m_\nu < 0.23$ eV at the 95% confidence limit [34]. Neutrino experiments such as KATRIN hope to reach sensitivities of this magnitude by looking for a cut-off in the energy spectrum of β decays [35].
5. **Are neutrinos Dirac or Majorana particles?** This concerns how neutrinos acquire mass, as if they are Dirac particles then they would acquire mass through interactions with the Higgs field, as the other particles in the SM do. However, if they were Majorana particles, then they would acquire at least some of their mass through self-coupling. This would violate lepton number, and mean that neutrinos are their own antiparticle. As discussed in Section 1.1.2, experiments are searching for this by looking for neutrinoless double β decay.
6. **Are there sterile neutrinos?** Experimental data from LEP tightly constrains the number of active neutrinos to those predicted by the SM [36]. However, in recent years there have been anomalous results which may hint at so-called sterile neutrinos [37–39]. These sterile neutrinos would not interact, but would still be involved in oscillations. Sterile neutrinos are included in the neutrino mixing formalism by extending the PMNS

matrix from a 3×3 matrix, to a $(3+N) \times (3+N)$ matrix, where N is the number of sterile neutrinos. Many short baseline experiments are searching for signatures which would suggest the presence of sterile neutrinos.

Next generation neutrino experiments, such as DUNE, hope to address some of these unanswered questions. For a review of the experimental aims of DUNE see Section 2.3.

1.2 Grand Unified Theories and nucleon decay

The subject of Grand Unified Theories (GUTs) is a complicated one and requires a level of mathematics which will not be covered here. However, as Chapter 6 concerns establishing background limits to nucleon decays, it is necessary to illustrate the basic premise of GUTs and the predictions on nucleon decay which they make; this will be presented in Section 1.2.1. The mechanism by which cosmic ray muons are able to produce signals which mimic nucleon decays will be presented in Section 1.2.2.

1.2.1 Overview of grand unified theories

The basic premise of GUTs is that they attempt to unite the strong, weak and electromagnetic forces. This is achieved by referring to very large energy scales of around 10^{16} GeV [40]. One of the first GUTs was that of Georgi and Glashow in 1974, which predicted that the three forces arose from a single interaction based on the $SU(5)$ gauge group [41]. One of the things which their theory predicted was that the proton would not be stable, and would have a lifetime $\tau_p \simeq 10^{30}$ years. This went against the long-held idea of baryon number conservation, which had been proposed by Weyl to explain why neutrons decayed but protons did not [42]. Proton decay had been considered since Weyl's prediction, but there had never been any prediction of it. One of the earliest limits on the lifetime of the proton had actually been made by Maurice Goldhaber, who noted that if the proton lifetime was less than 10^{18} years we would receive lethal doses of radiation from its decay [43].

With the prediction of proton decay, experiments began searching for it in underground labs. The proton lifetime predicted by Georgi and Glashow has now been conclusively ruled out, and inconsistencies have been found with their original theory, such as the gauge couplings not unifying at a common energy, and the neutrinos predicted by it being massless [43]. However, it began the search for a GUT, as well as the search for nucleon decay, and so

it is interesting from an historical standpoint. Extensions to the original theory have been made which attempt to address some of the issues mentioned above, these extensions predict proton lifetimes of the order $\tau_p \leq 10^{35}$ years [44, 45].

SUperSymmetry (SUSY) is an extension of the standard model which aims to rectify many of the inconsistencies seen in the SU(5) models. It does this by predicting that there are “superpartners” to the bosons and fermions predicted by the SM, which differ from their SM counterparts by a spin factor of $\frac{1}{2}\hbar$. An “s” is added to the names of the SM fermions (sup, selectron), and “ino” is added to the names of the SM bosons (wino, gluino), to denote the difference between SM particles and their superpartners [46]. The Minimally Supersymmetric Standard Model (MSSM) is the simplest SUSY theory and it predicts a single superpartner to each of the known SM fermions and bosons [47]. Supersymmetric SU(5) models have been popular for many years and even predicted the value of the Weinberg angle and that the top quark would have a mass of around 200 GeV [43]. Though the initial proton lifetimes of $\tau_p \simeq 10^{30-31}$ years have now been ruled out, it is possible to get proton lifetimes of $\tau_p \simeq 10^{34-36}$ years, which are well within the current experimental limits [48].

In the MSSM it is possible to have proton decay lifetimes of less than 1 s, through baryon and lepton number violating interactions, unless a symmetry called R -parity [49] is introduced (Equation 1.21).

$$P_R = (-1)^{3(B-L)+2s} \quad (1.21)$$

where B is baryon number, L is lepton number, and s is spin. SM particles have R parity $+1$, whilst supersymmetric particles have R -parity -1 . The baryon – lepton number ($B - L$) is generally conserved in GUTs, as a result of R -parity. However, it can be necessary to break R -parity in order to explain neutrino mass and mixing [43]. This has the unfortunate consequence of making the lightest neutralino unstable, meaning that it cannot be the Lightest Supersymmetric Particle (LSP), as it would decay too quickly to explain dark matter. This then means that the best candidate for the LSP is the unstable gravitino [43]. However, after allowing for R -parity to be broken, additional channels of nucleon decay become possible, such as $n \rightarrow K^+ + l^-$ and $p \rightarrow K^+ + l^- + \pi^+$. Chapter 6 will concern simulations of the muon-induced background to the $n \rightarrow K^+ + e^-$ decay channel, this decay was first proposed by invoking the non-conservation of $B - L$ [50, 51]. As will be shown in Table 2.3, the current lifetime limit on this channel is 3.2×10^{31} years [52]. It can be shown that the decay rate of the $n \rightarrow K^+ + l^-$ channel may be an order of magnitude larger than that of the proton [43, 53]. All decay channels which do not non-serve $B - L$ currently have very low limits on their lifetimes [26] and so warrant further study, as nucleon decay signatures may potentially be

observed with relatively low lifetimes.

Finally, it is also possible to construct GUTs which use higher order gauge groups, such as the SO(10) group. When considering SO(10) theories, supersymmetric models are normally considered, as ordinary SO(10) models have failed to be able to predict fermion masses and mixings [43]. However, the supersymmetric models are able to accommodate right handed neutrinos, explain the disparity between the quark and lepton mixing angles at SM energies, and predict the branching ratios of proton decays. As such, predictions of the proton lifetimes of the order of 10^{33-36} years exist, with an upper limit of 10^{38} years in the $p \rightarrow K^+ \bar{\nu}$ decay channel [54].

Over the years there have been many experiments which have searched for the signatures of nucleon decay, these have included NUSEX [55], FREJUS [52], Soudan I/II [56], and IMB [57]. The current experimental limit on the proton lifetime is held by the SK experiment, which sets a lower limit on partial lifetime in the $p \rightarrow e^+ \pi^0$ decay mode of 1.6×10^{34} years [58]. A successor to SK, called Hyper-Kamiokande (HK), is planned for the mid 2020s [59] and will increase the proton decay lifetime sensitivity by around an order of magnitude. DUNE will also search for nucleon decays, and will have a sensitivity to a number of nucleon decay modes which competes with HK despite being much smaller, due to the increased resolution of LArTPCs as compared to water Cherenkov detectors. The sensitivity of DUNE to nucleon decay will be discussed in Section 2.3.2.

1.2.2 Backgrounds to nucleon decay

In order to observe such rare processes it is necessary to place experiments in environments which have as little background as possible. For this reason, experiments which search for proton decay are placed far below the Earth's surface to reduce the cosmic ray flux. Such searches would be impossible if they were attempted on the Earth's surface. At large depths the hadronic flux is totally suppressed, however cosmic ray muons can survive to depths of over a mile underground. These high energy cosmic muons are able to produce signal mimicking events in underground detectors. Therefore, to observe the signal from a nucleon decay, these events have to be identified and rejected.

In the discussion of backgrounds to nucleon decay, reference will be given to the $p \rightarrow K^+ \bar{\nu}$ decay channel, where the signal is a lone K^+ in the detector. The kaon is isolated as the neutrino will not interact before escaping the detector, and the recoiling nucleus would have

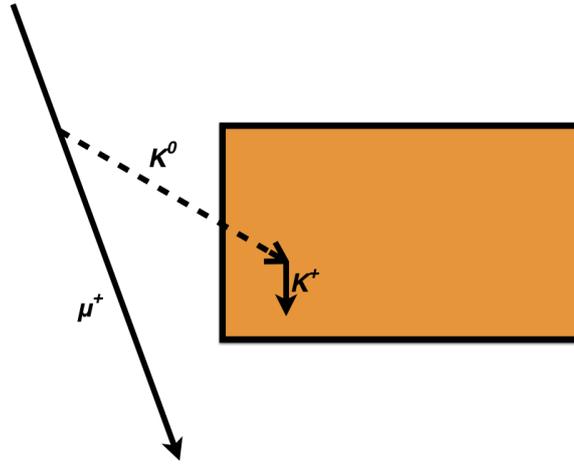


Fig. 1.2 Schematic of a muon event mimicking a nucleon decay. A cosmic muon produces a K_L^0 outside of the detector volume, which then interacts producing an isolated kaon. There would be no other charged tracks produced, and so the event would be indistinguishable from a real proton decay event.

too little energy to be detected. Though it is difficult to imagine a situation where a cosmic muon would directly produce an isolated kaon in the detector, one of its interaction products could do so. For instance, a muon could produce a K_L^0 outside of the detector, which then propagates into the detector, and undergoes charge exchange with a proton in the nucleus of an argon atom. This would produce an isolated kaon inside the TPC, with no associated tracks coming from outside the detector (the K_L^0 is neutral and so does not produce a track in LArTPCs). This event structure would look very similar to a nucleon decay event, and a schematic of such an interaction is shown in Figure 1.2. It is important to note that as LArTPCs are not magnetised it is very difficult to discern the charge of a kaon, and so a K^\pm are largely indistinguishable.

Atmospheric neutrinos are also able to produce signals which mimic nucleon decays if they interact within the active volume. The simulations described in Chapter 6 concern backgrounds induced by cosmic ray muons. The backgrounds induced by atmospheric neutrinos will not be considered in this thesis, as this background was previously shown to be negligible for the LBNE experiment [60]. This study showed that with truth cuts it was possible to reject all events from a sample of 10 million atmospheric neutrinos, representing 34.8 Mt yrs of LBNE detector live time. However, these studies will need to be repeated for DUNE in the future.

Though background inducing events like the one shown in Figure 1.2 are rare, the nucleon decay rate is very low, and so it is necessary to identify background events and reject them when searching for a signal. For this reason a fiducial cut is often applied, as many of the K_L^0 s will interact close to the detector edge. Should an interaction pass this cut, it is possible to apply strict energy criteria to the reconstructed energies. This is because the energies of the particles produced in a nucleon decay event are well defined. This will be discussed further in Chapter 6 in the context of the $n \rightarrow K^+ e^-$ decay channel. In the case of the $p \rightarrow K^+ \bar{\nu}$ decay channel, one would expect the kaon to have a momentum of about 340 MeV, and a total energy of about 600 MeV. The kaon would also be expected to decay at rest, and so the decay products from the kaon should have an energy equal to the rest mass of the kaon (493.677 MeV [26]). The derivation of how these energies are calculated will be shown in Section 6.3.3, with reference to the $n \rightarrow K^+ e^-$ decay channel.

Chapter 2

The Deep Underground Neutrino Experiment

2.1 Brief experimental overview

The Deep Underground Neutrino Experiment (DUNE) is a next-generation neutrino experiment, which came about through the recommendations of the P5 report in 2014 [61]. It was largely formed by the merger of two next-generation neutrino experimental proposals, the Long Baseline Neutrino Experiment (LBNE) [62–67] a US-based project, and the Long Baseline Neutrino Observatory (LBNO) [68] a European-based project.

The DUNE Far Detector (FD) will consist of four modules, each of which is a Liquid Argon (LAr) Time Projection Chamber (TPC) (LArTPC), with an approximate fiducial mass of 10 kton. The four FD modules will be 4850 ft below ground, at the Sanford Underground Research Facility (SURF). The first module will be a single phase (liquid) LArTPC and will be operational before the end of 2024. The subsequent modules will be a combination of single phase and two phase (liquid/gas) detectors, each becoming operational in 2025, 2027 and 2028. No details will be provided in this thesis concerning the two phase detector design, for details see [69]. A staged approach is used to ensure that physics measurements can begin before all of the modules are constructed.

The FD location at SURF is 1300 km away from Fermi National Laboratory (FNAL, Fermilab), from which DUNE will receive a neutrino beam, produced by a proton beam

with a power of 1.2 MW. The beam will become operational in 2026, meaning that two of the four FD modules will be constructed before the beam is finished. Therefore, the first two FD modules will have been fully characterised, and will have taken atmospheric data before the beam is operational. The proton beam power will increase to 2.4 MW during the lifetime of the experiment. The exact specifications of the beam design are currently being developed [70], though it will use the Main Injector beamline at FNAL after the PIP-II upgrade [71]. The neutrino beam will be able to be configured to run in either neutrino or antineutrino mode by “sign-selecting” the mesons which are produced after protons hit a graphite target. Assuming a beam exposure of 150 kt·MW·yr whilst running in neutrino mode, there will be over 800 (500) ν_e signal events in the FD assuming normal (inverted) mass hierarchy. When running in anti-neutrino mode, the neutrino rate will be over 200 (400) events in the FD. This compares to over 10,000 and 3,000 ν_μ signal events in the FD when running in neutrino or anti-neutrino mode respectively [72].

DUNE will also have a high resolution and high precision Near Detector (ND) located at Fermilab, approximately 600 m downstream from the target. The exact nature of the ND is still being decided, though the reference design is a Fine-Grained Tracker (FGT). This FGT consists of a Straw-Tube Tracking module, and an Electromagnetic Calorimeter, inside a 0.4 T dipole magnet. The DUNE ND will also have a set of muon identifiers in the dipole magnet. A detector of this structure will allow the DUNE ND to make precision measurements of the neutrino flux and interaction cross section, as well as the signal and background rates. There is also the possibility of building an additional LArTPC ND, though this would suffer from neutrino pile-up, due to the slow drift time of LArTPCs, and the high neutrino flux of the beam.

A schematic of the DUNE experimental setup is shown in Figure 2.1. Given this experimental setup, DUNE will have a wide range of physics opportunities, these are discussed in detail in Section 2.3. There is extensive documentation concerning all aspects of DUNE, this can be found in the Conceptual Design Report (CDR) documents [73, 72, 70, 69], and numerous annexes [74]. What follows is a high level summary of the information contained in these CDR volumes, and annexes.

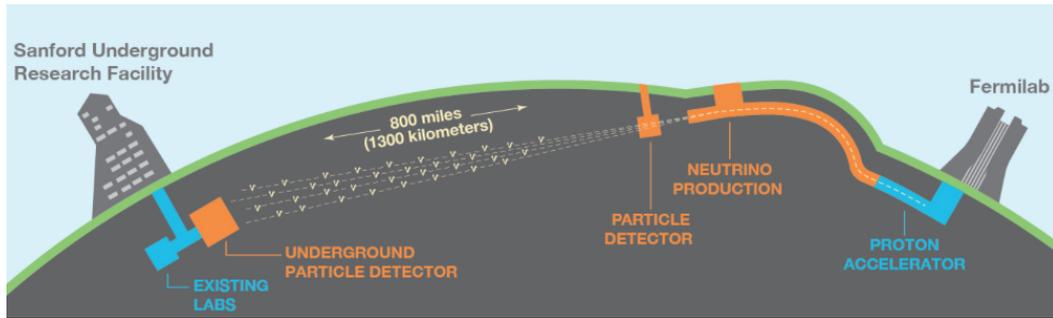


Fig. 2.1 A representation of the DUNE experimental setup. Fermilab, the host lab is shown on the right. The neutrino beam will be produced at Fermilab, and a high resolution near detector will also be located there. SURF is shown on the left, it is 1300 km away from Fermilab, and the far detector will be located there. Objects in orange need to be built during the lifetime of the experiment, whilst objects in blue already exist. The figure is not to scale, and is taken from [73].

2.2 The single phase detector design

As the work presented in this thesis has been performed on a prototype of the single phase design (Chapters 3, 4 and 5), and a simulation of the FD single phase design (Chapter 6), the DUNE FD single phase design will be briefly described below. As no work has been performed in relation to the two phase design, no information will be presented concerning this design. Detailed descriptions of both the single and two phase FD designs can be found in [69].

In the single phase design, the charge generation, electron drift, and charge collection, all occur in the LAr. This takes advantage of the low ionisation threshold for liquid Argon at only 23.6 ± 0.5 eV [75]. Upon ionisation a large fraction of the lowest energy electrons recombine with the ion produced in the interaction, however in higher energy ionisations additional electrons are produced by inelastic collisions. When electrons recombine they produce flashes of scintillation light with a mean wavelength of 128 nm, which can also be measured by photon detectors normally covered in wavelength shifting coatings. The strength of the applied electric field is chosen to optimise the amount of scintillation light produced, compared to the charge measured, as at very low fields little charge is collected; whilst at large electric fields very little scintillation light is produced. An electric field of 500 V cm^{-1} is normally chosen as a compromise between these two effects citepChepel:2012sj.

As the charge collection occurs in the LAr, this means that all of the TPC structure, including the electronics, are submerged. The active volume of 14.1 kt, measuring 12 m high, 14.5 m wide (in the drift direction), and 58 m long (in the beam direction), is too big to be

contained in a single TPC, and so many individual TPCs are required. This is because it is not practical to collect electrons after they drift over many metres, or to construct an Anode Plane Assembly (APA) frame and Cathode Plane Assembly (CPA) frame across 10 metres or more. The mine shafts would also not be able to accommodate APA and CPA frames which were over 10 metres long or wide. This limit on size also affects the material used to make the cryostat, and so it will be constructed using a stainless steel membrane design. This design is commonly used for liquefied natural gas storage and transport tanker ships, and so is an established technology. However, it had not been used in LArTPC experiments before the 35 ton prototype detector, described in Section 2.4.

The active volume of the DUNE FD is divided into 200 TPCs, each produced by applying voltages to APAs and CPAs, and contained within a field cage. Each TPC measures 6 m high, 3.6 m wide (in the drift direction), and 2.3 m long (along the beam direction). The TPCs are arranged such that there are 4 TPCs across the width of the detector, 2 TPCs stacked vertically on top of each other, and 25 TPCs along the length of the detector, to give the full active volume size of $12 \times 14.5 \times 58 \text{ m}^3$. This experimental setup means that the detector is 58 m long in the beam direction, offering a large region where neutrinos interaction can occur. The total drift distance of 14.5 m is split into four TPCs which are 3.6 m wide. Two sets of these four TPCs along the width of the beam are stacked on top of each other to produce the full height of the detector. A diagram showing this cross section of the single phase design is shown in Figure 2.2, where there would be an additional 24 cross sections going into the page, identical to the one shown.

As can be seen from Figure 2.2 there are 3 planes of APAs and 2 planes of CPAs, ordered such that there is a CPA plane in between any two APA planes, and there are APA planes on the outer faces of the detector. The separation of the APA and CPA planes is 3.6 m, producing the 3.6 m drift of the detector design. The ionised electrons drift towards the APA planes, in an electric field of 500 V cm^{-1} , produced by the voltage difference between the CPA and APA. In total there will be 150 APA frames, and 200 CPA frames installed in each single FD module.

Each APA frame is instrumented on both sides with 4 wire planes, the properties of which are outlined in Table 2.1. The first wire plane which a drifting electron would encounter is the grid plane (G), which acts to shield the other wire planes from distant moving charges and to make the field lines more stable near the wire planes. The grid plane wires are not read out. The next two wire planes are induction wire planes (U and V), and have a bias

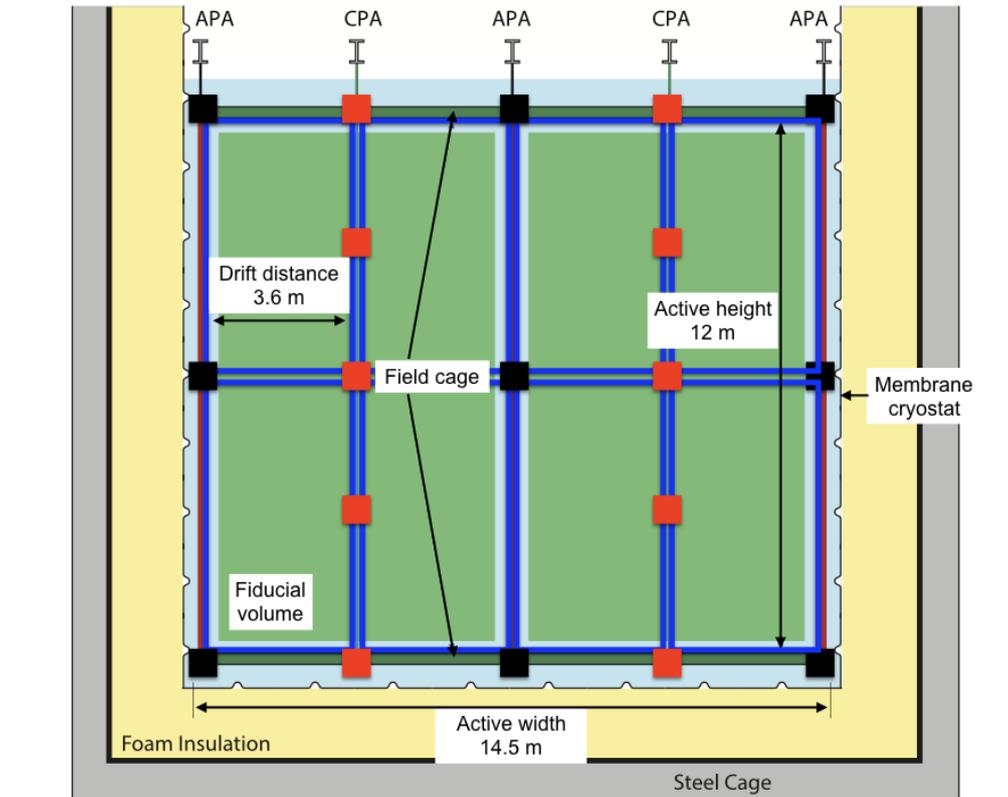


Fig. 2.2 A cross section of the DUNE signal phase detector design, showing 8 TPCs. The cross section shown is looking along the beamline, such that the detector extends a total of 58 m into the page. The TPCs (blue squares) are bound between a pair of APAs and CPAs, and are surrounded by a field cage (green bar). The TPCs are ordered such that there are two TPCs vertically, and four horizontally, such that there are a total of 8 TPCs shown. The planes of APAs are made from two APAs (vertical gaps between black squares), and the planes of CPAs are made from four CPAs (vertical gaps between red squares). There are a total of 3 planes of APAs, and 2 planes of CPAs, ordered from left to right || APA | CPA | APA | CPA | APA ||. The fiducial volume of LAr (green) is shown included into the full volume of LAr (light blue), which is bounded by a membrane cryostat (thin black), contained in foam insulation (yellow), which is further contained in a steel cage (thick black) and a concrete structure (grey). In the full detector there will be 24 more cross sections going into the page to produce a total of 200 TPCs. The figure has been modified from [69].

Table 2.1 The parameters of the wire planes in the DUNE FD single phase design.

Function	Orientation ($^{\circ}$)	Pitch (mm)	Num. Wires	Bias Voltage (V)
Grid (G)	0	4.79	960	-655
1 st induction (U)	35.7	4.67	800	-365
2 nd induction (V)	-35.7	4.67	800	0
Collection (Z)	0	4.79	960	860

voltage such that electrons will not be collected on them, but will instead induce charge as they drift past. Finally, the fourth wire plane consists of collection plane wires (Z), which record the charge deposited by ionisation electrons. A grounded mesh behind the collection wires prevents the electric field around these wires being distorted by the metal frame of the APA.

The induction plane wires are wrapped around the long edges of the APA frames. This means that on each APA frame there are two sets of grid and collection planes, one on each side of the APA, and two induction plane wires, both of which are sensitive to both sides of the APA. The wrapping of the induction plane wires means that readout electronics are only required at one of the short ends of the APAs. As the collection plane wires are only sensitive to particles in one TPC, they can resolve which TPC a measured charge on an induction plane wire was deposited in. An illustration of the wire wrapping on one of the DUNE FD APA frames is shown in Figure 2.3.

It is also envisioned that there will be a system of Photon Detectors (PDs), to measure interaction times of both beam, and non-beam, events. An individual PD will be comprised of a light guide, and 12 Silicon PhotoMultipliers (SiPMs). Each APA frame will contain 10 equally spaced PDs, which will be between the two planes of grounded mesh. The PDs will use a wavelength shifter on the surface of the light guide, to convert the 128 nm scintillation photons from the LAr, to photons with a wavelength of 430 nm. This wavelength shifted light will be collected by the SiPMs. The front end electronics for the SiPMs will reside outside of the cryostat, where a SiPM Signal Processor (SSP) digitises the signals from the SiPMs.

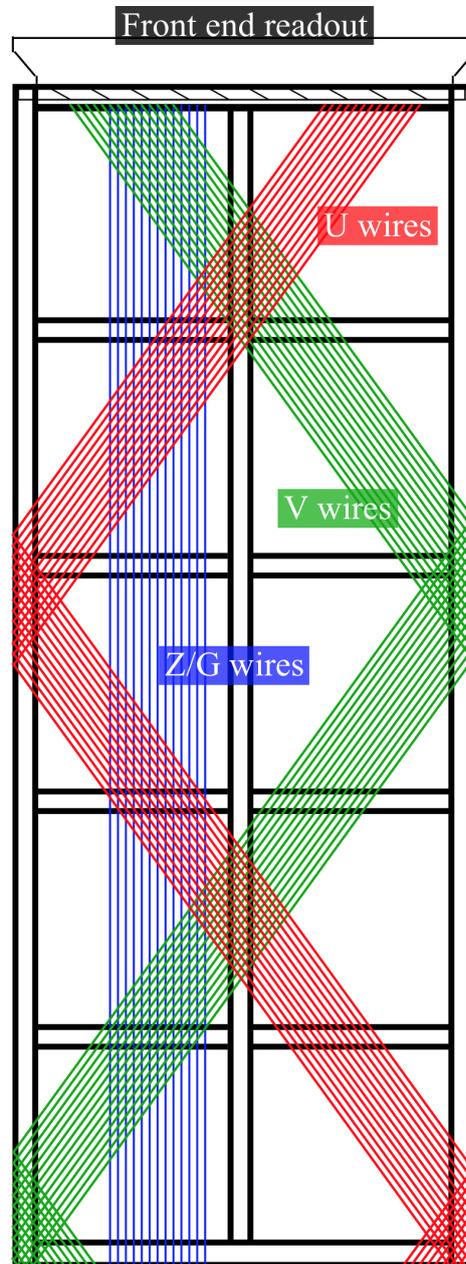


Fig. 2.3 An illustration of the wire wrapping in the DUNE single phase design. A small number of the wires in each plane are shown. The instrumented wire planes, in the order that the electrons drift past them, are shown as follows; the 1st induction plane (U) is red, the 2nd induction plane (V) is green, and the collection plane (Z) is blue. The grid plane (G), is in front of the 1st induction plane and is parallel to the collection plane. The figure is modified from [69].

2.3 The physics capabilities of DUNE

DUNE hopes to be able to deliver a wide ranging physics programme. These topics include, but are not limited to, precision measurements of neutrino oscillation physics, discussed in Section 2.3.1, and searching for nucleon decay in several important decay modes, discussed in Section 2.3.2. There are also additional physics goals, such as searches for supernova neutrinos, which are briefly outlined in Section 2.3.3.

2.3.1 Neutrino physics

The primary goals of DUNE concern neutrino physics, where many of these ideas were introduced in Section 1.1, and are [72]:

- Determine the neutrino mass hierarchy.
- Measure the charge-parity (CP) violating phase - δ_{CP} .
- Make precision measurements of the neutrino mixing parameters, such as θ_{13} , θ_{23} , and Δm_{31}^2 .

There are also secondary physics goals concerning neutrino physics, these include:

- Measuring neutrino oscillations using atmospheric neutrinos.
- Measuring a wide range of neutrino cross-sections, using the ND.
- Measuring nuclear effects, particularly neutrino final-state interactions, using the ND.

The beam exposure as a function of time which DUNE will receive is shown in Table 2.2, and is important when considering the figures presented below.

As presented in Section 1.1, both the matter effect, and δ_{CP} introduce an asymmetry between neutrino and anti-neutrino oscillations. As the matter effect is caused by the difference in the presence/absence of electrons/positrons in the Earth, it increases with distance. The result of this is that for baselines longer than approximately 1000 km the two effects can be resolved [77]. For this reason DUNE will have a baseline of 1300 km so that it is able to unambiguously determine the neutrino mass hierarchy, and determine δ_{CP} [78].

Figure 2.4 shows the oscillation probabilities of $\nu_{\mu} \rightarrow \nu_e$ for both neutrino and anti-neutrino modes at a baseline of 1300 km. The reason for requiring a broadband neutrino

Table 2.2 The beam exposure in units in units of kt·MW·years, as a function of time, assuming a staged DUNE construction. This table uses the fiducial mass of the detector (10 kt), the power of the initial proton beam (1.2-2.4 MW), and the time that the detector is exposed to the beam (years). The table is taken from [76].

Exposure (kt·MW·years)	Exposure (years)
171	5
300	7
556	10
984	15

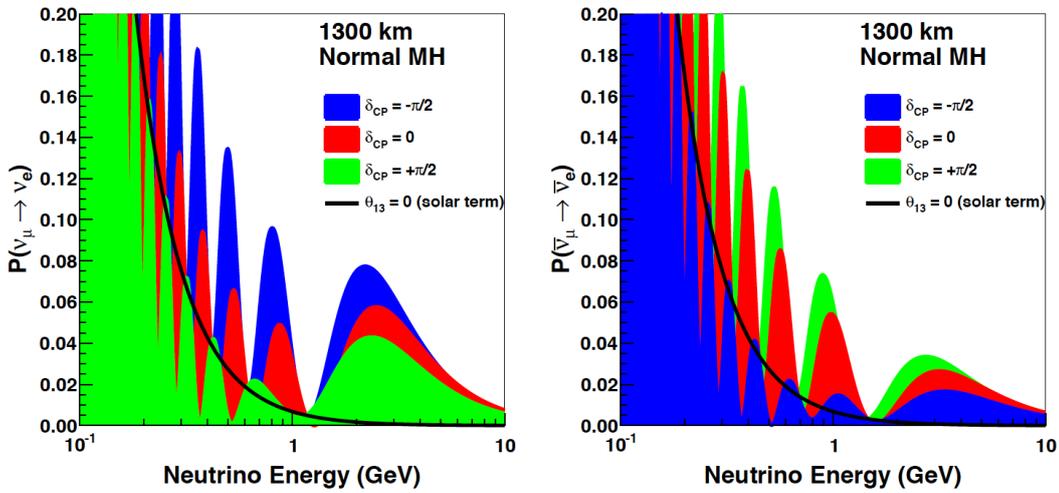


Fig. 2.4 The ν_e appearance probability at 1300 km as a function of neutrino energy, for a range of values of δ_{CP} . Left: the oscillation probability for neutrinos. Right: the oscillation probability for anti-neutrinos. Both figures assume normal mass hierarchy (ν_1 is the lightest state). The probabilities for different values of δ_{CP} are shown, $\delta_{CP} = -\pi/2$ (blue), $\delta_{CP} = 0$ (red), $\delta_{CP} = \pi/2$ (green). The figure is taken from [72].

beam at this baseline is clearly apparent, as it can be seen that whilst the energy of the first neutrino oscillation maximum is relatively unaffected by the value of δ_{CP} , the energies of the higher oscillation maxima are strongly affected. It is therefore vital that DUNE is able to accurately measure the rate of ν_e appearance at the lowest energies of the neutrino beam. It can also be seen that there are large differences in the expected oscillation probabilities for neutrinos and anti-neutrinos. Therefore, in order to measure the effect of δ_{CP} , the neutrino beam will have to be able to operate in both neutrino and anti-neutrino mode.

Figures 2.5 and 2.6 show the DUNE sensitivities to the mass hierarchy, and the value of δ_{CP} , with beam exposures corresponding to 7 and 10 years worth of data taking [76]. These

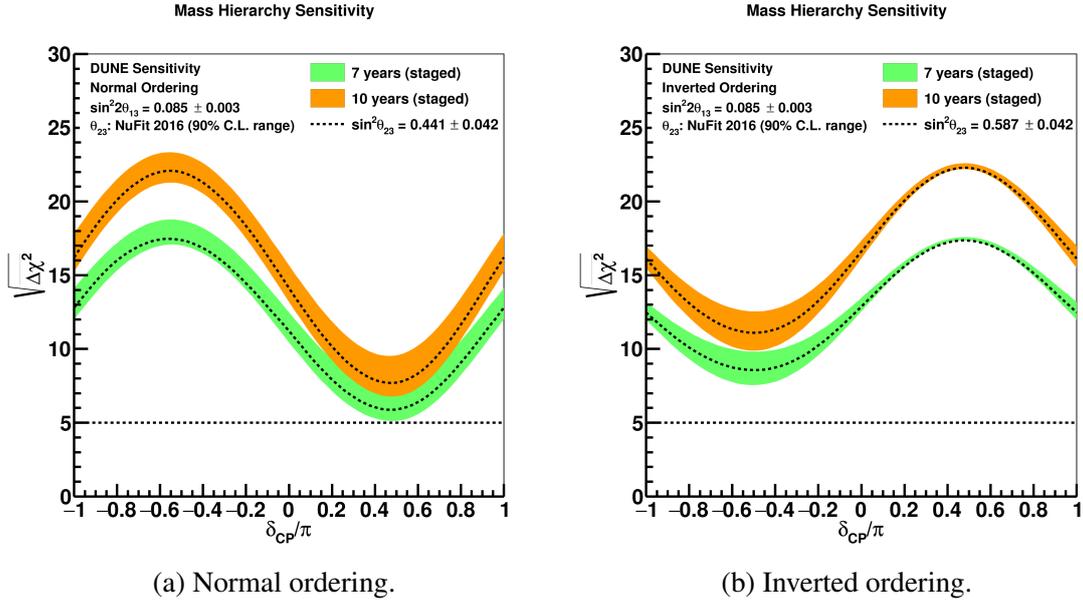


Fig. 2.5 The significance with which DUNE will be able to determine the neutrino mass hierarchy, for all values of δ_{CP} . Left: the sensitivity assuming normal ordering. Right: the sensitivity assuming inverted ordering. The shaded region shows the range of sensitivities for the 90% confidence level range for θ_{23} values, the dashed line shows the sensitivity for the NuFit central value of θ_{23} . The figure is taken from [79].

beam exposures are calculated given the phased approach to DUNE which was outlined in Table 2.2, meaning that a beam exposure of 7 years corresponds to a beam exposure 300 kt·MW·years, whilst a beam exposure of 10 years corresponds to a beam exposure of 556 kt·MW·years. The best fit values from NuFit 2016 [33] are used when making these figures, and equal running in neutrino and antineutrino mode is assumed. Figures 2.8, 2.9, 2.10 and 2.11 are taken from the CDR [72], and show how accurately given mixing parameters can be determined for increasing exposures.

Figure 2.5 shows the significance with which DUNE will be able to determine the neutrino mass hierarchy, for all values of δ_{CP} . It can be seen that the mass hierarchy can be determined with a significance of $\sqrt{\Delta\chi^2} = 5$, for all values of δ_{CP} after 7 years of data taking. It can also be seen that the mass hierarchy can be more conclusively determined if it is inverted.

Figure 2.6 shows the significance with which DUNE will be able to determine the value of δ_{CP} . It can be seen that even with 10 years worth of data, there are regions where the value of δ_{CP} cannot be determined accurately because of the complex phase being 0 when δ_{CP} is equal

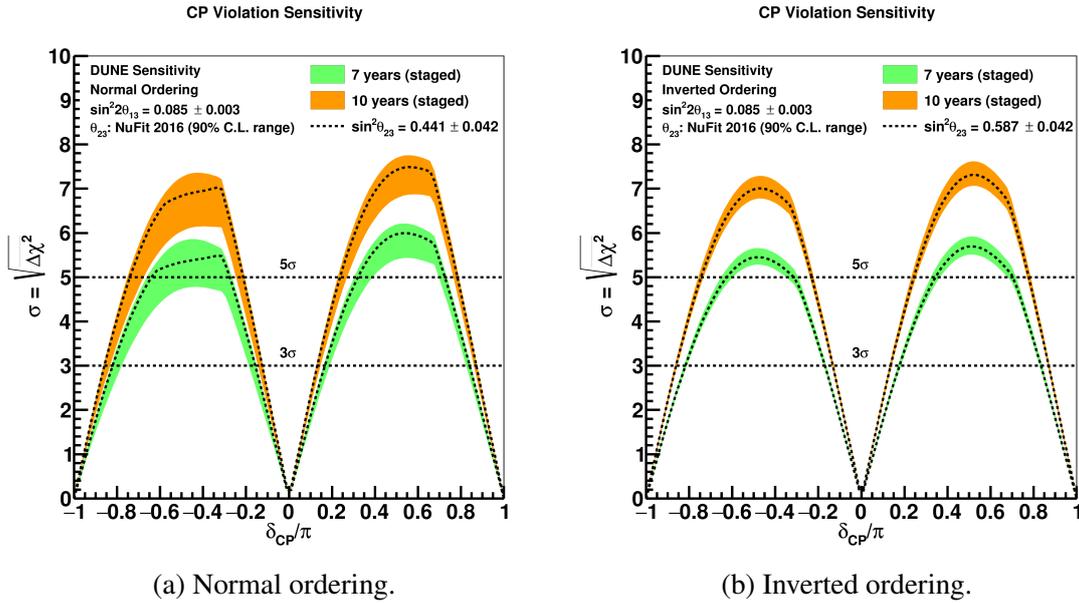


Fig. 2.6 The significance with which DUNE will be able to determine the value of δ_{CP} , for all values of δ_{CP} . Left: the sensitivity assuming the normal ordering. Right: the sensitivity assuming inverted inverted. The shaded region shows the range of sensitivities for the 90% confidence level range for θ_{23} values, the dashed line shows the sensitivity for the NuFit central value of θ_{23} . The figure is taken from [81].

to $-\pi$, 0, or π . Therefore, for values of δ_{CP} around these values, the significance to which δ_{CP} can be determined approaches 0. As such, even at very large beam exposures of over 800 kt·MW·years, corresponding to around 13 years of data taking according to Table 2.2, only 75% of the δ_{CP} values can be determined to a significance of 3σ [80]. However, even with a relatively modest exposure of 150 (550) kt·MW·years, DUNE can determine the value of δ_{CP} for over 50% of the range for δ_{CP} to a significance of 3σ (5σ) [80]. This shows that should the value of δ_{CP} be close to a CP-conserving case, it would be very difficult to determine it. However, if it is far away from these values, DUNE could make a measurement, with a significance of over 5σ , in a matter of years.

Figure 2.7 shows the resolution to which δ_{CP} can be determined when the value of δ_{CP} is both maximally CP-violating ($\delta_{CP} = 90^\circ$), and when it is CP-conserving ($\delta_{CP} = 0^\circ$). Somewhat paradoxically, the resolution is better when $\delta_{CP} = 0^\circ$. The reason for this, is that the region of values of δ_{CP} for which CP-violation would not be observed becomes increasingly small, as the beam exposure increases. This means that the resolution to which δ_{CP} can be determined increases. However, the more interesting result would be the one

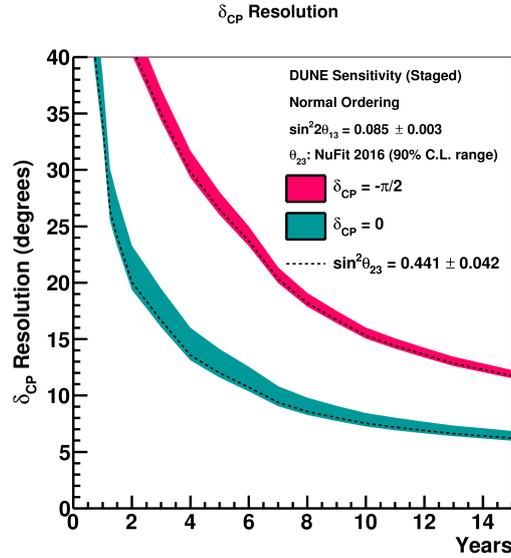


Fig. 2.7 The resolution with which DUNE will be able to determine the value of δ_{CP} , for increasing beam exposures (in years). Two bands are shown, one for a value of δ_{CP} which could cause maximal CP-violation ($\delta_{CP} = 90^\circ$), and another where there would be no CP-violation ($\delta_{CP} = 0^\circ$). A normal hierarchy is assumed, and the shaded region shows the range of sensitivities for the 90% confidence level range for θ_{23} values. The dashed line shows the sensitivity for the NuFit central value of θ_{23} . The figure is taken from [82].

which supported δ_{CP} having a value which causes maximal CP-violation.

DUNE also aims to perform precision measurements of the neutrino mixing parameters in order to improve sensitivity to any physics beyond the standard three-flavour oscillation model. As discussed in Section 1.1.3, the current best limit for the value of θ_{23} does not determine which octant it is in, specifically whether it is more than, or less than, 45° . Determining the octant of θ_{23} is important, as, should it be found that θ_{23} is equal to 45° , it would hint at an unknown symmetry. DUNE will determine the value of θ_{23} by combining measurements of $\nu_\mu \rightarrow \nu_\mu$ and $\nu_\mu \rightarrow \nu_e$, which are sensitive to $\sin^2 2\theta_{23}$ and $\sin^2 \theta_{23}$, respectively. Figure 2.8 shows the significance to which the octant of θ_{23} can be determined for different values of θ_{23} . Figure 2.9 shows the resolution to which the value of $\sin^2 \theta_{23}$ can be determined with increasing beam exposures.

The precision with which the values for $\sin^2 \theta_{13}$ and Δm_{31}^2 can be measured by DUNE, are shown in Figure 2.10 and Figure 2.11 respectively. The resolution to which DUNE can measure the value of $\sin^2 \theta_{13}$ is unlikely to surpass that of reactor experiments. However, DUNE will measure $\sin^2 \theta_{13}$ using ν_e and $\bar{\nu}_e$ appearance, as opposed to $\bar{\nu}_e$ disappearance, as

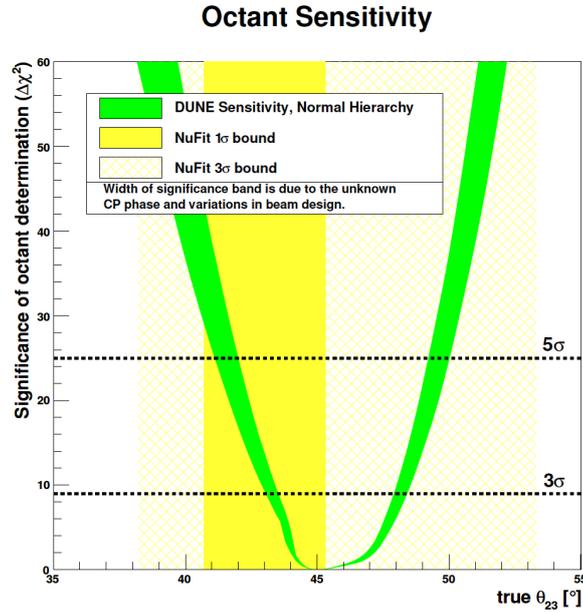


Fig. 2.8 The significance to which the octant of θ_{23} can be determined, for different values of θ_{23} . A beam exposure of 800 kt·MW·years is assumed, offering a 3 σ determination of the value of δ_{CP} for 75% of the values of δ_{CP} . The green band shows the effect of different potential beam designs, and variations in the true value of δ_{CP} [70]. The yellow regions show the 1 σ and 3 σ bands. The figure is taken from [72].

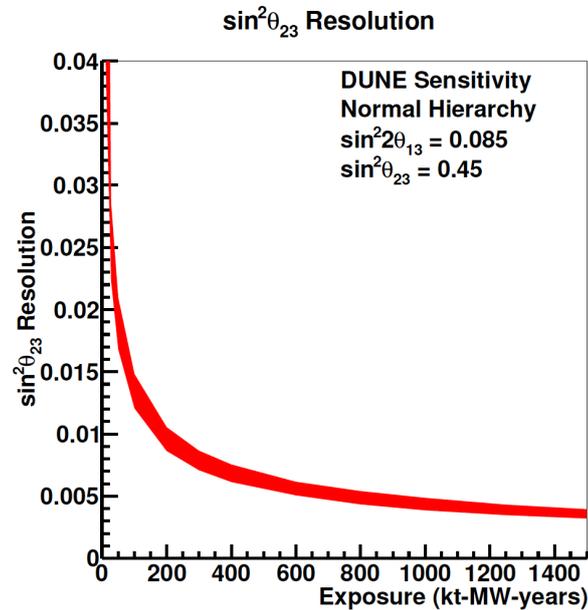


Fig. 2.9 The resolution with which DUNE will be able to determine the value of $\sin^2\theta_{23}$ for increasing beam exposures. A normal hierarchy is assumed, and the band shows the range of sensitivities given variations in the potential beam design [70]. The values for $\sin^2 2\theta_{13}$ and $\sin^2\theta_{23}$ are the current best fit values from [83]. The figure is taken from [72].

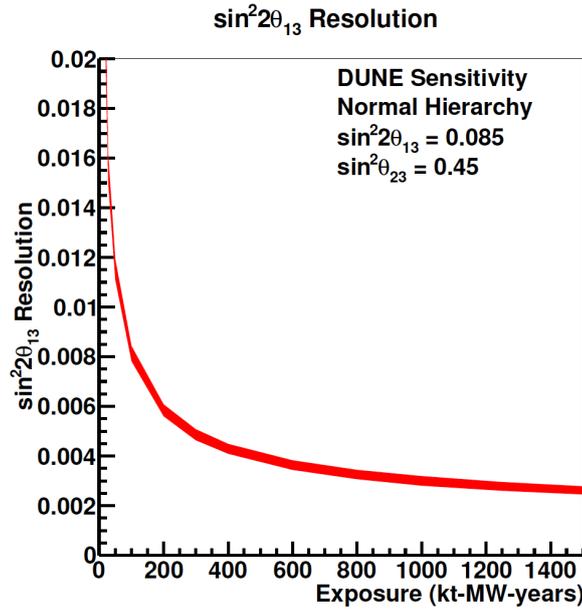


Fig. 2.10 The resolution with which DUNE will be able to determine the value of $\sin^2 \theta_{13}$ for increasing beam exposures. A normal hierarchy is assumed, and the band shows the range of sensitivities given variations in the potential beam design [70]. The values for $\sin^2 2\theta_{13}$ and $\sin^2 \theta_{23}$ are the current best fit values taken from [83]. The figure is taken from [72].

is done in reactor experiments. This complementary measurement of $\sin^2 \theta_{13}$ will provide an independent constraint on the three-flavour mixing matrix. DUNE will also be able to greatly improve the resolution to which the Δm_{31}^2 mass splitting can be determined [72].

It is also possible to measure many of the properties of neutrino mixing using atmospheric neutrinos. This is because atmospheric neutrinos contain all flavors of neutrinos and antineutrinos, and cover a wide range of L/E values. Also, atmospheric neutrinos are always available. This is particularly useful because, as discussed in Section 2.1, the DUNE schedule has two of the four FD modules becoming operational before the beam is ready. As is the case in experiments such as Super-Kamiokande, DUNE can observe the differences in upwards and downwards going neutrinos. The enhanced detector resolution of LArTPCs allows DUNE to have a comparable sensitivity to the mass hierarchy as the proposed Hyper-Kamiokande experiment, despite having a much smaller fiducial mass [72]. This can be seen in the plots shown in the most recent proposal for HK which involves placing a second detector in South Korea to take advantage of additional matter effects [84]. From this paper it can be seen that after 20 years of running HK will have almost identical sensitivities to the resolution of δ_{CP} and the neutrino mass hierarchy as the figures shown above for DUNE.

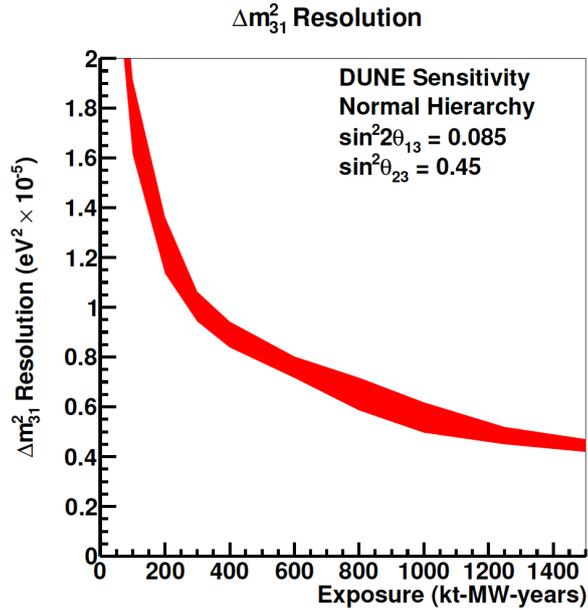


Fig. 2.11 The resolution with which DUNE will be able to determine the value of Δm_{31}^2 for increasing beam exposures. A normal hierarchy is assumed, and the band shows the range of sensitivities given variations in the potential beam design [70]. The values for $\sin^2 2\theta_{13}$ and $\sin^2 \theta_{23}$ are the current best fit values taken from [83]. The figure is taken from [72].

2.3.2 Nucleon decay

As presented in Section 1.2, many so called Grand Unified Theories (GUTs) predict some form of nucleon decay. Though nucleon decay has never been observed, it has not been completely ruled out. Should a nucleon decay event be observed; it would appear as a roughly 1 GeV energy deposition in the detector which would ideally be isolated and fully contained. In reality, the observed energy would be less than this, as significant amounts of energy can be lost due to the production of neutral particles such as neutrinos which would not deposit energy in the detector. As DUNE will be located deep underground, it will have a low background rate due to cosmic rays. This, combined with the high resolution of a LArTPC, means that DUNE offers a good opportunity to continue the search for nucleon decay.

Nucleon decay has been searched for by studying many decay channels, the current longest partial lifetime is in the $p \rightarrow e^+ \pi^0$ decay channel. In this decay, the total mass of the proton should be converted into the electromagnetic showers produced by the two particles, and their net momentum should be zero. This is a signal which can be clearly identified in water Cherenkov detectors, such as Super-Kamiokande (SK), and so is the main decay

mode which these detectors look for. This decay mode should also produce a clear signal in a LArTPC such as DUNE, though the limit that DUNE will place on this decay mode will not be competitive with that of Hyper-Kamiokande (HK), the successor to SK, due to the much smaller mass of DUNE. The $p \rightarrow K^+ \bar{\nu}$ decay mode is particularly interesting in DUNE, as kaons should be able to be accurately identified in a LArTPC. This is not the case in a water Cherenkov detector, as the kaons are not energetic enough to produce Cherenkov light. This is true for all decay modes which have a kaon in the final state, including neutron decay modes. It is also important to note that DUNE will search for all types of baryon number non-conservation, including bound neutron decays, di-nucleon decay modes and neutron-antineutron oscillations. An analysis concerning the $n \rightarrow K^+ + e^-$ decay channel will be presented in Section 6.3.

It is hoped that DUNE will be able to reach sensitivities to nucleon decay lifetimes of between $10^{33} - 10^{35}$ years. Figure 2.12 shows a comparison of current and potential lifetime limits for some decay modes, from a range of experiments. It can be seen that DUNE will provide very stringent limits to nucleon decay lifetimes, which will compete with HK in all decay modes containing kaons in the final state. Cherenkov detectors are able to perform relatively background free studies in the $p \rightarrow e^+ \pi^0$ channel and, due to its much larger mass, HK will be able to achieve a limit which is superior to DUNE. In the other channels however, the excellent spatial resolution of LArTPCs allows DUNE to compete with, and in many cases improve on, any limit which could be set by HK.

When estimating the expected sensitivity of DUNE to a range of nucleon decay modes, previous studies have been used [87, 88]. Some of the channels where one would expect a LArTPC, such as DUNE, to have an advantage in signal efficiency when compared to a water Cherenkov detector, such as SK/HK, are shown in Table 2.3. The accurate tracking of the kaon, and its subsequent decay products, is the reason for the increased signal efficiency that is expected in LArTPCs. The ability for LArTPCs to perform tracking to this accuracy was seen by the ICARUS collaboration, using the T600 detector [89]. Figure 2.13 shows an example event where a kaon enters the detector and is seen to decay to a muon, which subsequently decays to an electron.

Preliminary studies, at the time of writing the DUNE CDR documents, showed that after considering the backgrounds due to cosmic rays at depth, the backgrounds due to atmospheric neutrinos, and the impact of reconstruction failures, the number of background events for the $p \rightarrow K^+ \bar{\nu}_e$ decay mode should be less than $1 \text{ event} \cdot \text{Mt}^{-1} \cdot \text{year}$ [88, 91, 60].

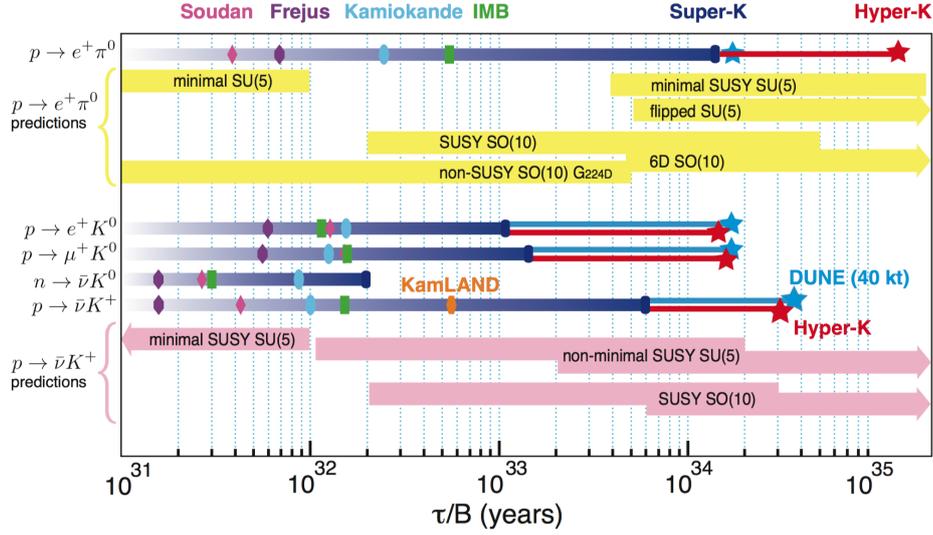


Fig. 2.12 A comparison of current, and future, nucleon decay lifetime limits, compared with the ranges predicted by Grand Unified Theories. Coloured bars are shown for published limits by a number of experiments [85, 86]. Stars are shown for projected limits by future experiments, these limits are calculated using Poisson statistics, and include predicted background rates. The lifetimes predicted by different models are shown for the $p \rightarrow e^+ \pi^0$ and $p \rightarrow K^+ \bar{\nu}$, decay modes. The figure is taken from [72].

Table 2.3 Nucleon decay limits in DUNE and Super-Kamiokande, in some favoured decay channels. The currently measured lifetime limits (Curr. limit), the estimated DUNE reconstruction efficiencies (DUNE ϵ), and the estimated DUNE lifetime limit in 2034 (DUNE limit) are shown. For comparison, the published Super-Kamiokande reconstruction efficiencies (SK ϵ), and the estimated Super-Kamiokande lifetime limit in 2034 (SK limit) are also shown. All lifetimes shown, are partial lifetimes, in units of 10^{33} years. The table is taken from [90], which uses [26].

Decay mode	Curr. limit	DUNE ϵ (%)	DUNE limit	SK ϵ (%)	SK limit
$p \rightarrow e^+ \pi^0$	16.7	45.3	21.4	54	50.8
$p \rightarrow \pi^+ \bar{\nu}_e$	0.016	41.9	19.8	54	N/A
$p \rightarrow K^+ \bar{\nu}_e$	0.051	97.0	29.7	11	2.1
$p \rightarrow \mu^+ \pi^0$	7.78	44.8	21.1	54	50.8
$p \rightarrow K^0 \mu^+$	1.6	46.7	22.0	100	22.2
$p \rightarrow K^+ \pi^+ e^-$	0.075	41.8	19.7	46	2.3
$n \rightarrow \pi^0 \bar{\nu}_e$	0.112	45.1	26.0	30	7.1
$n \rightarrow K^+ e^-$	0.032	96	55.4	100	1.8

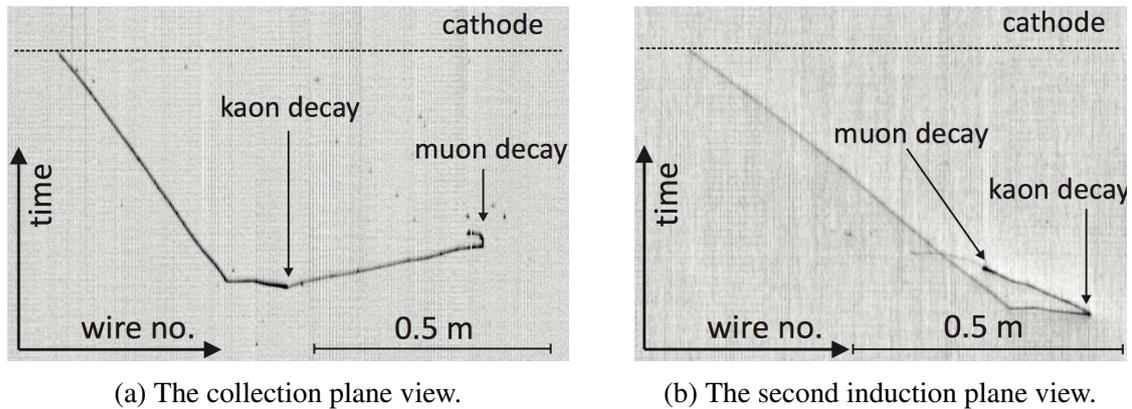


Fig. 2.13 A kaon event which was observed in the ICARUS T600 detector, in the CNGS data. Left: the signal on the collection plane. Right: the signal on the second induction plane. The kaon enters the detector and decays via $K \rightarrow \mu \nu_\mu$, the muon then decays via $\mu \rightarrow e \nu_e \nu_\mu$, where charges have been ignored. The figure is taken from [89].

With a background rate this low, the observation of a single, well reconstructed event, could provide evidence of nucleon decay [72].

2.3.3 Additional physics opportunities

Many of the largest detectors in the world would hope to observe neutrinos from a core-collapse supernova, should one occur during their lifetimes. DUNE is one of these experiments, and will have particularly good sensitivity to the electron flavour supernova neutrinos. This means that it should be possible to get a large, clean, supernova ν_e signal from DUNE, which is not possible using water Cherenkov detectors [92, 93]. Models predict that the full DUNE FD should observe about 3,000 events from a supernova at a distance of 10 kpc. The neutrino interactions from a supernova would consist of short electron tracks, which were potentially accompanied by a few gamma rays. The energy threshold for observation of these supernova neutrinos is still being studied, but it is thought that the threshold will be a few 10's of MeV [72]. The observation of a core-collapse supernova during the lifetime of DUNE could help answer some of the open questions which still remain after the observation of SN1987A [94, 95].

There are also other physics aspects which DUNE could make progress in, such as the indirect searches for WIMPs. This is because, should a flux of high energy neutrinos be observed to originate from the Sun, it could support the idea of dark matter annihilation. Searches for these interactions have been performed by IceCube [96, 97], but have not observed any signals. It is hoped that with the increased angular resolution of LArT-

PCs, the number of background events could be substantially reduced [72]. Other aspects where DUNE could make contributions are in searches for the diffuse supernova neutrino background [98], searches for neutrinos from accretion disks [99], and searches for black hole-neutron stars mergers [100].

Hence, there is a very compelling case for building a LArTPC with a fiducial mass of approximately 40 kt, at a large depth underground.

2.4 Path to building DUNE - The 35 ton prototype

The DUNE FD modules represent an increase in size by more than an order of magnitude, when compared to existing LArTPCs. It is therefore necessary to build a robust prototyping schedule, a brief outline of which is shown in Figure 2.14. It should be noted that DUNE will draw on the experience of all LArTPC experiments (MicroBooNE, SBND and many others), and so experiments not associated with DUNE will help shape the experiment through the improvements which they make to LArTPC technology.

The 35 ton detector is the first prototype under the umbrella of the DUNE experiment, for the single phase detector design. As such, the 35 ton detector provides a test bed for many of the novel aspects of the detector design required for the DUNE FD. This testing was done during two running periods, the Phase 1 run from November 2013 - February 2014, and the Phase 2 run from November 2015 - March 2016. Chapters 3, 4, and 5, will focus on some of the studies performed in preparation for, and during, the Phase 2 run.

The primary goal of the Phase 1 run was to verify that a detector with a membrane cryostat could achieve, and maintain, the high levels of LAr purity which are required to perform the physics which DUNE is expected to achieve. It was also designed to show that it was possible to achieve this high level of purity without the need to fully evacuate the detector prior to filling, but that this could instead be achieved by performing a “piston purge.” This “piston purge” had previously been demonstrated by the Liquid Argon Purity Demonstrator (LAPD) [102], which the 35 ton was built next to, at Fermilab’s PC-4 facility. The process by which the “piston purge” was performed is as follows. Firstly, room temperature argon is pumped into the detector, displacing the air in the detector. Once the purity is no longer seen to improve, a gas/liquid spray is used to slowly cool the detector. The injection of the spray introduces turbulence into the gas in the detector, which causes the entire cryostat to cool,

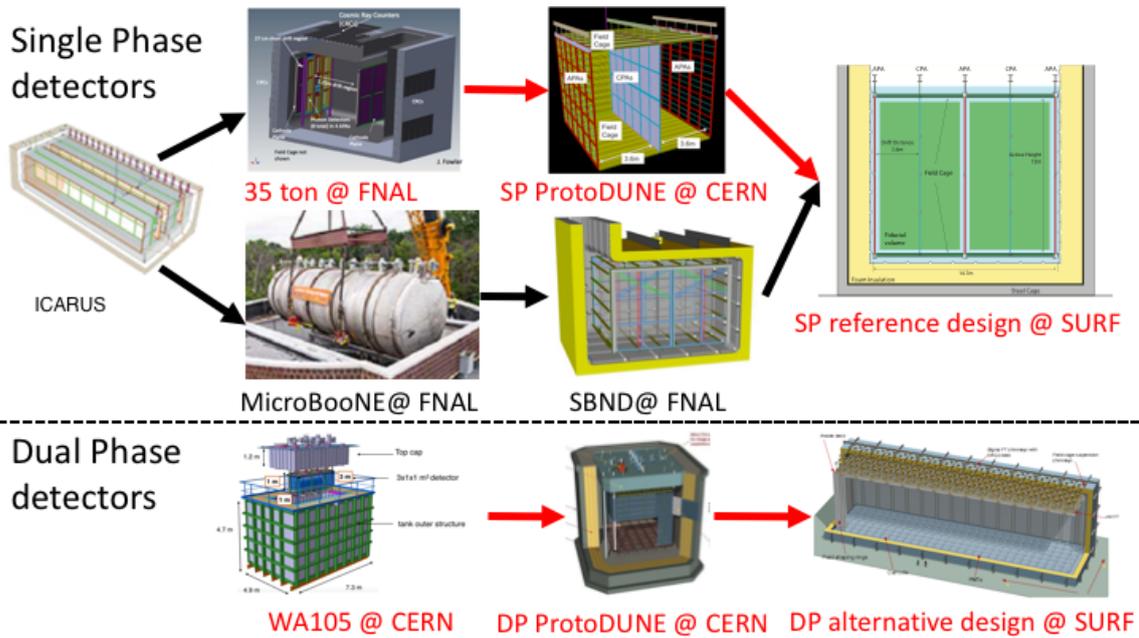


Fig. 2.14 An overview of the DUNE prototyping schedule, including complementary experiments. Top shows the experiments which utilise single phase LArTPC designs, whilst bottom shows experiments which utilise a two phase design. The experiments which are part of the DUNE prototyping program are shown in red. The figures are taken from [69, 101].

avoiding large temperature differences. Upon the completion of the cooldown, filling of the LAr commences, and LAr purification is performed using the installed recirculation loop. Using this purification method, electron lifetimes of 3 ms were observed after a few days of purification. These purity levels were maintained for many days at a time, thus demonstrating that membrane cryostat technology is capable of producing, and maintaining, high purity LAr [103, 104].

A detector similar to ProtoDUNE (outlined below), was originally planned to follow the Phase I run. However, funding constraints meant that this detector (at the time part of LBNE) got cancelled, and so the 35 ton cryostat was repurposed to contain a number of TPCs. This repurposed detector is the 35 ton Phase II run. The installed TPCs were designed to have many of the features which are present in the single phase reference design, these include:

- Modular APAs with wrapped wires, collecting charge from multiple drift volumes.
- Vertical and horizontal gaps between APAs. The Phase II run must show that it is possible to stitch particle tracks across, and through, APA frames.
- APAs and electronics which are immersed in the LAr.

- Waveguide-style photon detectors, installed inside the APA frames.
- A field cage built using printed circuit board.
- A DAQ which is capable of triggerless operation.

All of these features are central to the single phase DUNE detector design, and so demonstrating the successful operation of a detector with these properties is an important step in realising DUNE. Getting the Liquid Argon Software (LArSoft) framework (outlined in Section 2.5) ready for data taking required extensive simulation work, some of which is presented in Chapter 4.

In total four APAs were installed in the 35 ton cryostat, each collecting charge from two drift volumes, to give a total of 8 TPCs. As the total drift volume for the 35 ton was spatially limited to around 2.5 m, it was not possible to use TPCs with the drift length which will be present in the DUNE FD design. As such, it was decided that the TPCs would either have a “long” drift volume of 2.23 m, or a “short” drift volume of 0.27 m. These lengths were chosen so as to represent a reasonable drift distance in the “short” drift volume, whilst also maximising the “long” drift distance. Of the four APAs which were installed, two were “tall”, and were 120 cm in height, whilst the other two APAs were “short”, being 60 cm in height, and were stacked on top of each other, sandwiched between the two “tall” APAs. All of the APAs were 30 cm wide. This orientation allowed both vertical, and horizontal gaps to be produced, over which the reconstruction algorithms could attempt to stitch tracks. A total of 8 photon detectors were installed in the APAs, with 3 PDs in each of the “tall” APAs, and 1 PD in each of the “short” APAs. Figure 2.15 shows a schematic of the 35 ton detector.

A system of Cosmic Ray Counters (CRCs) can be seen on the outer edges of the cryostat walls in Figure 2.15. These were installed so that sets of cosmic muons could be recorded which were either parallel, or perpendicular, to the APA frames, as it was envisioned that these muons would be useful for later studies. There was also an additional set of CRCs on top of the cryostat to collect muons which were nearly vertical. The location of the 35 ton was not in a beamline, because, as discussed earlier, it was not originally intended to house TPCs. As such, only cosmic ray data was collected in the Phase II run, and so the muons identified by these CRCs produced a very valuable subset of the data, as will be seen in Chapter 5. The locations, and numbering scheme for the CRCs, is shown in Figure 2.16, and will be used extensively in Chapter 5.

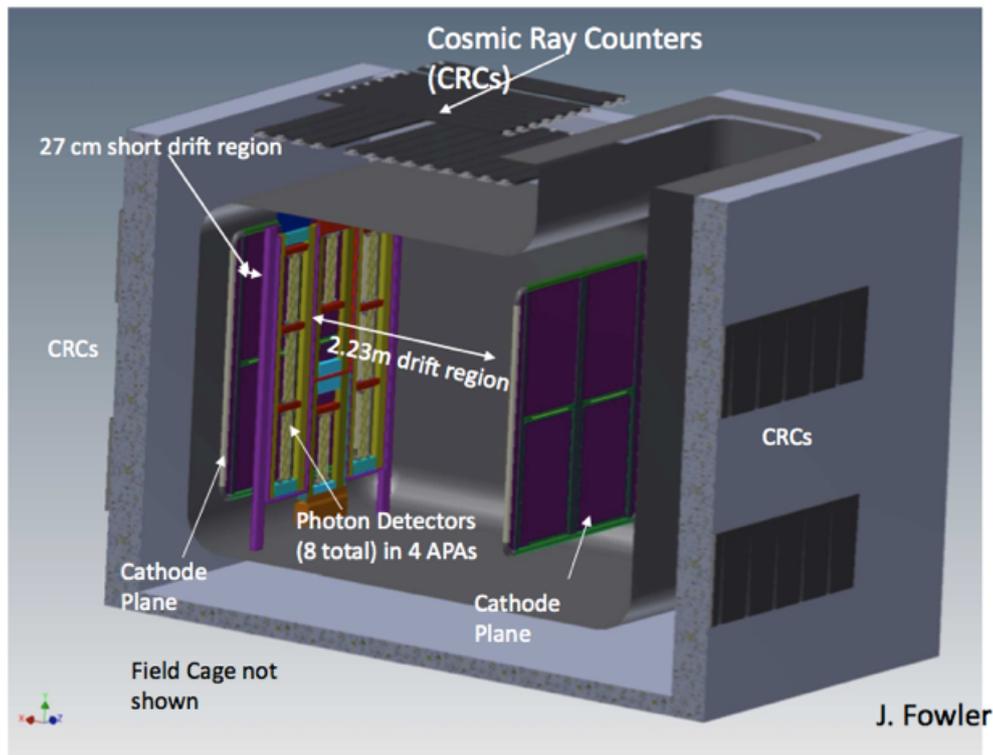


Fig. 2.15 A schematic representation of the 35 ton prototype detector during the Phase II run. The cathode planes (dark purple), photon detectors (yellow), APA frames (bounded light purple regions), cosmic rays counters (outer black rectangles), and the two drift volumes (2.23 m and 0.27 m) are shown. The photon detectors are in the middle of the APA frames, with the wire planes being on the outside of the APA frames. The 35 ton vessel (spotted grey), and steel container (black) are also shown. The field cage is not shown. Cryogenic piping for the detector is placed within the cryostat, to the right of the cathode plane associated with the “long” drift. The figure is taken from [69].

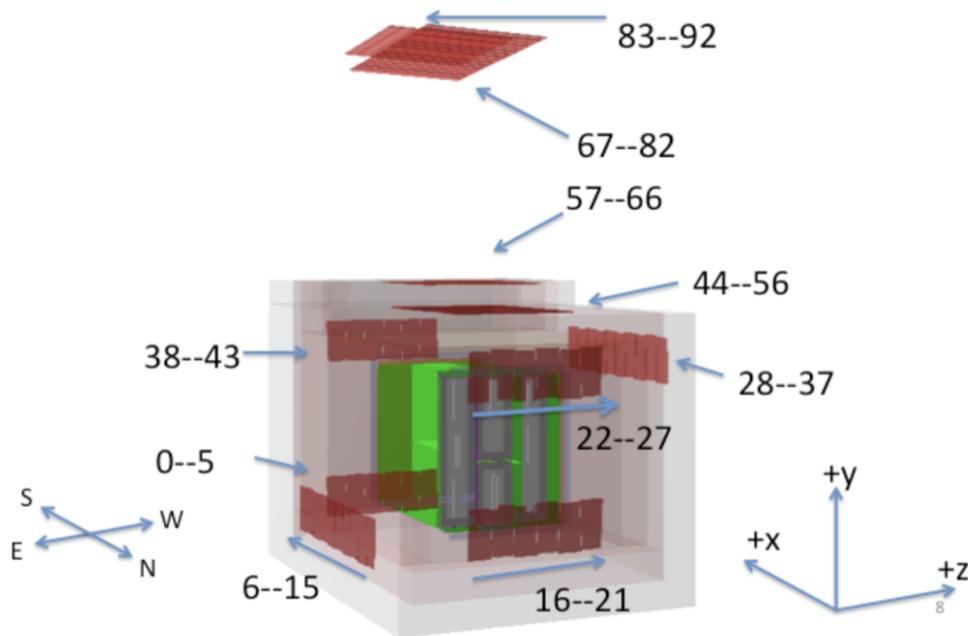


Fig. 2.16 A representation of the cosmic ray counter locations in the 35 ton, with the geographic, and LArSoft, coordinate systems shown. The other detector components can be seen inside the cryostat, such that the cosmic ray counters on the north wall are behind the short drift volume. The east - west counters are numbered 6-15 and 28-37 respectively. The north lower - south upper counters are numbered 16-21 and 38-43 respectively. The north upper - south lower counters are numbered 22-27 and 0-5 respectively. The “telescope” (vertically through-going) counters are numbered 44-92 and are split into four groups.

The Phase II run built on the experience gained during the Phase I run, using the same process for the initial cooldown, and LAr recirculation. The run also served to show that the electron lifetime which could be achieved was not limited by the presence of an instrumented detector. The operation of the installed TPC also served to test many of the novel features of the FD reference design. Some of the results from the Phase II run are shown in Chapter 5.

Following on from the 35 ton Phase II run, the single phase ProtoDUNE detector is under construction at CERN, and is due to take data in the second half of 2018. The ProtoDUNE detector is a small version of the full DUNE single phase reference design, as the APAs, and the drift length, are the same size as those in the reference design. The detector will be in a charged particle test beam, and so will be able to fully test the assumptions made in the DUNE CDRs about the detector performance characteristics. The detector will contain two sets of three APAs, 7.2 m apart, with one set of CPAs between them. This will give a total of 6 TPCs, each with a drift length of 3.6 m.

There is also a series of prototypes for the two phase design. The WA105 project involves both a small scale demonstrator, with an active mass of $1 \times 3 \times 3 \text{ m}^3$, which ran from late 2016 to early 2017, and a larger detector measuring $6 \times 6 \times 6 \text{ m}^3$. As was the case with the single phase prototypes, the demonstrator is not in a beam and so has only taken cosmic data, whilst the larger detector is in a charged particle test beam at CERN. The larger detector will serve as a full scale demonstration of the two phase detector design, though the drift distance is half of that in the final DUNE FD. Data-taking for this detector will be during the second half of 2018, at roughly the same time as the single phase ProtoDUNE detector.

2.5 The DUNE software

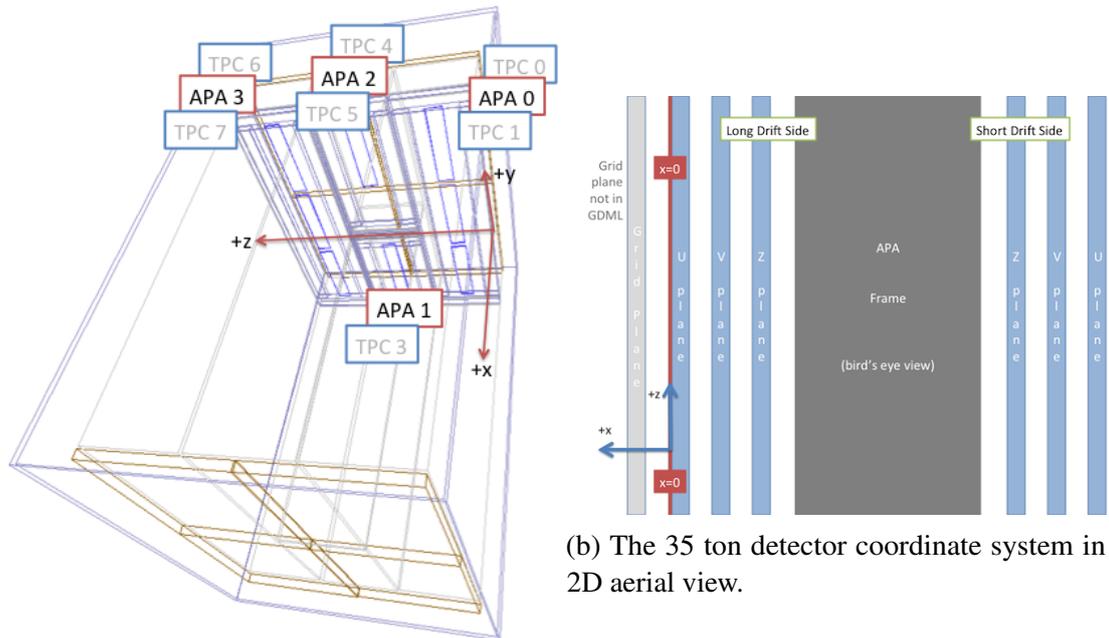
The software package used by DUNE is called LArSoft [105, 106]. LArSoft is a simulation, reconstruction and analysis package for LArTPCs that is being used by many experiments in the US neutrino program. It has been developed to be detector agnostic, meaning that much of the code is shared between experiments. To this end, it is envisioned that it will be used as a platform for constant development in existing experiments, and those still in the planning phases, such as DUNE. LArSoft is built around the Fermilab-supported *analysis reconstruction framework (art)*. External packages such as ROOT [107] and GEANT4 [108] are incorporated into LArSoft, meaning that the user does not have to coordinate specific

versions of the packages.

There are numerous mechanisms by which particles can be generated, as many external packages have been incorporated into the software. One such package is Generates Events for Neutrino Interaction Experiments (GENIE) [109], which is used to study neutrino interactions and nucleon decays. Another package, Nuance [110], is a neutrino interaction generator specifically for LAr. Finally, Cosmic RaY shower library (CRY) [111, 112], and COsmic Ray SIMulations for KAscade (CORSIKA) [113], are cosmic ray events generators, which are used to simulate the expected event rates for surface detector locations, in the absence of a neutrino beam. A muon generator using a Gaisser's parameterisation [114] has been incorporated for use by surface detectors [115]. Recently the MUon Simulations UNderground (MUSUN) [116, 117] generator, which takes the output of MUon SIMulation Code (MUSIC) [116, 118, 119], has also been incorporated, see Section 6.2 for further details. It is also possible to use a single particle generator, where the particle type, initial momenta, initial positions, and initial directions, can all be set by the user.

The coordinates and angles in LArSoft are defined as follows;

- x - The drift direction, which is normally perpendicular to the beam direction.
 - In the 35 ton prototype where there is no beam, positive x is in the opposite direction to that which electrons drift in the large TPC, where $x = 0$ is the position of the APA frames in the long drift volume.
 - In the far detector geometry $x = 0$ is defined as the midpoint across the full 14.5 m active width.
- y - The vertical direction, with maximal y being the highest point.
 - In the 35 ton $y = 0$ is halfway between the gap created by the two centre APAs, which are mounted one above the other.
 - In the far detector $y = 0$ is defined as the midpoint between the two vertical layers of TPCs.
- z - Defined so as to have a right handed co-ordinate system.
 - In both the 35 ton, and the far detector geometries, $z = 0$ is at the edge of the leftmost APA frame (when looking down the detector to maximal x position).
- θ - The angle that a vector makes from the x axis in the xy plane.



(a) The 35 ton detector coordinate system in 3D.

(b) The 35 ton detector coordinate system in a 2D aerial view.

Fig. 2.17 The LArSoft coordinate system as it is represented in the 35 ton detector in 3D (left) and 2D (right). The location of the origin is shown relative to the TPC detector components. Left: the four APAs (purple outlines), and eight TPCs are shown, where the even numbered TPCs are on the “short” drift side, 27 cm drift, and the odd numbered TPCs are on the “long” drift side, 223 cm drift. The CPAs (light brown outline) are also shown. Right: the location of the origin with respect to the wire planes which are wrapped around the APAs. It can be seen that $x = 0$ is defined as the location of the U plane in the “long” drift volume. The figure is taken from [120].

- ϕ - The angle between the z axis and the vector.

Figure 2.17 shows two schematic representations of how the location of the origin appears in the 35 ton prototype.

The simulation of particles is usually split into five processes, to reflect the different areas in which development often progresses. These stages are as follows;

- Particle generator.
- Particle transport using GEANT4.
- Full detector simulation, including detector responses.
- Full event reconstruction.

- Analysis.

The advantage of separating the computational processes in this way, is that improvements can be easily applied to a file without rerunning the entire simulation/reconstruction chain. This is especially important when large Monte Carlo, or data, samples are produced for general use within the collaboration, because it allows users to concentrate on improving a specific part of the computational process. A very general analysis is performed on these all-purpose samples, which provides users with all of the Monte Carlo truth information, along with all of the reconstructed quantities. This general analysis produces a file which can be used outside of the LArSoft framework, so that users can perform analyses independently of LArSoft, should they wish to.

Significant focus will be given to the reconstruction of TPC data in later chapters, and so it is necessary to briefly illustrate the mechanisms by which TPC data is reconstructed in LArSoft. Much of the information presented below is summarised in [106, 121].

The data collected by an experiment will have detector effects such as, an electronics response function, and the digitisation of signals. The full detector simulation also introduces these detector effects into simulated data, so that the reconstruction process does not treat simulated and recorded data differently. Therefore, the first step of the reconstruction algorithms is to remove these detector effects. Once these effects are removed, the signal is estimated using the value of *signal/noise* which would produce the measured signal. This process, called deconvolution, does not conserve pulse height, and is not guaranteed to preserve the normalisation. The deconvoluted signals are all unipolar distributions, which means that Gaussian distributions can be fitted to them, when trying to reconstruct hits. This is shown in Figure 2.18, and explained further below.

The deconvoluted signals are reconstructed into hits by identifying regions that are above a threshold value, and then attempting to replicate the signal in these regions by fitting to Gaussian distributions. For isolated hits, this is typically achieved using only one Gaussian distribution, however, for large energy depositions, over a large period time, where many particles are involved, multiple Gaussian distributions are often required. Large energy depositions are also possible when the direction of the particle aligns with the inclination of a wire plane, this means that all of the deposited energy may be deposited on a single wire. Examples of reconstructed hits are shown in Figure 2.18. These figures are taken from separate events where CRY was used to generate the particles, and so do not correspond to a continuous simulated event. They have been selected only as a demonstration of the

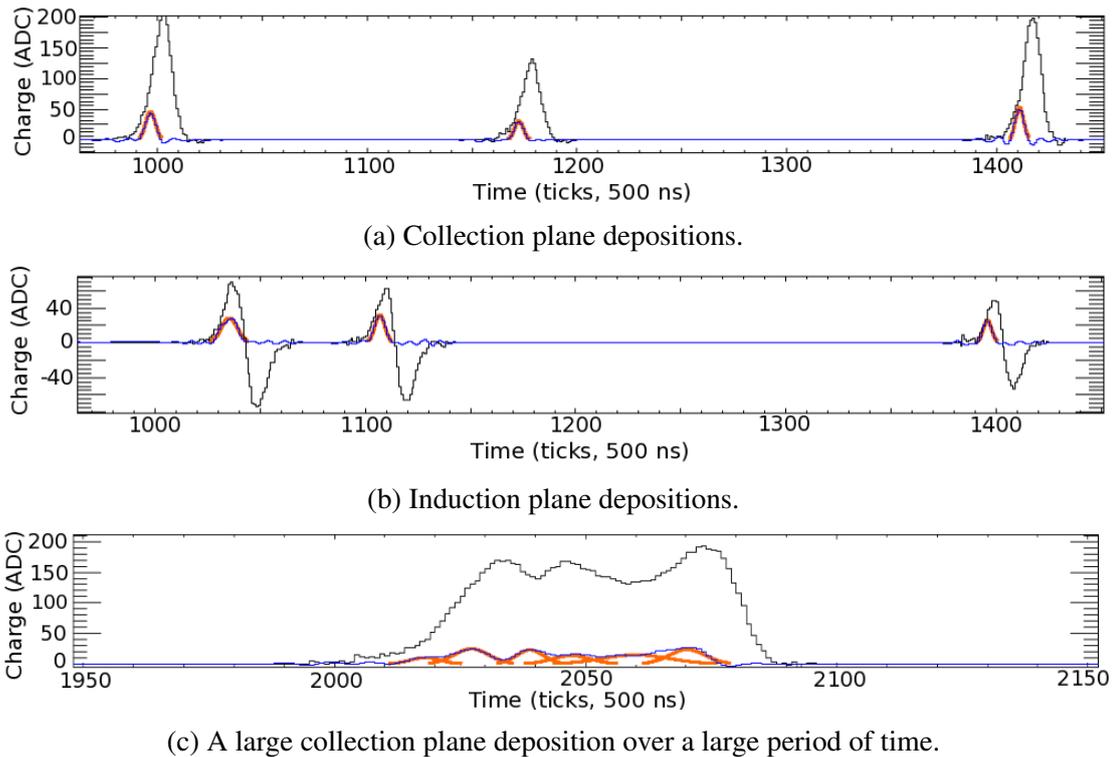


Fig. 2.18 The raw and deconvoluted signals with reconstructed hits for simulated energy depositions. The depositions are from particles generated by CRY and are not from a single event, as such they have been selected for demonstration purposes only. The plots are shown with increasing charge (ADC) on the y axis, and increasing time (ticks, 500 ns) on the x axis. The black lines represent the raw signals, the blue lines represent the deconvoluted signals, and the orange lines represent the reconstructed hits. Top: depositions on a collection plane wire, it can be seen that the raw signal is unipolar. Middle: depositions on an induction plane wire, it can be seen that the raw signal is bi-polar, whilst the deconvoluted signal, and reconstructed hits, are unipolar. Bottom: a complex deposition on a collection plane wire, where multiple reconstructed hits are required to reproduce the deconvoluted signal.

process of hit reconstruction. Figures 2.18a, and 2.18b, show multiple time-separated energy depositions on a collection plane wire, and an induction plane wire, respectively. A more complex energy deposition on a collection plane wire is shown in Figure 2.18c, where energy depositions from many particles, at similar times, have created a complicated energy deposition, which requires many reconstructed hits to explain.

As noted in Section 2.2, and Section 2.4, the DUNE FD, and the 35 ton detector, both have wrapped wires on the induction planes. Hence, the location of the reconstructed hit on an induction wire is ambiguous, as a single wire has many wire segments, as shown in Figure 2.3. An important feature of this ambiguity, is that the TPC in which an induction plane hit occurred cannot be identified, unless it is combined with another hit. These ambiguities do not extend to the collection plane wires, as they are not wrapped, and so consist of only a single wire segment, in a single TPC. Hits are combined across the three planes by identifying wire segments on each plane which intersect, and have hits at common times. In the traditional reconstruction process only hits that make these so-called “triple points” are considered disambiguated, with other hits being identified as noise hits, causing them to be discarded.

The inclination of the wire planes has to be carefully chosen so as to minimise both the number of wires required, and the number of times that wire triplets intersect. This is shown in Figure 2.19, where the wire inclinations used in the 35 ton detector are compared to those in the DUNE FD reference design. The inclination of induction plane wires in the 35 ton detector was $45^\circ \pm 0.7^\circ$, meaning that many wire triplets cross twice, and some wire pairs cross three times. The inclination of induction plane wires in the FD reference design is 36° , meaning that wire triplets only ever cross once. When wire triplets cross multiple times, the triplet which has the smallest distance between the common intersection point, and the two-wire intersection points, is chosen as the location of the hit. This is shown as the “Good intersection” on the right panel in Figure 2.19. The different wire pitches in the 35 ton detector were necessary so that one of the triple points could be evaluated to be the better candidate for the hit location, as with a wire pitch of 45° it can be impossible to distinguish between different triple points. The inclination of wires in the FD was chosen to be 36° to remove the possibility of multiple intersection points, as given the geometry of the APAs, multiple intersection points are impossible, and so disambiguation is much simpler. The lower inclination results in more induction wires being required though, making it more expensive to instrument the detector. It is also important that all wires on a given APA are either read out at the top, or base, of the APA. This is because there must be minimal space

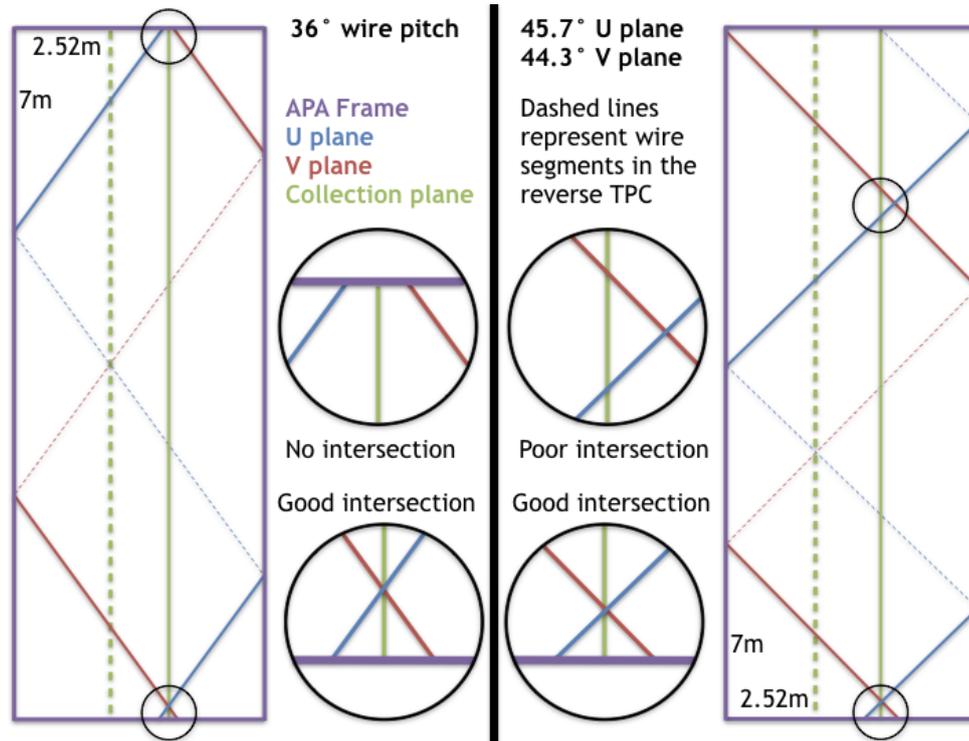


Fig. 2.19 The effect that different wire pitches have on the ability to perform disambiguation in APAs with the far detector geometry. Left: induction plane wires with a wire pitch of 36° , this is the wire pitch that is used in the FD reference design. Right: induction plane wires with a wire pitches of $45^\circ \pm 0.7^\circ$, this was the wire pitch that was used in the 35 ton detector. The left panel shows that only one “triple point” can be made with the three wires shown, and so disambiguation is trivial. The right panel shows that two “triple points” can be made with the three wires shown. The “triple point” where the three wires have a common intersection point is labelled as a “good intersection,” and it is this intersection point which would be chosen for the disambiguated hit.

between TPCs in the DUNE FD, in order to reduce the amount of internal dead space, and reading out APAs from the sides would introduce large regions of dead space. As a result of this, APAs at the top of the cryostat are read out from the top of the APA, whilst APAs at the bottom of the cryostat are read out from the bottom of the APA.

Once the hits have been disambiguated they are combined to make clusters in each of the three planes. The clustering process is usually performed in wire-tick space on each plane separately, where all the hits from a single track or shower, should make a single cluster in each plane. It is possible to seed the start of clusters by using imaging techniques such as a Harris transform [122], or to identify straight lines by using Hough transforms [123], though these are rarely used. As hits from a physical entity are unlikely to remain on a single

channel, or all come at identical times, clusters are often spread out over many channels, for a range of times, especially when performing clustering for showers.

Once clusters have been identified in each plane, they can then be merged into 3-dimensional tracks and showers. The two most common tracking algorithms are PMTrack [89] and Pandora [124], and the most common showering algorithm is EMShower [125]. As the tracking algorithms will be discussed in Chapters 4 and 5, a very brief overview of them will be given.

PMTrack is a multi-trajectory fit with some capabilities of pattern recognition, it iteratively builds its outputs by looking at the projection of multiple clusters in all projections, and builds 3D objects which best represent the 2D inputs. In doing this it is able to reconstruct both track and showers; and is also able to determine particle hierarchy. The theorem at the heart of PMTrack is to build and optimise 3D objects whilst minimising the distance between hits in 2D. [106, 126].

Pandora utilises a multi-algorithm approach to reconstruction to gradually build up a 3D picture of the event, whereby candidate vertices are identified and ranked from the initial 2D inputs. This technique was originally developed for the International Linear Collider but has since been developed to solve generic pattern recognition problems, and so is ideally suited to the events produced by LArTPCs. The number of individual algorithms contained in the Pandora package is very large, as whilst some are simple others are very complex. By combining the large algorithm base in multiple ways Pandora has multiple pathways for different types of event classifiers; such as “cosmics”, “tracks” and “showers” [106].

Once 3D objects have been reconstructed, the calorimetric quantities need to be determined, this is often done separately for each plane. Two models exist for calculating $\frac{dE}{dx}$ in LArSoft, Birks model [127] and a modified Box model [128]. The Box model is calculated under the assumption that during the recombination phase electrons and ions have a Gaussian distribution around the trajectory of the particle; and that the charge mobility is identical for electrons and ions. This model was originally developed to model the yield of scintillation light as a function of the stopping power of a material. When performing calculations on LAr it is assumed that the electrons are thermalised with the material. The Box model [129] applies boundary conditions to the charge distribution, and assumes that electron diffusion and ion mobility are negligible during recombination. It is found that neither Birks model or the Box model provide a coherent explanation of data in all energy ranges. This is because whilst Birks model is accurate at relatively low values of $\frac{dE}{dx}$, it is technically difficult to apply it for

large ionisations. In contrast, the Box model is not accurate at low values of $\frac{dE}{dx}$, but it does not have technical limitation at large ionisations. As a result it is necessary to modify one of these formalisms to provide a coherent picture of ionisation. This is normally done by modifying the parameters of the Box model at low values of $\frac{dE}{dx}$ to give rise to the so-called Modified Box model [128]. Therefore, the Modified Box model has been used in the simulations considered in this thesis, though the formalism for the Birks model is provided for illustration.

Both models incorporated in LArSoft calculate the $\frac{dE}{dx}$ of a hit using the deposited charge (dQ), and the track pitch (dx) of the hit, as well as the conversion of ADC value to number of electrons ($C_{ADC \rightarrow e^-}$), the conversion of GeV to number of electrons ($C_{GeV \rightarrow e^-}$), a correction due to the electron lifetime ($C_{lifetime}$), the LAr density (ρ), the electric field (E_{field}), and the tuneable electron recombination factors ($Recomb_{A/B}$). The series of equations used in Birks model are shown in Equation 2.1, whilst those used in the modified Box model are shown in Equation 2.2.

$$\frac{dE}{dx} = \frac{(dQ/dx)_{Cor}}{\alpha - [\beta \times (dQ/dx)_{Cor}]} \quad (2.1a)$$

$$(dQ/dx)_{Cor} = \frac{dQ}{dx} \times \frac{C_{lifetime}}{C_{ADC \rightarrow e^-}} \quad (2.1b)$$

$$\alpha = Recomb_A \times C_{GeV \rightarrow e^-} \times 10^{-3} \quad (2.1c)$$

$$\beta = \frac{Recomb_B}{\rho \times E_{field}} \quad (2.1d)$$

$$\frac{dE}{dx} = \frac{e^\alpha - Recomb_A}{\beta} \quad (2.2a)$$

$$\alpha = \frac{10^3 \times \beta}{C_{GeV \rightarrow e^-}} \times (dQ/dx)_{Cor} \quad (2.2b)$$

$$(dQ/dx)_{Cor} = \frac{dQ}{dx} \times \frac{C_{lifetime}}{C_{ADC \rightarrow e^-}} \quad (2.2c)$$

$$\beta = \frac{Recomb_B}{\rho \times E_{field}} \quad (2.2d)$$

When performing calorimetry, it is also important that the interaction time is known, so that the x positions of hits can be corrected, as they will initially be reconstructed assuming an interaction time of 0 s. This assumption is made because the beam trigger is placed at a time

of $T = 0$ when considering beam induced events. An unknown interaction time causes the hit and track positions to be calculated incorrectly. The correction to the reconstructed positions has to be made by hand, as the positions of the reconstructed hits cannot be modified in the event record.

Chapter 3

The 35 ton camera system

A camera system which was designed and built by the University of Sheffield, was installed in Run II of the 35 ton prototype. The reason for this was to monitor the cryostat for any potential high voltage (HV) breakdowns. The occurrence of a HV breakdown in liquid Argon (LAr) is possible due to the large electric fields which are required for experiments such as DUNE. For example, when the drift field of 500 V cm^{-1} will be applied to the DUNE far detector (FD), it will require a bias voltage of -180 kV [69] on the cathode plane. Therefore, there will be a very large electric field near the cathode, and as the breakdown field has recently been measured to be around 40 kV cm^{-1} [130], the cathode will require some form of monitoring.

To this end, a Complementary Metal-Oxide Semiconductor (CMOS) camera is used within the LAr to observe any potential HV breakdowns in the 35 ton cryostat. This has never been done before using cameras which are immersed in LAr. Previously, Charge-Coupled Device (CCD) or CMOS cameras have either used viewing ports which were built into the detector, or were inside enclosures which had a raised temperature [130, 131, 102, 132, 133]. As well as looking for HV breakdowns, the installed cameras also visually monitored the TPC and cryogenic components as filling occurred. The information contained in this section is a summary of the paper detailing the performance of the cameras in the 35 ton prototype [134]. The author was responsible for the fabrication and initial setup of all equipment which was used at Fermilab, including ensuring that the camera system met all safety standards. The author made only very minor contributions to the studies which were performed to select the cameras which were used, or in the design and construction of the camera mounting modules.

3.1 The selection and characterisation of cameras

As the cameras in the 35 ton are submerged in LAr, they are required to work at cryogenic temperatures. The cameras are also required to be sensitive enough to visible light that they can observe the sparks created by the HV breakdowns, this means that the cameras must have low thermal noise. CMOS cameras are used as there are many cameras which are rated to work down to temperatures of -40°C , however it should be noted that this is much hotter than the temperature of liquid Argon at 87K. CMOS cameras generally have multiple transistors per pixel meaning that their area for light collection is generally lower than for CCD cameras, though it does mean that each pixel can be read out individually. CMOS cameras are also found to have lower leakage currents, improved mobility, and lower thermal noise, at temperatures around 100K [135, 136].

When a shock test, consisting of submersion in LAr, is performed on a range of CMOS cameras, it is found that the most reliable camera is a *Floureon* car reversing camera. It is believed that the simplicity of this camera allows it to be relatively unaffected when operating at low temperatures. This is because it has a simple internal circuit consisting of resistors, capacitors, a crystal oscillator, and a flash memory chip.

It is important to determine the effect which operation at cryogenic temperatures has on the performance of the cameras. To do this, the frame rate (defined as how often the camera refreshes), and resolutions in both time and space (defined as the minimum separation in both time and space that a flashing LED can be resolved), whilst submerged in LAr, are compared with a complementary set of measurements made at room temperature in air. It is found that the frame rate, and resolution in time, are both unaffected at 50 Hz, and 20 ns, respectively. The timing resolution of the cameras was determined using LED's in a dark box, which were connected to a pulse generator that generated pulses of variable widths. It was found that in both air and LAr, the cameras were able to trigger on pulses once they were longer than 20 ns, due to the difference between successive frames. The spatial resolution, which was also found to be unaffected within systematic errors, was measured by varying the distance between two optical fibres, connected to a single LED, until the light from both optical fibres could no longer be resolved.

When attempting to use the cameras to locate HV breakdowns, it is necessary to develop a triggering mechanism. The camera signals are read out using a digital video recorder (DVR), which is remotely accessible using SwannView Link, a commercially available surveillance program. The SwannView Link software detects movement between successive frames, and

so in normal use will begin recording if, for example, a person enters the field of view. The length of time for which a video is recorded, and the number of frames which are recorded before the trigger, can be configured by the user. The monitoring for HV breakdowns is remarkably similar to this application, as when a breakdown occurs the field of view would be illuminated by a large flash of light, which was not present in the previous frame. When determining if a breakdown has occurred, a threshold in the number of pixels which change between consecutive frames is used. It is also possible to select only a limited number of pixels, if for example, some regions of the cameras field of view are rapidly changing, or if the expected change would only be in a very precise location. This was the case with some of the cameras in the 35 ton, where there was elevated thermal noise in some regions of the camera pictures, and any breakdowns would be highly localised.

The ability of the cameras to measure HV breakdowns, using the above triggering system, was tested prior to installation in the 35 ton. In the tests, a high voltage was applied across a printed circuit board until breakdown was observed. When breakdowns occurred, the triggers which the cameras recorded showed sparks which were highly localised and lasted over multiple frames.

3.2 The design of the camera system

As will be noted at the start of Chapter 5, the 35 ton detector was filled with LAr for over 3 months, and so the shock tests performed with liquid nitrogen were not sufficient to guarantee continuous running of the cameras during the entire 35 ton run. It is also necessary to protect against the prospect of a power failure, which could cause the cameras to be turned off for a large period of time. Should this occur, the cameras would have to be able to be turned on at cryogenic temperatures. During testing, it was found that whilst some of the *Floureon* cameras were able to do this, this was not the case for all cameras. Only cameras which were found to be able to be turned off and on again in the cold were used in the 35 ton. The process of turning a camera off and on again is referred to here as a camera being “power cycled.”

As the failure rate of the cameras was non-zero, and because the 35 ton run was scheduled to be much longer than any cameras had been tested in the cold, it was decided that it was prudent to build a self-contained module to house the cameras. A heater, consisting of two resistors placed either side of the camera, was installed inside the camera module. This is because it was found that some of the cameras which failed the shock tests, could be power

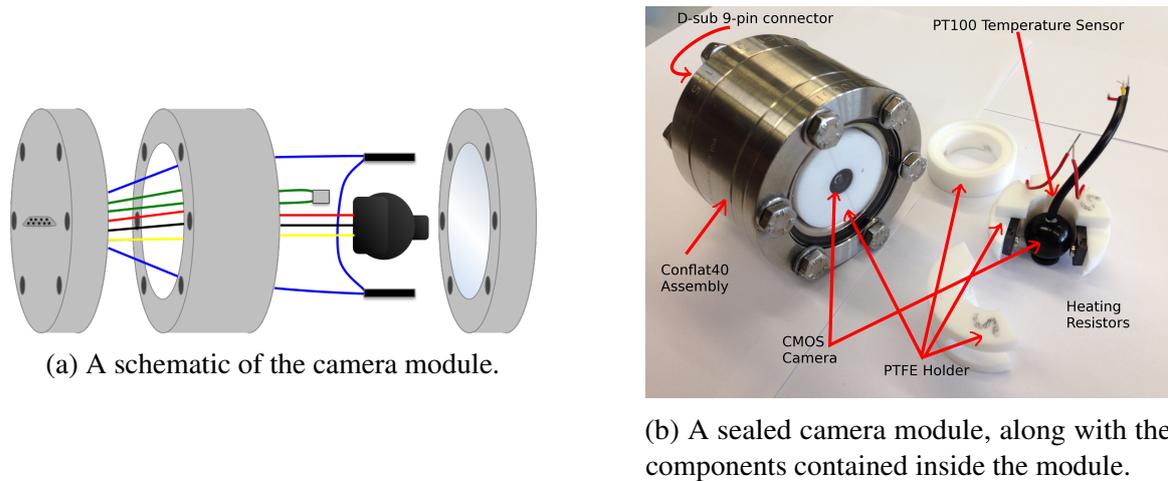


Fig. 3.1 The components which made up a camera module used in the 35 ton camera system. Left: a schematic representation of the components in the camera modules, with wires from the PT100 temperature sensor (green), the heating resistors (blue) and camera (red, blue, yellow), connected to a 9-pin D-sub feed through, shown on the left of the image. Right: the physical components, both inside and outside, a module. A camera, a heater and a temperature sensor, can all be seen. A PTFE holder is used to ensure that the components remain in the desired locations, with the camera pressed up to the glass viewing panel. Bolts are used to ensure that the camera modules are leak tight.

cycled when their internal temperature was raised by around 20 K. A temperature sensor was also placed in the modules, so that the increase in temperature could be measured during normal operation, and also during heating. It was found that the cameras caused the internal temperature of the module to rise by roughly 15 K, though the temperature of the glass viewing panel, and the external camera module, were unchanged. It was also found that with the heater operational, the internal temperature of the module could rise by as much as 80 K after 17 minutes. During operation, the heaters and power supply to the cameras were controlled using a custom-built, remotely controllable device, built at the University of Warwick. The temperature from the PT100 sensor was processed by another custom-made device, the output of which was read into a computer using a NI USB-6009 device.

A schematic of a camera module is shown in Figure 3.1a. Figure 3.1b shows a picture of a sealed camera module, with the individual components inside a module next to it.

It was necessary to mount the camera modules inside the 35 ton. This was achieved using a custom-designed mounting bracket, an example of which is shown in Figure 3.2. The mounts were fixed to existing cryogenic pipework, and could be freely rotated, so that the



Fig. 3.2 Two camera modules which are mounted in the 35 ton prototype detector. The mounts are affixed to a 3” cryogenic pipe. The multi-coloured ribbon cables which were used to carry the camera signals outside of the cryostat can be seen in the picture.

cameras could be pointed towards specific regions of interest. This meant that the camera mounts had two degrees of freedom, as well as being able to freely placed on the cryogenic pipes.

However, before the cameras could be installed in the 35 ton cryostat, several safety reviews had to be completed, in order to comply with safety standards set out by Fermilab, the host lab for the 35 ton prototype. These consisted of a review of the cable rack which was used during operations, plus individual reviews for the components which were custom-built. There was also a further review covering the camera system as a whole when it was operational. Only after the successful completion of all of these reviews was the camera system allowed to be installed and run unsupervised inside the 35 ton cryostat.

In order to satisfy the safety requirements adequate grounding of the rack, and all of its components, had to be displayed. This requirement extended to ensuring that the cables used in the system, which were assembled at Fermilab, would not introduce any potential ground loops into the 35 ton system. This meant that detailed diagrams of the individual components and how they linked together to form the camera system had to be produced. The detailed diagram which was produced representing the entire camera system is shown in Figure 3.3.

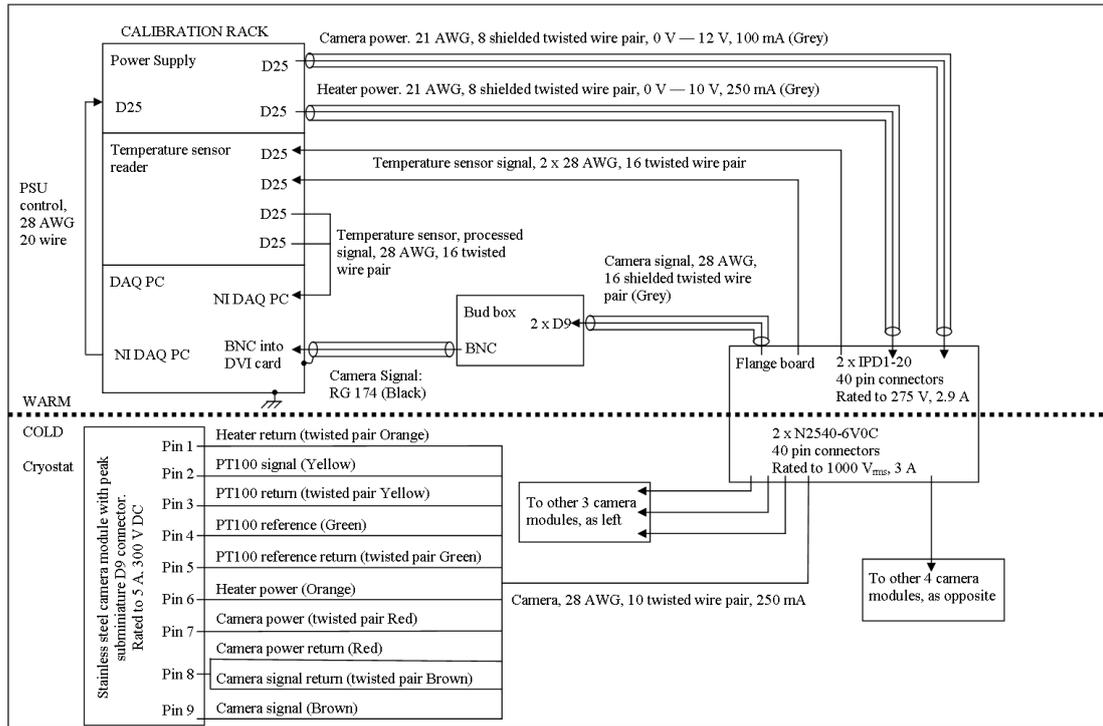


Fig. 3.3 The system diagram for the 35 ton camera system. The contents of the rack which was used in the camera system (top left), the wires which were terminated at the flange board (right), and the signals that each of the wires connected to the cameras carried (bottom left), are shown. In addition to this, the wire gauge (in units of American Wire Gauge), along with the voltages and currents which they carried, and the number of wires which were bound together, are shown for the connections between each subsystem and the flange board. Shielded wires are shown as wires with cylindrical covers around them, such as the wire labelled “Camera power.” In total there were eight camera modules installed in the cryostat, connected in two groups of four cameras, to the flange board by two 40 pin connectors.

3.3 Performance in the 35 ton

In total eight cameras were installed inside the 35 ton cryostat, the locations of which were selected to maximise the potential of observing any HV breakdowns which may occur. It was decided that the cameras would be focused on the following parts of the detector;

- The top right hand corner of the cathode.
- The bottom right hand corner of the cathode.
- The top left hand corner of the cathode.
- The bottom left hand corner of the cathode.
- The location of the high voltage feed through.
- The ullage - the gap between the top of the TPC, and the roof of the cryostat
- The cool down sprayers - this was only for use during cool down, to check that they were operating as expected.
- The phase separator - this was for use during operation to ensure that it was running as expected.

Upon installation in the cryostat, some calibration images were taken for each camera, these are shown in Figure 3.4.

As previously mentioned, the cameras used were car reversing cameras, and so this is why there are distance lines on some of the images in Figure 3.4. From Figure 3.4, it is also evident that there is a large variation in the picture quality of the different cameras. This is attributed to aspects of the cabling, as some cables were slightly longer than others, or had less reliable electrical connections. It could also be due to differences in the cameras, because, as discussed in Section 3.2, the stability of the cameras which were tested was not uniform. This variability in the camera stability, also manifested itself in the amount of signal degradation which was observed over time. Figure 3.5 shows the signal degradation over the course of the run for two of the eight cameras.

From Figure 3.5, it can be seen over the course of the run the fields-of-view for both cameras became more pixelated, and that there was an increase in the number of pixels which were either saturated or dead. The large number of saturated pixels was particularly evident when the cryostat was not illuminated, and meant that the region which could be

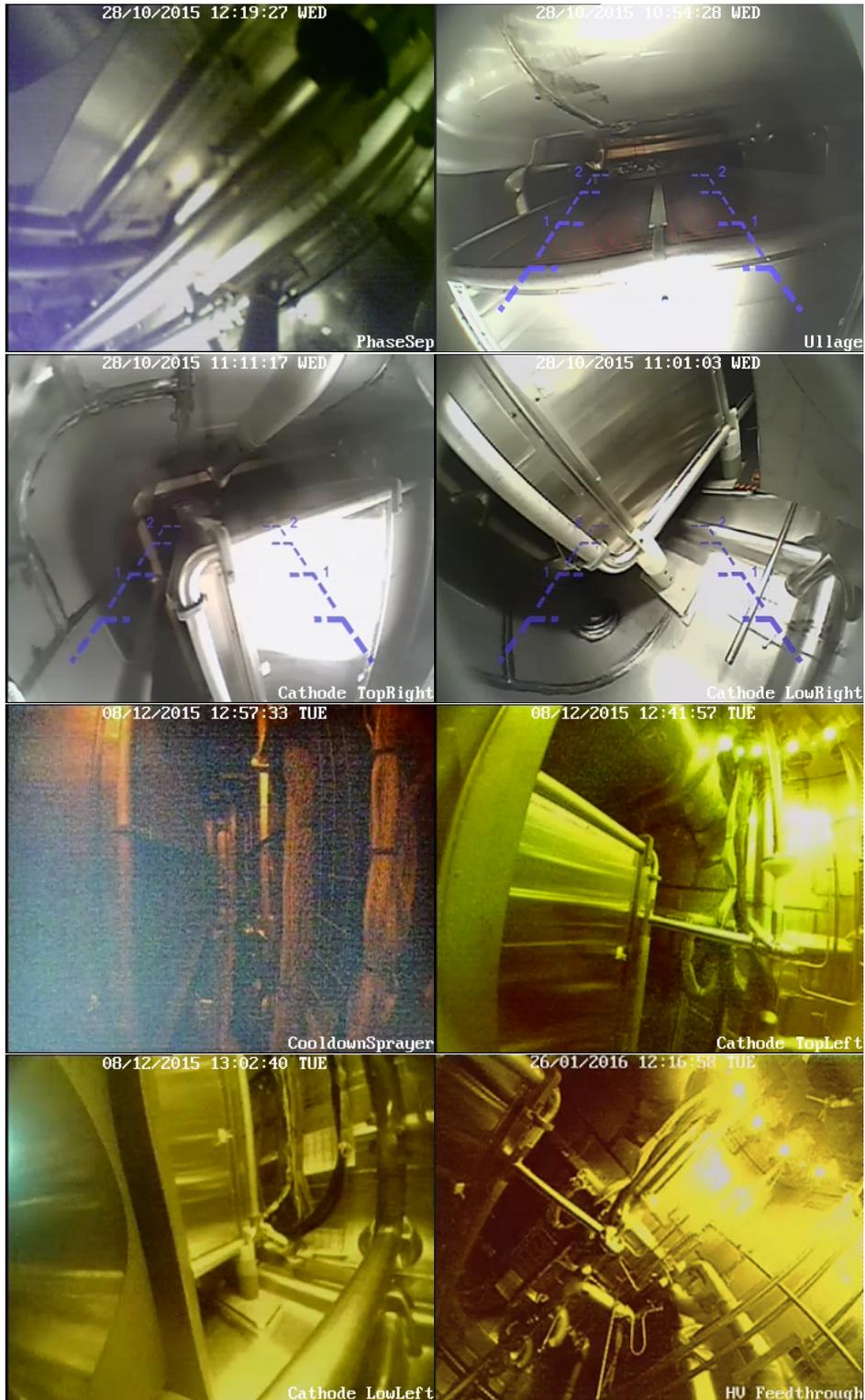


Fig. 3.4 The calibration images for the 8 cameras used in the 35 ton camera system. The images from left to right, top to bottom, show the following: phase separator, ullage, cathode top right, cathode bottom right, cool down sprayers, cathode top left, cathode bottom left, and high voltage feed through. The upper four images were taken with a halogen light illuminating the cryostat, prior to it being sealed up. The lower four images were taken after the cryostat was sealed, and was illuminated by a ring of LEDs. All images are left-right inverted due to software.



(a) The degradation in signal quality of the camera focused on the ullage.



(b) The degradation in signal quality of the camera focused on the lower left corner of the cathode.

Fig. 3.5 The signal degradation over time for two cameras in the 35 ton camera system. Top shows the signal degradation observed for the camera focused on the ullage. Bottom shows the signal degradation for the camera focused on the lower left corner of the cathode. For both cameras, left shows the field-of-view prior to cool down, centre shows the field-of-view immediately after cool down, and right shows the field-view after 10 weeks of operation. These are full colour images, as recorded by the DVR, and so no post-processing has been performed on the images.

used to trigger on HV breakdowns was severely limited for some cameras. As shown, this degradation was highly camera specific, with the resolution of the ullage camera deteriorating severely, whilst the camera focusing on the lower left corner of the cathode is relatively unchanged over the course of the run. The cause of this highly specific camera degradation is thought to be due to individual camera variability, as during testing large variations in stability were observed and whilst only the most resilient cameras were installed, they could not be exposed to the cold for many months prior to installation due to time constraints.

A striking difference between the left-most images, and the central and right-most images, in Figure 3.5, is the loss of colour in the images which are produced. This is attributed to a partial failure of the on-board encoding circuits, which resulted in the colour signal streams not being functional when exposed to cryogenic temperatures. This loss of colour was seen as soon as the cameras were exposed to cryogenic temperatures, and was also observed in the initial tests performed in Sheffield.

It is important to note that the cameras were able to operate safely within their modules, and did not impinge on any other systems in the cryostat. They were also found to be able to be power cycled after large periods of time in the cold, including after not being operational for a large period of time. One such period, which lasted for 9 days, was when the cameras were powered off whilst extensive noise hunting was performed. Upon being power cycled, all 8 cameras were immediately brought back online, without the need for the inbuilt heaters.

During normal operation of the 35 ton system, the high voltage operated stably at 60 kV, and so no breakdowns were observed. However, after the cathode was raised to 135 kV in low-purity argon, three of four breakdowns were detected by the camera system, and data was written to disk as expected. Unfortunately, the cameras were not able to pinpoint the exact locations of the breakdowns, and so their effectiveness at discerning the locations of HV breakdowns is still largely untested. This could be due to a lack of sparks being produced in correlation with the breakdowns, or the spark could have been outside of the field of view of the cameras. However, despite this the cameras proved to be a very valuable visual monitoring tool in the 35 ton cryostat, and there has been interest in including analogous systems in future LArTPCs, such as SBND [137].

Chapter 4

Simulations of the 35 ton prototype

4.1 Determination of interaction times

As outlined at the end of Section 2.5, it is important to know the interaction time of a track when performing reconstruction. When performing simulations, the simplest interaction time to assign to a reconstructed object is the Monte Carlo truth time of when the particle was created. The generation time can be used, as the time taken to travel the distances considered in simulations (less than 100 ns), is small when compared to the resolution of the detector (500 ns). When matching a reconstructed object with a GEANT4 particle, the particle which contributed the most overall deposited charge to the whole track is chosen. In order to calculate this, the energy contribution of each particle to the total energy for each hit on the track is calculated. The particle which contributed the most overall charge to the track is then calculated by summing the energy contributions for all particles, over all hits in the track. The ability to calculate the true interaction times of 3D objects, such as tracks, is vital when wanting to benchmark how well other algorithms estimate interaction times, or to determine the efficiency of the tracking algorithms, as described in Section 4.3.

It was envisioned that there would be at least two ways in which interaction times could be assigned to tracks in the 35 ton detector, one using the external cosmic ray counters, and another using reconstructed scintillation light collected by the photon detectors. The cosmic ray counters were used extensively in the 35 ton data, as described in Section 5.4. However, in simulations the scintillation light was used, as this would have been more powerful during continuous running. This is because not all particles entering the detector would pass through the counters, but one would expect almost all of them to produce reconstructable scintillation light. The flashes of light are reconstructed using a pre-built library, which models the expected number of photoelectrons that would be measured on

each photon detector, given the 3D position of the source of the flash. Using this library, it is then possible to reconstruct the location of a flash in three dimensions, given the relative amounts of light that each photon detector collects. For example, less scintillation light will be collected for a flash that originated further away from the photon detectors. This library also takes into account the expected quantum efficiencies of each photon detector.

When trying to produce an association metric, a sample of 10,000 isolated positive muons generated with CRY at $T = 0$ is used. Isolated positive muons are used, as then the events should only contain one muon track, and one reconstructed flash. This should mean that matching the track and flash is trivial. The positive muons are generated outside of the detector with a constant y position, above the uppermost scintillation counters, and flat distributions in x and z . The muons have a range of angles, though they are all “downward-going”. It was clear that the photon detector reconstruction in the simulations worked well, when using the pre-built libraries, as the reconstructed flash source normally lay very close to the track which caused it. It was found that a Point of Closest Approach (PoCA) calculation, between points on the reconstructed track, and the reconstructed flash centre, gave an effective metric by which the flash and track could be associated. Other metrics, such as the distance between the flash and track centres, and the perpendicular distance between the flash centre, and the line joining the start and end of track were investigated, but found to provide less reliable metrics. The latter of these metrics is less effective because the reconstructed tracks are rarely straight lines, due to particles scattering as they travel through the LAr. Therefore, calculating the separation between the track and the flash for each point along the track, will result in a smaller minimum separation being calculated. A comparison of these metrics is shown in Figure 4.1.

Another metric by which flashes could be assigned to reconstructed tracks, is by utilising the relationship between the number of measured photoelectrons in the simulation, and the distance from the APAs at which they were produced. When considering two flashes of scintillation light that are produced at different distances from the APAs, it would be expected that more photoelectrons would be collected when the photons were produced closer to the APAs. This relationship is shown in Figure 4.2, where it can be seen that as drift distance increases, the number of measured photoelectrons decreases exponentially. It is possible to fit an exponential curve to this decay, and using this the distance between the APAs and the source of the flash can be predicted from the number of measured photoelectrons. This predicted position can then be compared to the expected x position of a reconstructed track, where the expected track position is calculated using the difference in flash time and hit

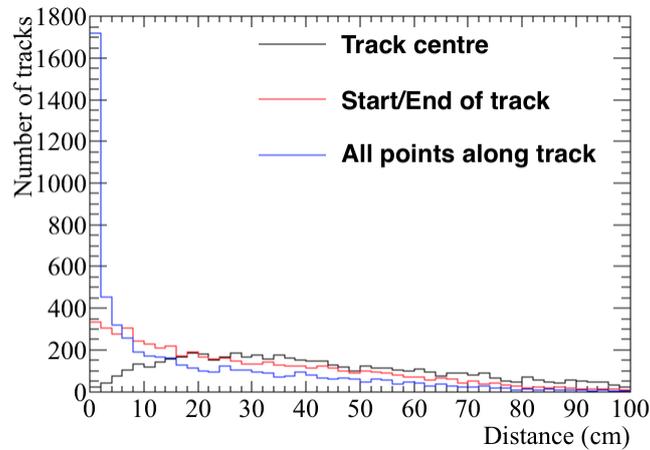


Fig. 4.1 The number of events as a function of the calculated distance between a reconstructed track, and a reconstructed flash, for various metrics. Black: the distance between the track centre and the flash centre. Red: the perpendicular distance between the flash centre, and the line joining the start and end of the track. Blue: the point of closest approach between the flash centre, and all hits along the track.

times and the electron drift velocity, this is shown in Figure 4.3. The difference in these two quantities can then be used as a second metric, as it gives a measure of how well matched the reconstructed flash and track are. The number of flash/track combinations which are “well matched,” can be seen by the collection of points around the $y = x$ line in Figure 4.3. These points correspond to flash/track combinations where the predicted and reconstructed x positions are identical, and so the track and flash have been accurately matched. It can be seen that this does not hold at large drift distances, this is because of the large degeneracy in the number of photoelectrons as a function of drift distance as in Figure 4.2.

Using these metrics, it is possible to attempt to assign reconstructed flashes to reconstructed tracks. Only flashes which are within one drift window of a given track are considered, as flashes outside of this time window cannot have been caused by the reconstructed track. Once a flash has been assigned to a track, it is possible to determine how well the matching has performed, by comparing the Monte Carlo truth interaction time with the photon detector interaction time. When doing this, it is more useful to use a CRY sample which spans multiple drift windows, as then incoming particles will create scintillation flashes at random times, as opposed to all at $T = 0$ as in the positive muon sample initially considered. The sample generated using CRY contains many particles, over a wide range of times, and is not limited to only producing positive muons. This means that it better represents the cosmic flux which will be observed by the 35 ton detector. The comparison between the Monte Carlo

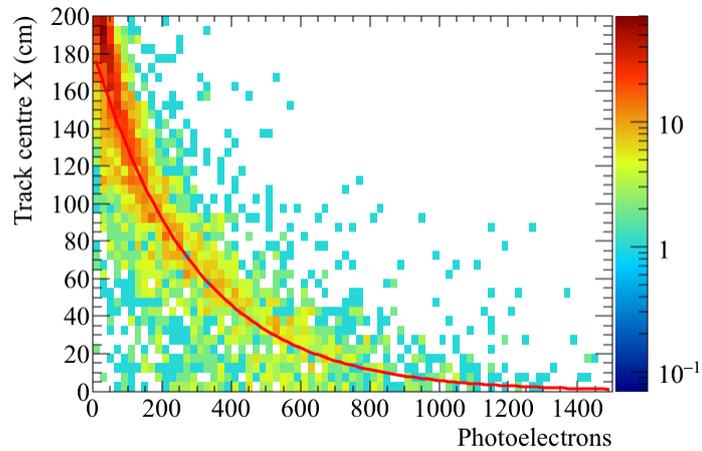


Fig. 4.2 The central x position of a reconstructed track versus the number of detected photoelectrons. The red line corresponds to a parameterisation of the distribution, which is used to predict the distance between the APAs and the flash source, given the number of measured photoelectrons in the flash. Using this parameterisation, the flash can be matched with a reconstructed track, whose centre is the same distance from the APAs.

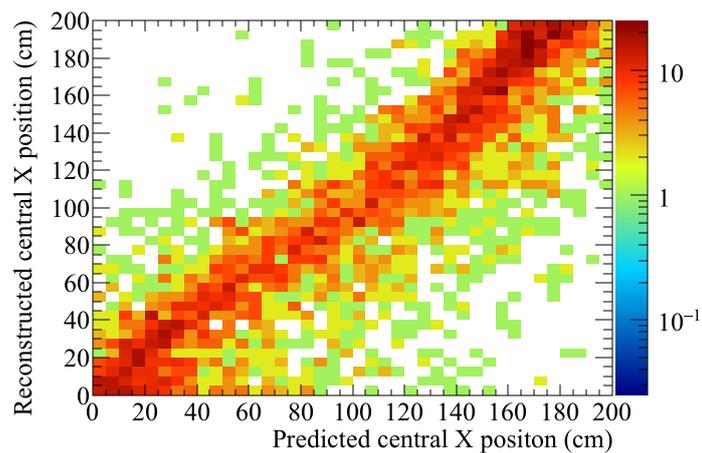


Fig. 4.3 A comparison of the x position predicted using the relationship in Fig 4.2, and the central x position of the reconstructed track.

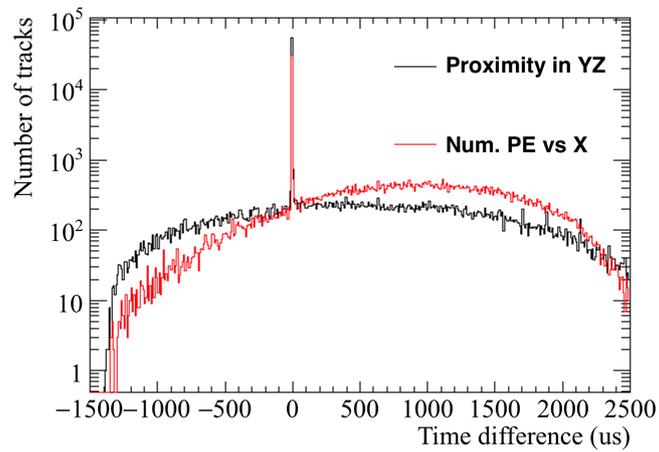
truth interaction time, and the photon detector interaction time, is shown in Figure 4.4.

Figure 4.4a shows a clear peak at a time difference of 0 ms between the Monte Carlo truth and photon detector times, showing that the interaction time which is calculated using the correlation between the photon detector and track information is often very accurate. However, there is also a significant number of tracks for which the photon interaction time does not match the Monte Carlo truth interaction time. This is due to, for example, the broad width of the distribution shown in Figure 4.2, meaning that the number of photoelectrons cannot necessarily accurately predict the distance between the APAs and the flash source. Improving the accuracy of the photon detector interaction time is required, as applying an incorrect interaction time correction will result in the x position of hits being incorrect. The peak seen at a time difference of 0 ms has been expanded in Figure 4.4b, where it can be seen that there is a systematic offset of $0.6 \mu\text{s}$, this is due to an electronics offset applied in the simulation to the photon detector system.

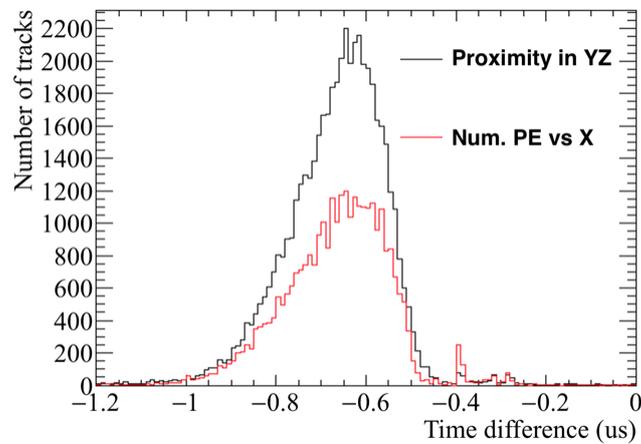
From Figure 4.4 it can be seen that the metric using the proximity of the flash centre to the track trajectory, yields the best track/flash matches. This is likely caused by the large spread in the number of photoelectrons collected at fixed drift distances, as shown by Figure 4.2, which causes a large degeneracy between the number of measured photoelectrons, and the distance between the APAs and the flash source. The two metrics can be combined to give a prediction for the interaction time, though given the increased sensitivity from the proximity metric, this should be given greater weighting. In physics data, the metric using the number of collected photoelectrons is particularly sensitive to the absolute light level in the detector. This is because a high residual light level in the detector would reduce the proportional change in the number of photoelectrons collected for increasing drift distances. This metric also relies on the existence of a sample of tracks with known x positions, upon which to calibrate the change in the number of photoelectrons collected for increasing drift distances. It may be difficult to obtain this dataset in a real detector, though the cosmic ray counters around the 35 ton may be able to provide such a sample.

4.2 Calibrating calorimetric constants

Having the correct calorimetric responses is vital when trying to calculate $\frac{dE}{dx}$, as the measured change in charge has to be correctly converted to the change in energy. The parameter which has to be tuned in order to ensure that this is done correctly, is the number of electrons that



(a) The difference in interaction times.



(b) Zoomed in at low time differences.

Fig. 4.4 The number of events as a function of the difference between Monte Carlo and photon detector times. Top: the difference in interaction times over a large range of times. Bottom: the peak at a time difference of 0 is expanded, showing a systematic offset of $0.6 \mu\text{s}$ due to an electronics offset in the simulation.

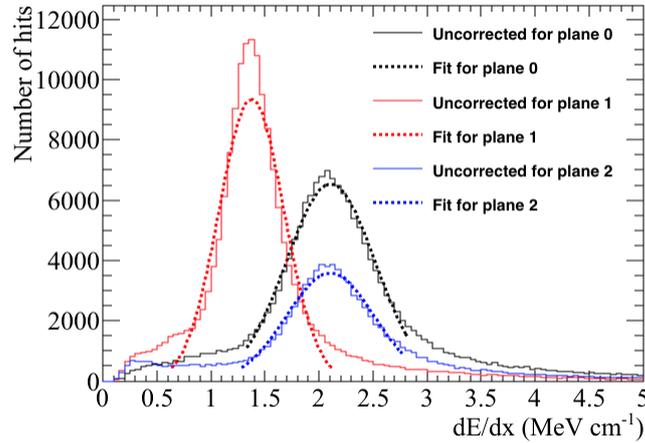
each ADC count corresponds to. This was presented in Equations 2.1 and 2.2, as $C_{ADC \rightarrow e^-}$. Each plane will have a different response function, and so each plane has to be treated separately. These parameters have to be tuned in such a way as to make a known particle energy deposition have the correct $\frac{dE}{dx}$, the easiest deposition to tune against is the minimum ionising particle (MIP) peak, which in LAr should have a value of $1.8 \text{ MeV}\cdot\text{cm}^{-1}$. To do this, the sample of 10,000 positive muons made to calibrate the photon detector track/flash assignment will be used, as many of these particles will be MIPs.

To select the MIPs in the sample, only tracks caused by through-going muons are used. The $\frac{dE}{dx}$ value for all hits, in all tracks, is then calculated, and the different planes are considered separately. A Gaussian distribution is fitted around the peaks for each of the planes, to discern the peak value of the $\frac{dE}{dx}$ distribution for that plane. If the MPVs are not equal to $1.8 \text{ MeV}\cdot\text{cm}^{-1}$ [26], the ADC to electron parameters are scaled by the factor between the measured MPV and the MIP peak. As the relationship between $\frac{dE}{dx}$ and $C_{ADC \rightarrow e^-}$ is not linear, an element of trial and error is required to find the correct ADC to electron parameters. An example of the calibration being applied is shown in Figure 4.5. Calibration of the response functions is required whenever the electronics gains, or signal shaping functions, are changed.

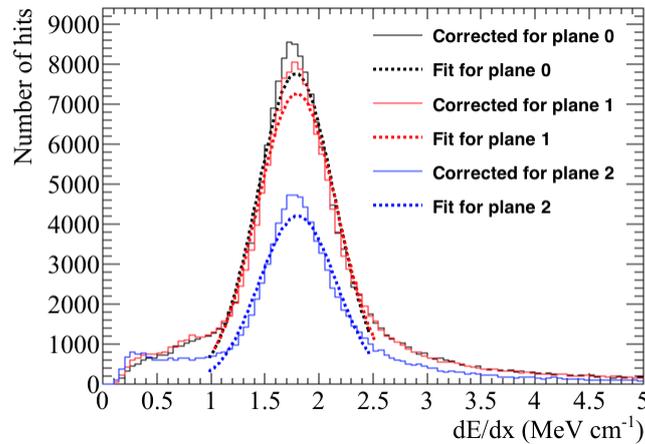
4.3 Discerning reconstruction efficiencies

Knowledge of the strengths and weaknesses of different tracking algorithms is vital when using them for physics analyses. To this end, it is useful to develop a metric by which they can be compared. In order to do this, a series of conditions have to be applied to the reconstructed tracks from a large set of simulated particles, which are reconstructed using different tracking algorithms. It is interesting to observe the effect that event complexity has on the reconstruction algorithms, and so efficiencies will be calculated for the positive muon sample, and the CRY sample, which were used in Section 4.1. The sample referred to as the positive muon sample contains single positive muons generated at $T = 0$, with a constant y position above the scintillation panels, and flat distributions in x and z . The sample referred to as the CRY sample, contains multiple particles of multiple particle types generated at times spanning multiple drift windows, at zero altitude above sea level.

The criterion upon which to determine whether a particle is well reconstructed has to be carefully chosen, as every definition will have limitations. A final definition of a well reconstructed track should depend on the accuracy of Particle IDentification (PID) power for



(a) Before calibration is performed.



(b) After calibration is performed.

Fig. 4.5 The number of hits as a function of the hit $\frac{dE}{dx}$, before and after calibration of the response functions for the conversion of ADC to number of electrons, for each plane. Top: the distribution of hit $\frac{dE}{dx}$ before calibration. Bottom: the distribution of hit $\frac{dE}{dx}$ after calibration. The collection plane, labelled as plane 2, is shown in blue, whilst the two induction planes, labelled plane 0 and plane 1, are shown in black and red respectively.

each of the algorithms, however the DUNE software is not currently at this level, and so a simple definition related to track length is used. However, the definition of this matched track length can still lead to uncertainties. For example, consider a particle that travels 100 cm in the active volume of the detector, but is reconstructed as 2 separate tracks (tracks 1 and 2), with lengths 77 cm and 23 cm respectively. Firstly, should these tracks be merged, or left separate? If the reconstruction algorithms have found them to be separate tracks, then it is likely that it would be difficult to ascertain that they are from the same particle in real data, and so in considerations here they are not merged. Secondly, one has to determine what the definition of a well reconstructed track should be. One definition of a well reconstructed track, would be that the reconstructed track length is between 75% and 125% of the track length in the detector from Monte Carlo truth, in which case track 1 would be considered well reconstructed. Another definition to consider however, would be whether the track length in the detector from Monte Carlo truth is between 75% and 125% of the reconstructed track length, in which case neither track would be considered well reconstructed. These definitions have used exactly the same tracks, and seemingly identical definitions of what constitutes a well reconstructed track, but they have got very different results. As such, it is wrong to say which definition gives the correct result, but instead the result of each should be considered equally. It should also be noted, that these are just two of a wide range of definitions one could use to quantify whether a track is well reconstructed. In the discussions presented here, the former definition of a well reconstructed track is used, such that a track is considered well reconstructed if:

- The reconstructed track length is more than or equal to 75% of the Monte Carlo track length.
- The reconstructed track length is less than or equal to 125% of the Monte Carlo track length.
- Only one reconstructed track can be matched per Monte Carlo particle.

When calculating the reconstruction efficiencies, the number of well reconstructed tracks, is divided by the total number of particles in the active volume, from Monte Carlo truth. When calculating these efficiencies, it is important to consider much more than just the Monte Carlo truth track length. To this end, efficiencies as a function of several parameters of the tracks are calculated:

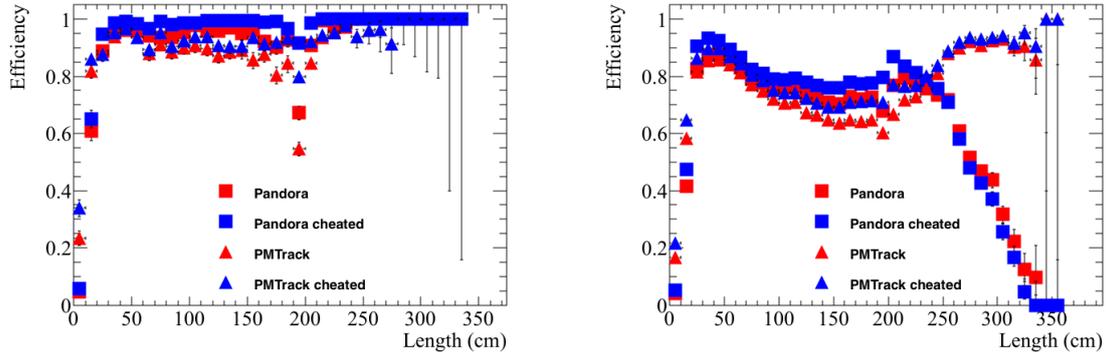
- Track length,
- Energy deposited in the active volume of the detector,

- The angle θ of the track, defined as the angle that a vector makes from the x axis in the xy plane,
- The angle ϕ of the track, defined as the angle between the z axis and the vector.

In all efficiency plots, the Monte Carlo truth quantity, not the reconstructed quantity, is shown, so as to reflect how the change in these quantities affects the reconstruction efficiency. It is also useful to observe the effect that failed disambiguation, and incorrect interaction time determination, has on the reconstruction efficiency. To show this, two reconstruction paths are run on the particles. One reconstruction path uses no truth information, and so the disambiguation is performed as outlined in Section 2.5, and the interaction time is determined using the simulated photon detectors, as described in Section 4.1. The second reconstruction path uses “cheated disambiguation,” whereby Monte Carlo truth information is used to select the wire segments which hits occurred on, and Monte Carlo truth information is also used when calculating the interaction time of the track.

The calculation of reconstruction efficiencies also serves as an effective method upon which reconstruction algorithms can be further developed, as it identifies aspects which do not work as expected. For example, when the reconstruction efficiencies for the CRY sample were initially calculated, they were significantly lower than for the positive muon sample (at 10%), but only when disambiguation was not cheated. It transpired that this was because the disambiguation was only selecting the largest collection of hits on each plane, for each TPC. This is not a problem when only 1 particle is simulated, as one would expect that there only be one particle on a given plane, in each TPC. However, when considering a CRY sample which lasts of the order of 10 ms, there will almost certainly be multiple particles on each plane, in each TPC. Removing the hits from all but one of these multiple particles, will cause them to have no reconstructed track, and thus cause the efficiency to drop significantly. Upon making the disambiguation algorithm no longer have this restriction, the reconstruction efficiencies of the positive muon and CRY samples were observed to become much more similar, as presented here.

The reconstruction efficiencies given the current state of the most commonly used reconstruction algorithms, Pandora [124] and PMTrack [89], are shown in Figures 4.6, 4.7, 4.8, 4.9 and 4.10. Efficiencies are shown for both the positive muon and CRY samples, where it can be seen that the efficiency tends to be lower for the CRY sample. It is thought that this is due to the more complex event structure in the CRY sample, as multiple primary particles will be present in the detector at any given time. The relatively slow drift velocity of LAr may mean



(a) Reconstruction efficiencies for the positive muon sample.

(b) Reconstruction efficiencies for the CRY sample.

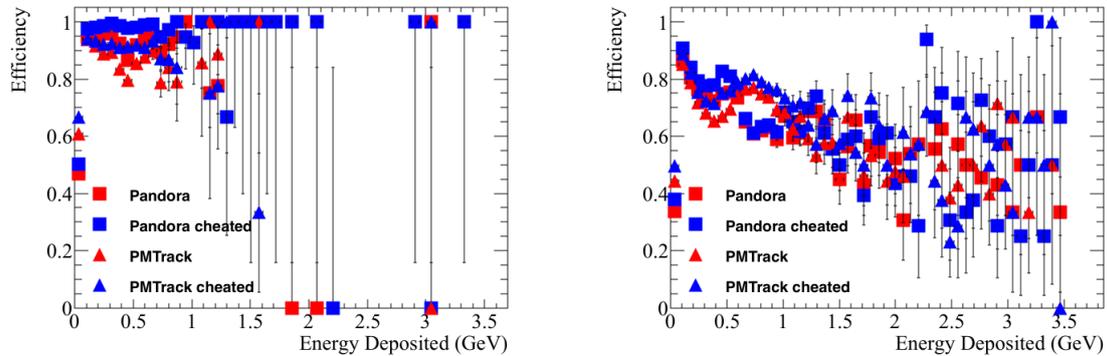
Fig. 4.6 The reconstruction efficiencies for simulated events as a function of the track length in the detector from Monte Carlo truth. The efficiencies are shown for “non-cheated” reconstruction (red), and “cheated” reconstruction (blue), for both Pandora [124] (squares) and PMTrack [89] (triangles).

that these tracks cross in wire-tick space. Tracks crossing in wire-tick space could cause reconstruction errors, as the overlaps may be mistaken for interactions, which would cause the tracks to be split, resulting in the interaction time calculated from the photon detectors to be incorrect. This error, in the calculation of interaction time using the photon detectors, was seen in Figure 4.4. The reconstruction efficiencies for the CRY sample are more realistic, as events will rarely be isolated in the 35 ton detector, due to the large flux of cosmic particles on the Earth’s surface.

A striking feature of Figure 4.6 is the rapid decrease in reconstruction efficiency seen in the CRY sample, for particles with track lengths in the detector of more than 250 cm from Monte Carlo truth, when using Pandora. The cause of this, is that tracks are reconstructed separately in the long, and short drift volumes, before being merged when they are found to be co-linear in the yz plane. This is not a problem in the positive muon sample, as the x position of the hits calculated using Equation 4.1 will be correct.

$$x_{Hit} = T_{Hit} \times v_{Drift} \quad (4.1)$$

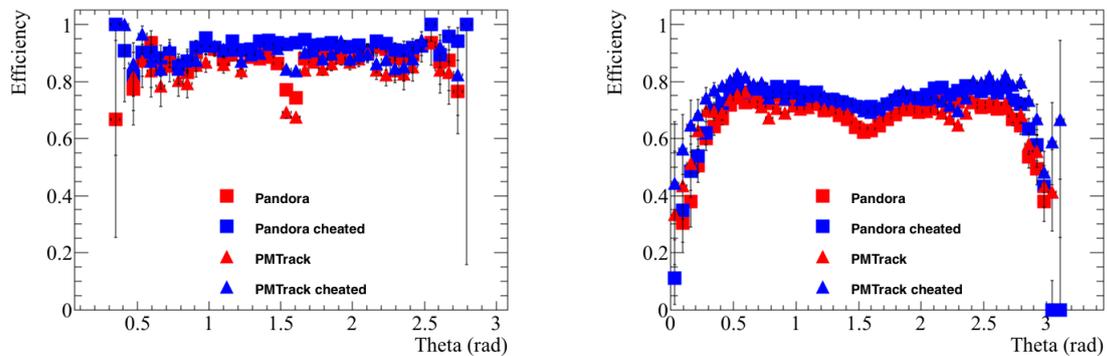
In Equation 4.1, x_{Hit} is the calculated x position of the hit, T_{Hit} is the measured time of the hit, and v_{Drift} is the electron drift velocity. An electron, in an electric field of $500 \text{ V}\cdot\text{cm}^{-1}$, in LAr, drifts at a speed of $0.161 \text{ cm}\cdot\mu\text{s}^{-1}$. However, when the same is done for hits in the CRY sample, using particles with large interaction times, the x positions will have



(a) Reconstruction efficiencies for the positive muon sample.

(b) Reconstruction efficiencies for the CRY sample.

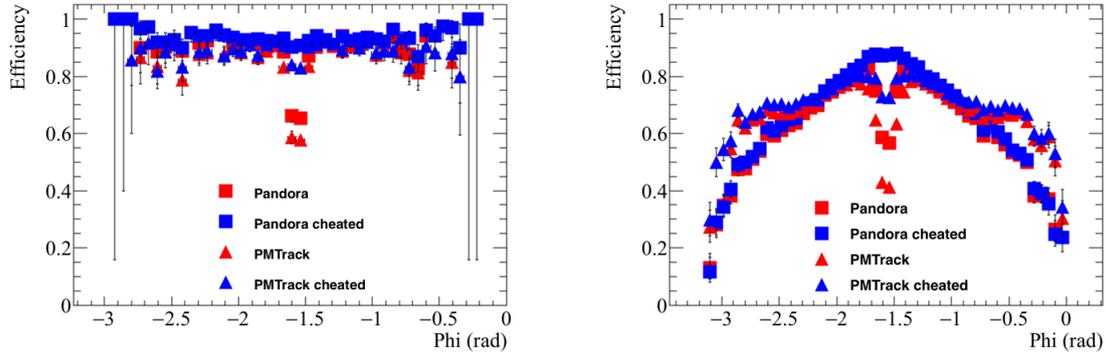
Fig. 4.7 The reconstruction efficiencies for simulated events as a function of the deposited energy from Monte Carlo truth. The efficiencies are shown for “non-cheated” reconstruction (red), and “cheated” reconstruction (blue), for both Pandora [124] (squares) and PMTrack [89] (triangles).



(a) Reconstruction efficiencies for the positive muon sample.

(b) Reconstruction efficiencies for the CRY sample.

Fig. 4.8 The reconstruction efficiencies for simulated events as a function of the θ track angle from Monte Carlo truth track. The efficiencies are shown for “non-cheated” reconstruction (red), and “cheated” reconstruction (blue), for both Pandora [124] (squares) and PMTrack [89] (triangles).



(a) Reconstruction efficiencies for the positive muon sample.

(b) Reconstruction efficiencies for the CRY sample.

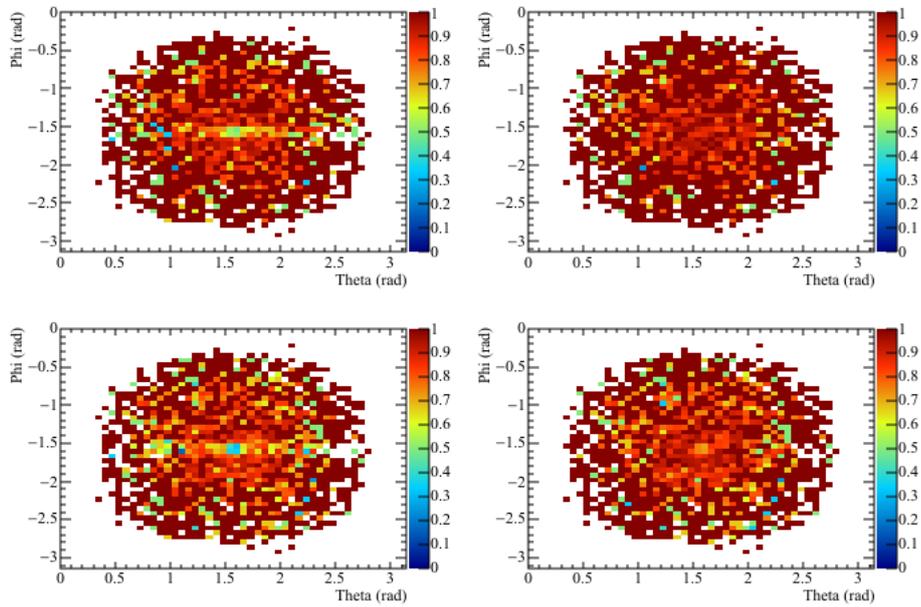
Fig. 4.9 The reconstruction efficiencies for simulated events as a function of the ϕ track angle from Monte Carlo truth track. The efficiencies are shown for “non-cheated” reconstruction (red), and “cheated” reconstruction (blue), for both Pandora [124] (squares) and PMTrack [89] (triangles).

offsets proportional to the interaction time of the particle, unless the hit time is corrected by Equation 4.2.

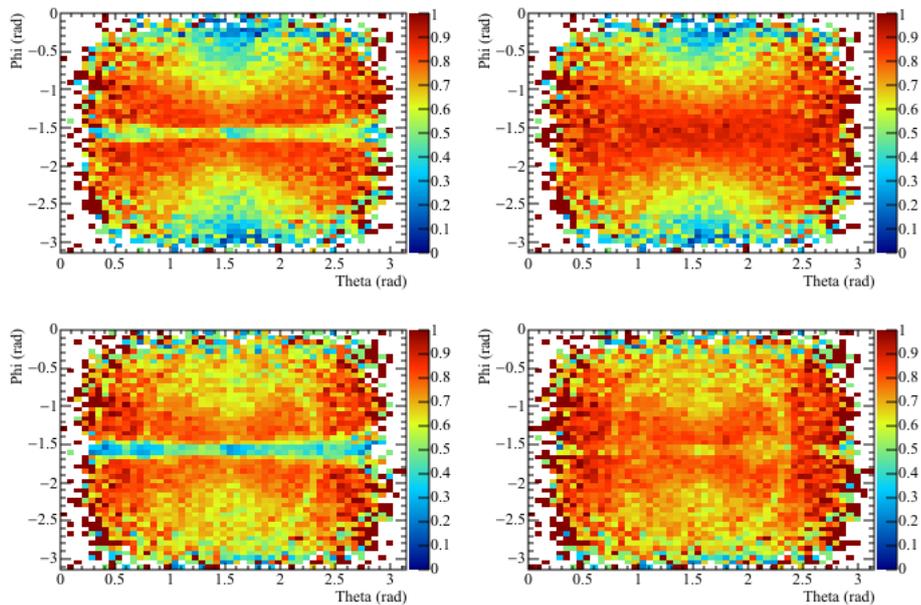
$$T_{Hit} = T_{Measured} - T_{Interaction} \quad (4.2)$$

In Equation 4.2, T_{Hit} is the corrected hit time, $T_{Measured}$ is the measured time of the hit, and $T_{Interaction}$ is the calculated interaction of the particle which caused the hit.

The scale of these offsets in x positions can be seen by considering a particle which is generated at a time of $T = 12.5$ ms, as then the offset in the reconstructed x position calculated by Equation 4.1, would be more than 20 m! Obviously the hits could not have occurred at these positions, as the drift distances are roughly 30 cm in the “short” drift volume, and 225 cm in the “long” drift volume. However, as track segments are reconstructed separately in the “short” and “long” drift volumes, before later being merged, it is possible for there to be a discontinuity in x of more than 40 m, if these x offsets are not corrected for. As the interaction time of the track is calculated using the output of the tracking algorithms, it is not possible to prevent this by using the interaction time at present. It is however, possible to subtract this discontinuity in x from the total reconstructed track length when the stitched track is stored in the event. This will give the correct track length, though the user will still have to correct the positions of individual hits, using the calculated interaction times. This is what is done by PMTrack, hence it not exhibiting this rapid decrease in reconstruction efficiency for long tracks. The interaction time can be found from, among other things, the



(a) Reconstruction efficiencies for the positive muon sample.



(b) Reconstruction efficiencies for the CRY sample.

Fig. 4.10 The reconstruction efficiencies for simulated events as a function of the θ and ϕ track angles from Monte Carlo truth track. The track angle θ is shown on the x axis, and the track angle ϕ is shown on the y axis. The colour z axis shows the reconstruction efficiency. In both figures, “normal” (left) and “cheated” (right) disambiguation is shown, for both Pandora [124] (top) and PMTrack [89] (bottom).

Monte Carlo truth generation time, or the photon detectors, as discussed in Section 4.1.

It is clear from Figure 4.6 that particles with track lengths in the detector of less than 30 cm are poorly reconstructed. The extremely low efficiency for particles with track lengths of less than 10 cm, can be largely attributed to protons with track lengths of less than 1 cm in the active volume of the detector, or muons which clip the edges of the detector. These particles produce tracks which are too short to be reconstructed using the current reconstruction process. Particles which produce tracks of this length will need to be reconstructed when looking for supernovae bursts, though special algorithms will be written to do this, as the traditional hit finding and clustering algorithms may discard them due to the isolated nature of the hits. Another issue, is that the low energies of these particles may mean that the signals that they produce are below threshold, and so will not even be reconstructed, or, if hits are reconstructed, they may be too close to a more energetic track and get absorbed into them. The reconstruction of tracks is affected by the number of wires which they cross, though this should not matter for particles with track lengths of more than 5 cm in the active volume, as they will have crossed roughly 10 wires in each plane, which should produce enough unique hits for a cluster to be reliably reconstructed. This can be seen to be the case for PMTrack when considering the positive muon sample, as the efficiency for particle track lengths between 10 and 20 cm, is roughly the same as that for track lengths between 20 and 30 cm. However, when considering the CRY sample, there is still a significant decrease in efficiency. This is attributed to secondary particles which are produced in hadronic interactions with the concrete surrounding the detector. Many of these particles will travel only very short distances in the active volume, though those that travel slightly larger distances are likely to cause energy depositions that will be confined to the detector edges. The tracking algorithms may struggle to accurately reconstruct these tracks, as significant portions of the track will be close to the detector edge, where the field is poorly modelled, and hits may be more difficult to disambiguate.

The trend of increasing efficiency for longer track lengths, which was seen in Figure 4.6, can also be seen in Figure 4.7, where the reconstruction efficiency increases as the amount of deposited energy increases. This is because particles which deposit more energy, will tend to have travelled further in the detector. The amount of energy that particles deposit is limited by the size of the detector, as particles with an energy of more than 1 GeV are energetic enough to be MIPs. This results in few particles depositing more than 1 GeV in the detector, and is why the uncertainty in the reconstruction efficiency increases for energy depositions in the detector volume of more than 1 GeV. The larger range in the amount of energy deposited

seen in Figure 4.7b, is due to the larger number of muons in the CRY sample which create large electromagnetic showers upon entering the LAr.

It is also interesting to note the pronounced decreases in reconstruction efficiencies for particular angles, shown in Figures 4.8 and 4.9. The decrease in efficiency at $\phi = \frac{\pi}{2}$, seen in Figure 4.9, can be attributed to the drop in efficiency for particles with track lengths between 190 cm and 200 cm, seen in Figure 4.6. This is because the vertical height of the detector is approximately 195 cm, and near vertical tracks will hit few collection wires, meaning that determining the triple points needed by the disambiguation is very difficult. This is verified by the large increase in efficiency achieved by cheating the disambiguation, as seen in Figure 4.8a, where the reduction in reconstruction efficiency is seen to become much less pronounced. Similarly, the decrease in efficiency at $\theta = \frac{\pi}{2}$ can be attributed to particles which are perpendicular to the collection plane wires, as this track orientation also results in few collection wires being hit.

The information from Figures 4.8 and 4.9 is combined in Figure 4.10, where the sharp drop in efficiency at $\phi = \frac{\pi}{2}$ for the “non-cheated” CRY sample is particularly visible. The effect of cheated disambiguation is clear in Figure 4.10b, where the dip in efficiency as a function of θ at fixed $\phi = \frac{\pi}{2}$ is completely removed. The same is not true however, for the dip in efficiency as a function ϕ at fixed $\theta = \frac{\pi}{2}$, though the reduction in efficiency was not as severe as that seen for fixed values of $\phi = \frac{\pi}{2}$. The effect of “cheated disambiguation” can still be seen though, as the reconstruction efficiency in Figure 4.10b can be seen to improve for values of ϕ . There are still however, noticeable decreases in the reconstruction efficiency for values of ϕ close to 0 or π , when using Pandora. The improvement in the performance of the reconstruction algorithms that comes from “cheating” the reconstruction, is part of the motivation for the wire pitches in the DUNE FD being 36° , as opposed to the $45 \pm 0.7^\circ$ used in the 35 ton. This is because, as discussed in Section 2.5, the shallower wire pitch makes disambiguation easier. Though disambiguation will be easier in the different geometry, further efforts to improve disambiguation are still required, as are continued efforts to reconstruct the shortest tracks.

4.4 Performing particle identification

Being able to perform reliable Particle IDentification (PID) is a key requirement for the DUNE experiment, and so efforts have been made to establish a procedure by which this can

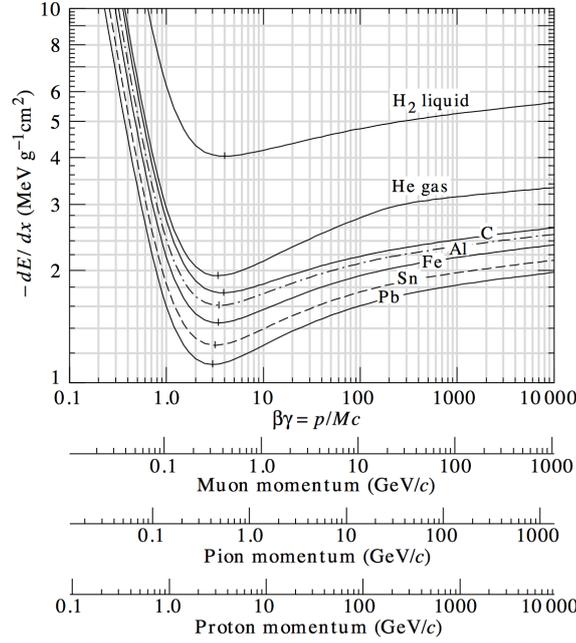


Fig. 4.11 The mean energy loss per unit track length of different particle masses in different materials [26]. The $\frac{Z}{A}$ of liquid argon is slightly less than that of carbon.

be achieved. The predominant method of performing PID in LAr is to use the relationship between $\frac{dE}{dx}$, and the residual range of the track. The residual range of a track is defined as, the distance between a given point on the track and the stopping point of the track. The rate at which energy is lost ($\frac{dE}{dx}$) can be seen to be highly dependent on the value of β , where a track will stop once it has low β and hence zero residual range. The relationship between $\frac{dE}{dx}$ and β is quantified by the Bethe-Bloch equation [138, 139], which is shown in Figure 4.11 and presented in Equation 4.3. From Figure 4.11 it can be seen that the rate of energy loss for a given value of β is highly dependent on the medium which a particle is passing through; in the case of LAr, it has a $\frac{Z}{A}$ which is slightly less than carbon.

$$-\frac{dE}{dx} = Kz^2 \frac{Z}{A} \frac{1}{\beta^2} \left[\frac{1}{2} \ln \frac{2m_e c^2 \beta^2 \gamma^2 T_{max}}{I^2} - \beta^2 - \frac{\delta(\beta\gamma)}{2} \right] \quad (4.3)$$

where $K = 4\pi N_A r_e^2 m_e c^2$, N_A is Avogadro's number, r_e is the classical radius of the electron, m_e is the electron mass, z is the charge of the particle, Z is the atomic number of the material, A is the atomic mass of the material, T_{max} is the maximum energy transferred to the ionisation electron, I is the mean excitation energy, and $\delta(\beta\gamma)$ is the density effect correction to ionisation energy loss.

Table 4.1 Stopping power parameterisation for various particle types in liquid argon. The table is taken from [128], where NIST refers to the the National Institute of Standards and Technology.

Particle	A (MeV·cm ^{-(1+b)})	b
Pion	8	-0.37
Kaon	14	-0.41
Proton	17	-0.42
Deuteron	25	-0.43

The sharp increase in energy loss per unit length can be seen to occur at different momenta for different particle masses in Figure 4.11, meaning that the peak value of $\frac{dE}{dx}$ changes significantly for particles of different masses. For example, it can be seen that the momentum at which the increase in $\frac{dE}{dx}$ occurs is very different for muons and protons. However, it can be seen the momentum at which this increase in $\frac{dE}{dx}$ occurs is very similar when considering muons and pions. For this reason, it should be relatively simple to discern a proton track from a muon track, though it may be difficult to discern a muon track from a pion track.

The particle mass dependence can be seen by plotting the $\frac{dE}{dx}$ against the residual range of the particle on a log-log plot, as shown in Figure 4.12. A power law dependence is found to describe the relationship [128], as shown in Equation 4.4.

$$\frac{dE}{dx} = AR^b \quad (4.4)$$

where A and b are constants for a particle and a given material. The dependence on b is found to be weak, and so can be set to -0.42 for all particle masses. This means that the main discriminant used is the A parameter, which has a strong dependence on particle mass. Table 4.1 shows the values of A and b which are calculated from Figure 4.12. It is found that the error introduced by fixing the b parameter is small compared to the error from ionisation fluctuations [128].

Once the b parameter is set to be constant for all particle types, it is possible to calculate a value of the A parameter for each hit on the track using Equation 4.5, where R_i is the residual range of the track at that point.

$$A_i = \left(\frac{dE}{dx} \right)_i \times (R_i)^{0.42} \quad (4.5)$$

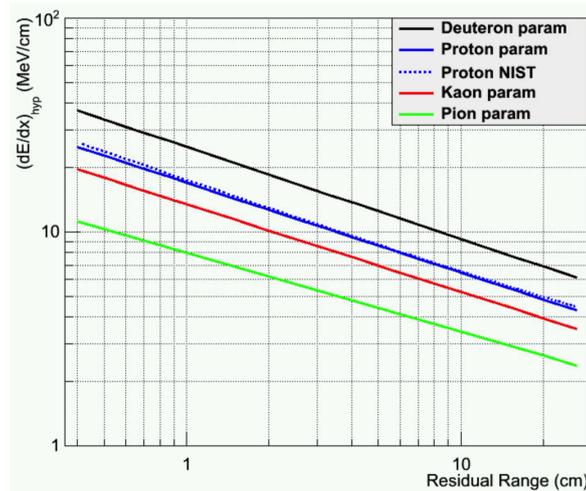


Fig. 4.12 Stopping power for different particle masses as a function of residual range in liquid argon. The figure is taken from [128].

The particle type discriminant, called PIDA, can then be calculated for a track by finding the average value of A_i for the track. As the particle mass dependent increase in $\frac{dE}{dx}$ only occurs near the end of the track, the PIDA variable can only be calculated for particles which stop in the detector, as all other particles will have MIP-like $\frac{dE}{dx}$ distributions and so cannot be identified in this way. As shown by the plotted range of Figure 4.12, the average value of A is normally calculated for the last 30 cm of the track.

The PIDA method was tested in [128], where the PIDA values were calculated for simulated particles which stopped in the detector, using Monte Carlo truth information over the last 30 cm of the particle tracks. This is shown in Figure 4.13, where a clear separation can be seen between the peaks for muons, pions, kaons, and protons. Though the muon and pion peaks are relatively close together, they can still be resolved in the plot due to little overlap. It is interesting to note how tight the PIDA distributions found in the paper are, which allows the different particles types to be cleanly separated in the truth study. An incorrect tuning of the electron recombination effects will cause the distributions in Figure 4.13 to become broader [128]. The dependence of $\frac{dE}{dx}$ on the recombination effects ($Recomb_{A/B}$) were presented in Equations 2.1 and 2.2. Also, an incorrect calibration of the detector will introduce a systematic shift in the expected values of PIDA [128], this is why the work presented in Section 4.2 is important.

From Figure 4.13, it can be seen that the most distinct PIDA distributions are that of muons and protons, these are also two of the most common charged particles in cosmic

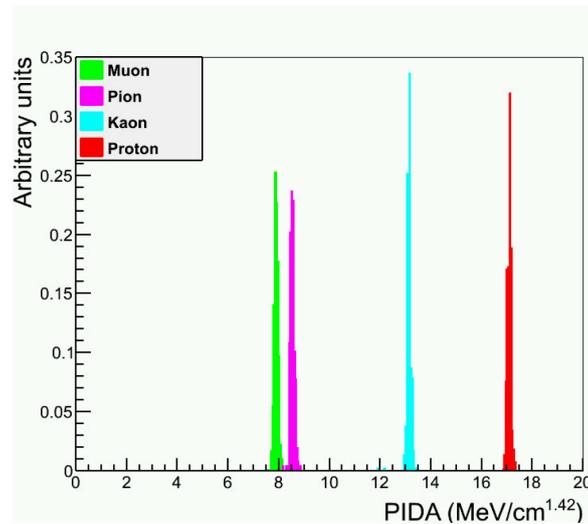
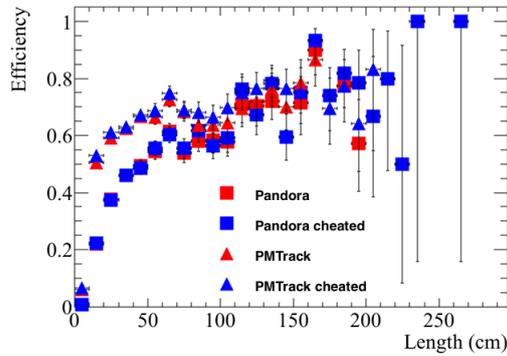


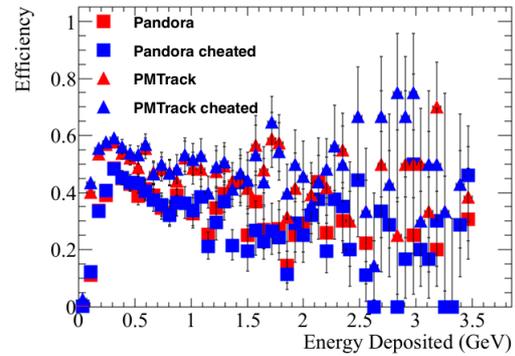
Fig. 4.13 The distribution of PIDA values, calculated using Monte Carlo truth, for different particle masses. The figure is taken from [128], note the units for PIDA should be $\text{MeV} \cdot \text{cm}^{-0.58}$.

rays. For these reasons, particle identification using the PIDA variable, will be attempted in simulations of the 35 ton detector. As outlined in Sections 4.1 and 4.2, in order to do this the interaction times of particles have to be well known, and the calibration constants must be tuned, so as to ensure that the effects of recombination are properly accounted for. It is also useful to use the information found in Section 4.3 about the efficiency with which tracks are reconstructed. In this regard, it is useful to produce additional figures showing the reconstruction efficiencies of protons in the CRY sample (Figure 4.14).

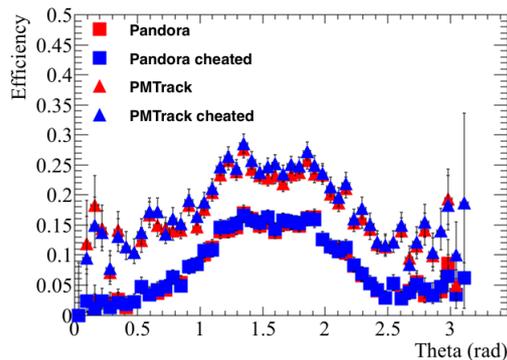
Figure 4.14 shows that the average reconstruction efficiency for PMTrack is higher than that for Pandora when considering protons. This can be easily seen in Figure 4.14c, where the efficiency for PMTrack is roughly 10% higher than that of Pandora, for all values of θ . The reconstruction efficiency is still much lower than the overall efficiency shown in Figure 4.8, for both the positive muon and CRY samples though. This shows that the overall reconstruction efficiency for protons is quite low. Comparing Figures 4.6b, and 4.14a, it is evident that the reconstruction efficiency for protons with track lengths of more than 10 cm is reasonably similar to that of the overall reconstruction efficiency for the CRY sample, when using PMTrack. However, the reconstruction efficiency is significantly lower for protons with tracks of less than 10 cm. When using Pandora to reconstruct protons, the reconstruction efficiency is lower for all track lengths. It is found that 60% of simulated protons have track lengths of less than 1 cm, and that none of these particles are reconstructed. It is this large number of very short tracks which causes the overall reconstruction efficiency to be



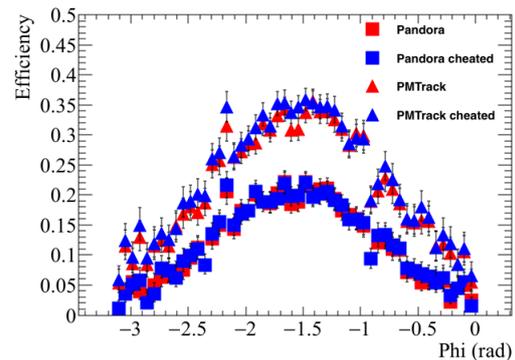
(a) The reconstruction efficiency as a function of the track length in the detector from Monte Carlo truth.



(b) The reconstruction efficiency as a function of the deposited energy in the detector from Monte Carlo truth.



(c) The reconstruction efficiency as a function of the θ track angle from Monte Carlo truth.



(d) The reconstruction efficiency as a function of the ϕ track angle from Monte Carlo truth.

Fig. 4.14 The reconstruction efficiencies for protons in a sample generated using CRY. Top left: the efficiencies as a function of the track length in the detector from Monte Carlo truth. Top right: the efficiencies as a function of the deposited energy from Monte Carlo truth. Bottom left: the efficiencies as a function of the θ track angle from Monte Carlo truth. Bottom right: the efficiencies as a function of the ϕ track angle from Monte Carlo truth. The efficiencies are shown for “non-cheated” reconstruction (red), and “cheated” reconstruction (blue), for both Pandora [124] (squares) and PMTrack [89] (triangles).

relatively low. When particles with track lengths of less than 1 cm (10 cm) are removed, the average reconstruction efficiency for PMTrack rises to 37% (58%). This shows that, when the shortest tracks are not counted, the reconstruction performs reasonably well.

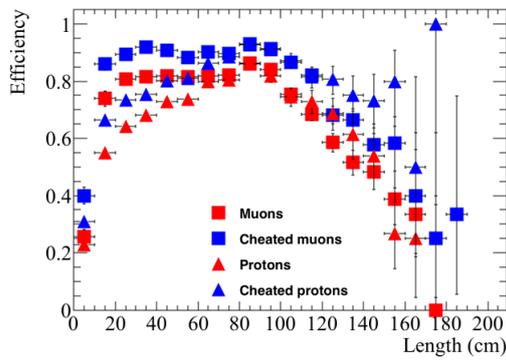
It is also useful to produce samples where the primary particle is a single muon or proton, located in the active volume of the detector. This allows a sample of isolated tracks to be made, upon which the capabilities of the PIDA metric can be tested. It also allows the reconstruction efficiency to be found for particles in isolation. The properties of the generated particles are illustrated in Table 4.2. The values of the simulated quantities were found by changing the given parameters by an amount taken from a random sampling of a Gaussian distribution of FWHM equal to the error listed. These simulation parameters were chosen to produce samples which would contain both exiting, and stopping particles, whilst generating the particles in the LAr would ensure that there should always be a reconstructable track in the detector. The reconstruction efficiencies when using the PMTrack reconstruction method are shown for the simulated particles in Figure 4.15.

Particles with track lengths of less than 1 cm have been excluded from these plots, which is why the angular reconstruction efficiencies for protons in Figures 4.15c and 4.15d, are higher than those seen in Figures 4.14c and 4.14d. This was done as none of these particles were reconstructed, due to the very short distances which they travel. After discounting these very short tracks, the efficiencies generally follow similar patterns observed in the earlier efficiency plots, though there is a decrease in efficiencies for the longest track lengths which is not observed in other samples. This is attributed to the initial positions of the particles being within the detector volume, as this means that any particle travelling over 100 cm would have a very peculiar trajectory, as the edge of the detector should never be more than 100 cm away from the starting position. The only exception to this, is if a particle travelled along the x axis to the other end of the detector. As discussed earlier, this is a very problematic orientation to reconstruct, as all of the charge would be deposited over a large range of time, on very few collection plane wires.

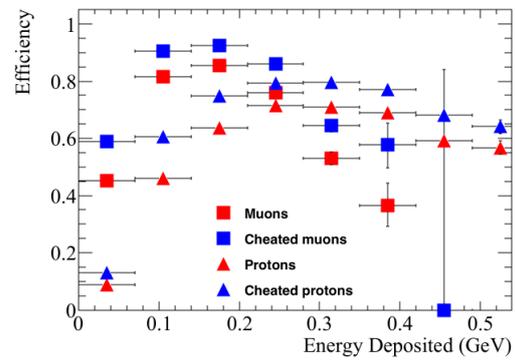
As the increase in $\frac{dE}{dx}$ is only visible when the particle stops in the detector, it is necessary to remove exiting particles from the sample. This is done by applying a fiducial cut on the end point of the reconstructed track. It is important to only place this on the end point of the track, as one does not want to remove particles which enter the detector and then stop. When calorimetry is performed, the end point of the track is determined using, among other metrics, the increase in $\frac{dE}{dx}$, and so the residual range of the track (a stored data member of

Table 4.2 The properties of initial particles simulated in the muon and proton samples. The angle θ_{xz} , is defined as the angle that a vector makes in the xz plane with the x axis. The angle θ_{yz} , is defined as the angle that a vector make in the yz plane with the z axis.

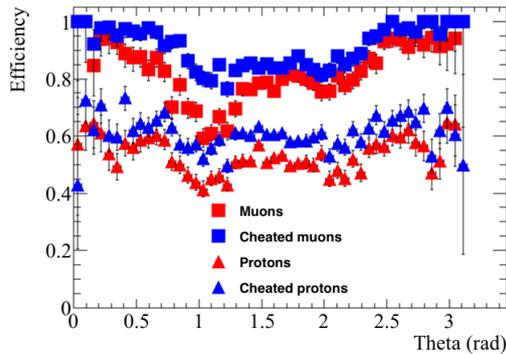
	Muon properties	Proton properties
Initial position (cm)	$(100 \pm 50, 0 \pm 30, 80 \pm 20)$	$(100 \pm 50, 0 \pm 30, 80 \pm 20)$
Initial momentum (GeV)	0.3 ± 0.1	0.8 ± 0.5
Initial θ_{xz} ($^\circ$)	0 ± 180	0 ± 180
Initial θ_{yz} ($^\circ$)	-45 ± 45	-45 ± 45



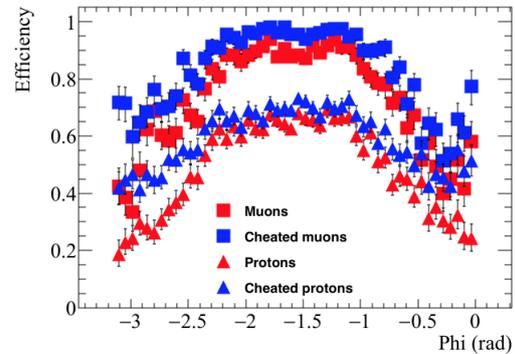
(a) The reconstruction efficiency as a function of the track length in the detector from Monte Carlo.



(b) The reconstruction efficiency as a function of the deposited energy in the detector from Monte Carlo truth.



(c) The reconstruction efficiency as a function of the θ track angle from Monte Carlo truth.



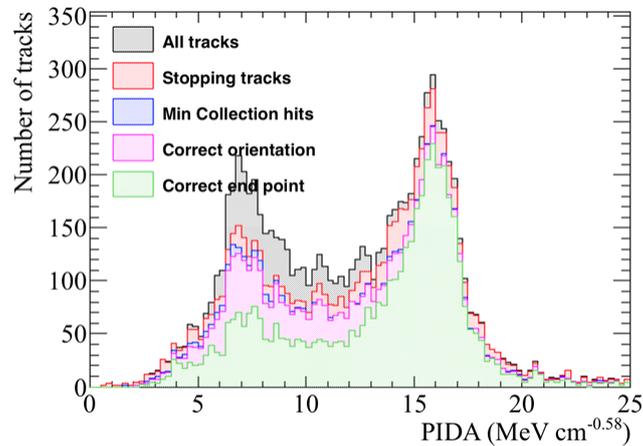
(d) The reconstruction efficiency as a function of the ϕ track angle from Monte Carlo truth.

Fig. 4.15 The reconstruction efficiencies for the simulated isolated muon and proton samples in the 35 ton detector. Top left: the efficiencies as a function of the track length in the detector from Monte Carlo truth. Top right: the efficiencies as a function of the deposited energy from Monte Carlo truth. Bottom left: the efficiencies as a function of the θ track angle from Monte Carlo truth. Bottom right: the efficiencies as a function of the ϕ track angle from Monte Carlo truth. The efficiencies are shown for 'non-cheated' reconstruction (red), and 'cheated' reconstruction (blue), for both Pandora [124] (squares) and PMTrack [89] (triangles).

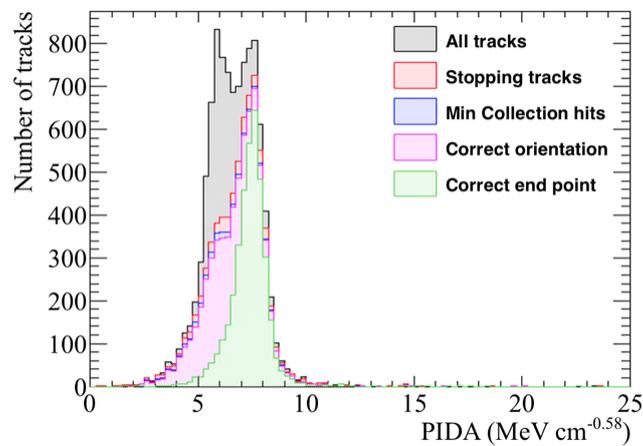
the track object) should always refer to the distance to the end of the particle's trajectory. For this study, a fiducial cut of 5 cm is used. This means that any track with hits within 5 cm of the edge of the detector volume is discarded, and counted as an exiting particle. This should mean that very few tracks due to exiting particles are identified as stopping in the detector, as it would require the reconstruction algorithms to miss a large section of the track. This will mean that some stopping particles are incorrectly assigned as exiting particles, causing the identification efficiency to drop, but it is necessary to ensure that exiting particles are not included in the final distributions. A further cut that is applied, is the requirement that the track contains a minimum of 10 collection plane hits, this is to ensure that an adequate number of points are taken upon which to find an average value of PIDA for the track. Similar cuts are described in [128], and the resulting distributions of PIDA values for the isolated muon, and isolated proton samples, are shown in Figure 4.16.

As can be seen from Figure 4.16, using Monte Carlo truth information can make the distributions much cleaner, particularly when discounting particles for which the reconstruction algorithms do not track to their end point. A track is identified as having a correctly reconstructed end point, if the reconstructed end point is within 2.5 cm of the end point of the particle from Monte Carlo truth. It is reassuring to see that few tracks are reconstructed backwards, as if this were not the case then performing particle identification would be very difficult. This is because, it would indicate that the calorimetry and tracking algorithms are not performing well. However, improvements can still be made, as both plots in Figure 4.16 contain many tracks which do not extend to the end points of the particles from Monte Carlo truth. This can be seen as when the requirement that the reconstructed track end point is consistent with the end point from Monte Carlo truth, the low tails of the PIDA distributions are significantly reduced. This is most noticeably the case in Figure 4.16b, where the peak at low values of PIDA is significantly reduced. It is observed that the PIDA distributions are cleaner when information from all three wire planes are used, as opposed to only using the collection plane, and so this is presented here. This shows how important it is to calibrate the electronics responses of all three wire planes, and how additional wire planes can improve calorimetry, as well as the accuracy of reconstruction algorithms [140].

Figure 4.17 shows the relationship between the $\frac{dE}{dx}$ and residual range of a track, for both protons and muons. The much steeper increase in $\frac{dE}{dx}$ at low residual ranges for protons, compared to muons, is clearly visible when comparing Figures 4.17a and 4.17b. The contamination in the proton sample at low PIDA can be seen in Figure 4.17a, where there is a clear sample of tracks for which the $\frac{dE}{dx}$ does not increase for low residual ranges. These plots are

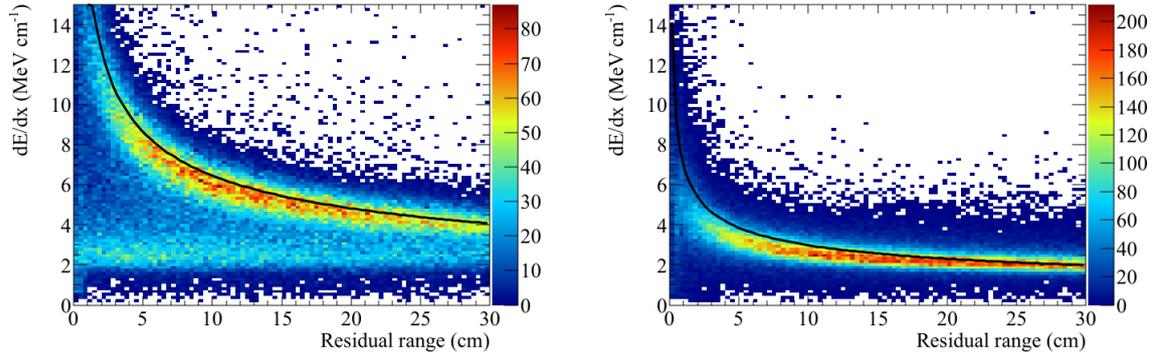


(a) The PIDA values calculated for the isolated proton sample.



(b) The PIDA values calculated for the isolated muon sample.

Fig. 4.16 The calculated PIDA values for the simulated isolated proton (top) and muon (bottom) samples in the 35 ton detector. A series of criteria designed to select only tracks due to stopping particles which have a required number of collection plane hits is applied. The tracks are then further refined using truth information such as the true end point of the particle.



(a) The $\frac{dE}{dx}$ versus residual range plot for the simulated isolated proton sample.

(b) The $\frac{dE}{dx}$ versus residual range plot for the simulated isolated muon sample.

Fig. 4.17 The measured relationship between $\frac{dE}{dx}$ and residual range for the simulated isolated proton (left) and muon (right) samples in the 35 ton detector. The black curves show the parameterisation given by the Bethe-Block equation (Equation 4.3 for each particle species). These plots are made after applying all of the cuts outlined in Figure 4.16, meaning that only hits from tracks whose end points are consistent with the end points from Monte Carlo truth are plotted.

filled after tracks whose end points do not correlate with the end points from Monte Carlo truth are removed, and so the tail of low $\frac{dE}{dx}$ values is due to particles for which the simulated detector did not find increased energy depositions as the particle stopped. It is therefore possible that at least some of these protons do not in fact stop, but interact inelastically when they still have a significant amount of kinetic energy. When this occurs, GEANT4 and the tracking algorithms, will create a new particle or track, which would be tracked to the true end point of the particle. However, the “end-point” of the initial particle will be before the particle actually stopped, and so it will have a MIP-like distribution in Figure 4.17a.

It is useful to summarise the information shown in Figure 4.16 in a table, so that an efficiency of identifying stopping particles can be found. This is shown in Table 4.3 for protons, and in Table 4.4 for muons. In these tables a “PIDA range” is defined which aims to contain as many of the reconstructed tracks as possible, whilst being roughly centred on the PIDA values shown in Figure 4.13. These PIDA ranges are 14-18 $\text{MeV}\cdot\text{cm}^{-0.58}$, and 5-9 $\text{MeV}\cdot\text{cm}^{-0.58}$, for protons and muons respectively. The efficiency shown in these tables is defined as the number of tracks in the PIDA range, divided by the total number of stopping particles. This means that it is possible to have an efficiency of more than 100%, if more reconstructed tracks have PIDA values within the PIDA range than there are stopping particles from Monte Carlo truth. This is the case in Table 4.4, where there are initially

Table 4.3 A summary of the PIDA values calculated for the simulated isolated proton sample, as sequential cuts are applied.

Applied cut	Proton sample			
	Tracks	In PIDA range	Efficiency	Purity
Total stopping particles	13295			
Reconstructed tracks	8761	3009	22.6%	98.7%
Survives 5 cm fiducial cut	7552	2894	21.8%	99.9%
Minimum of 10 collection plane hits	6186	2507	18.9%	99.9%
Correct track orientation	6022	2491	18.7%	99.9%
Correct tracking end point	4588	2327	17.5%	100%

more reconstructed tracks than stopping particles within the PIDA range. The purity shown in these tables is defined as the percentage of tracks in the PIDA range associated with particles that stop in the detector from Monte Carlo truth. As many of the tracks shown in the “reconstructed tracks” row in Table 4.4 are not due to stopping particles, the initial purity is low, though this increases markedly after the fiducial cut is applied.

As can be seen in Table 4.3, the efficiency with which protons can be identified does not change significantly as the sequential criteria are applied, but, as shown in Figure 4.16a, the peak at low values of PIDA decreases significantly. The same cannot be said for the muon sample however, as when the criteria that the tracking end point matches the end point from Monte Carlo truth is applied, a significant section of the tail within the PIDA range is removed, significantly reducing the PIDA efficiency. However, the resulting distribution is more similar to that shown in Figure 4.13, showing that the particles which survive the cut are those that are very well reconstructed. The cut to remove tracks that do not have the correct end points from Monte Carlo truth reduces both sets of efficiencies, but, if all particles were reconstructed with the correct end points, then one can imagine that the number of tracks within the PIDA ranges would increase, and the distributions would become more symmetrical, as shown in Figure 4.16b. Both tables also exhibit high purities, which shows that the fiducial cut, designed to removing exiting particles, is effective, with only 2 exiting protons being mis-identified in the proton sample.

From Table 4.3, it can be seen that there are more stopping protons than primary protons, as only 10,000 primary protons were generated. The effectiveness of the PIDA algorithm at identifying only primary protons is shown in Table 4.5. Comparing both tables, it can be seen that the efficiency with which the primary protons can be identified is larger than the

Table 4.4 A summary of the PIDA values calculated for the simulated isolated muon sample, as sequential cuts are applied.

Applied cut	Muon sample			
	Tracks	In PIDA range	Efficiency	Purity
Total stopping particles	6880			
Reconstructed tracks	9883	8907	129%	67.4%
Survives 5 cm fiducial cut	7126	6259	90.9%	90.2%
Minimum of 10 collection plane hits	6580	5876	85.4%	89.9%
Correct track orientation	6436	5767	83.8%	90.1%
Correct tracking end point	3832	3699	53.8%	100%

Table 4.5 A summary of the PIDA values calculated for the primary particles in the simulated isolated proton sample, as sequential cuts are applied.

Applied cut	Proton sample			
	Tracks	In PIDA range	Efficiency	Purity
Total stopping particles	7798			
Reconstructed tracks	5920	1937	24.8%	98.9%
Survives 5 cm fiducial cut	5044	1878	24.1%	99.9%
Minimum of 10 collection plane hits	4485	1711	21.9%	99.9%
Correct track orientation	4363	1707	21.9%	99.9%
Correct tracking end point	3246	1595	20.4%	100%

secondary protons, as the efficiencies shown in Table 4.3 are lower than those in Table 4.5. It is thought that this is due to the low reconstruction efficiency for particles with the shortest track lengths, as many of the secondary protons will have short track lengths in the detector from Monte Carlo truth, as discussed in Section 4.3. A similar table is not produced for primary muons, as there were no secondary muons produced in the isolated muon sample, and so Table 4.4 is itself the efficiency with which the primary muons can be identified.

Upon verifying that the PIDA metric can reliably determine particle type when they are simulated in isolation, the next step is to observe the accuracy with which particles can be identified in a CRY sample. The sample used here differs from the CRY sample used earlier, in that only events which contain a proton track in the detector from Monte Carlo truth are reconstructed. This is done to reduce simulation time and storage space, as this cut will still provide a substantial number of muons, whilst ensuring that a large proton sample can be

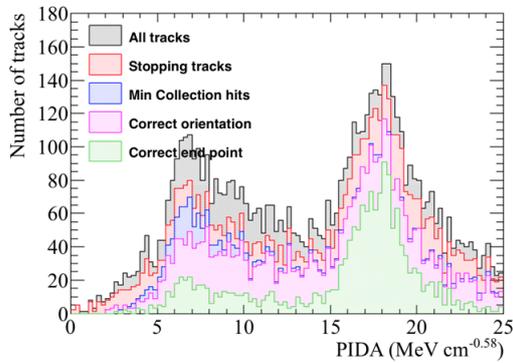
reconstructed. This sample is thus called the “proton enriched CRY sample’.”

The process of calculating PIDA values for tracks is identical in all samples, though as discussed in Section 4.3, the much more complicated event structure in the CRY sample affects the reconstruction efficiency, and so will likely also affect the accuracy of the calorimetry. The calorimetry will be affected in two ways. Firstly, the reduced performance of the reconstruction algorithms will mean that some particles are not reconstructed at all, whilst those that are reconstructed may be more likely to have missing hits, and so the end points may be reconstructed less accurately. This will cause the tail of low $\frac{dE}{dx}$ values which were seen in Figure 4.17a to be more pronounced. Secondly, though the photon detector time determination is very accurate for a large number of tracks, it is also incorrect for a number of tracks, as shown in Figure 4.4. This will cause the x position correction to be miscalculated, which will in turn increase the calculated $\frac{dE}{dx}$, and hence PIDA values. For this reason, it is also useful to present the PID efficiency when using cheated reconstruction, so that the effect of performing incorrect time determination and disambiguation can be seen.

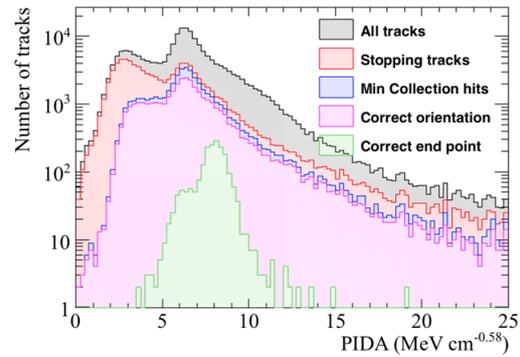
The PIDA values calculated for protons and muons in the proton enriched CRY sample are shown in Figure 4.18. The PIDA distributions when the reconstruction is both “cheated,” and not “cheated” are shown.

Figure 4.18 shows that the proton tracks which are reconstructed in the proton enriched sample are similar to those reconstructed in the isolated proton sample, shown in Figure 4.16a. This is reassuring, as it shows that even when protons are surrounded by multiple cosmic rays, the reconstruction algorithms are still able to accurately reconstruct proton tracks. The benefit of performing “cheated” reconstruction can be seen by comparing Figures 4.18a and 4.18c, where the number of proton tracks that survive the application of all cuts is seen to increase. When the reconstruction is performed using non-cheated disambiguation, and a cheated interaction time determination, the distribution of events surviving the application of all cuts is seen to be very similar to Figure 4.18c. This implies that the increase in the number of potentially identifiable proton tracks seen in Figure 4.18c, is largely due to the increased accuracy of interaction time determination.

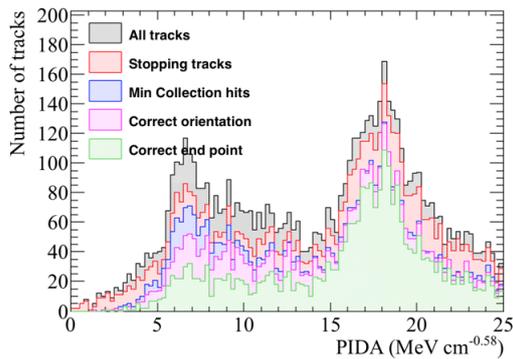
Though the plots shown in Figure 4.18 concerning the identification of protons are encouraging, the complementary plots concerning the identification of muons are much less encouraging. This is because, though there is quite a large peak at a PIDA value of around $8 \text{ MeV}\cdot\text{cm}^{-0.58}$ after the application of all cuts, this cannot be seen before the last cut is



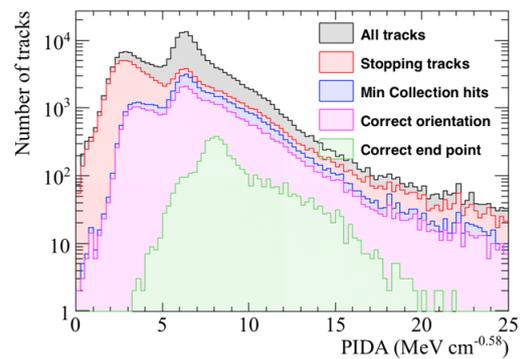
(a) The PIDA values calculated for protons, using the photon detectors to calculate an interaction time.



(b) The PIDA values calculated for muons, using the photon detectors to calculate an interaction time.



(c) The PIDA values calculated for protons, using cheated reconstruction.

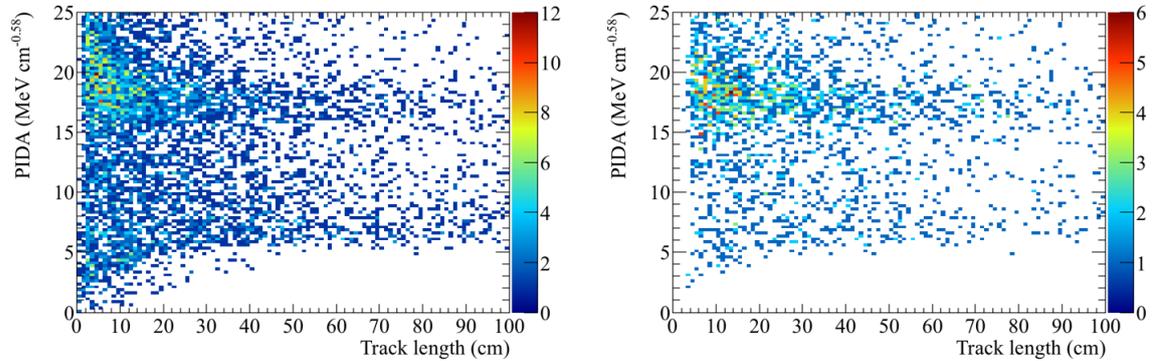


(d) The PIDA values calculated for muons, using cheated reconstruction.

Fig. 4.18 The calculated PIDA values for the simulated proton enriched sample in the 35 ton detector. Top left: the PIDA values calculated for protons using the photon detectors to measure the interaction time. Top right: the PIDA values calculated for muons using the photon detectors to measure the interaction time. Bottom left: PIDA values calculated for protons using a cheated interaction time determination. Bottom right: the PIDA values calculated for muons using a cheated interaction time determination. A series of criteria designed to select only tracks due to stopping particles which have a required number of collection plane hits is applied. The tracks are then further refined using truth information such as the true end point of the particle.

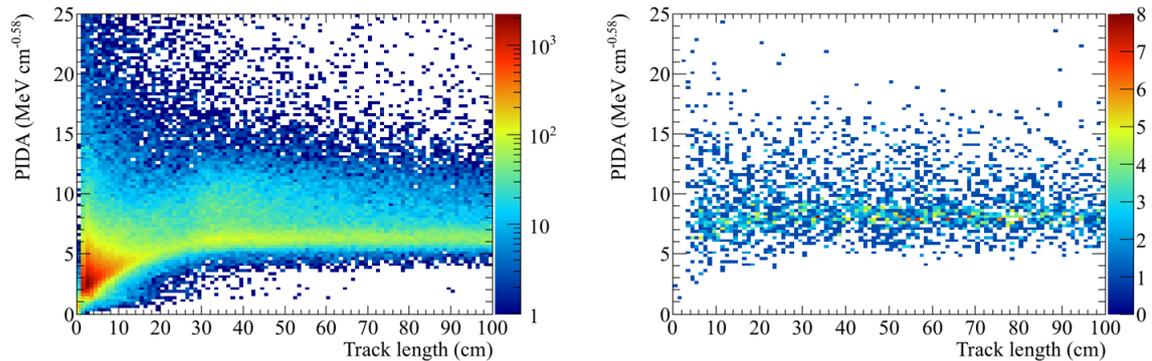
applied. This is because of two large peaks with PIDA values of around $3 \text{ MeV}\cdot\text{cm}^{-0.58}$ and $6 \text{ MeV}\cdot\text{cm}^{-0.58}$. A large peak at a PIDA value of $6 \text{ MeV}\cdot\text{cm}^{-0.58}$ was observed in Figure 4.16b, though this was largely removed by requiring that the muon track stopped in the detector. Unfortunately, this is not seen to be the case when considering muons in the enriched proton sample, though the size of the peak is significantly reduced by requiring that the track stops in the detector. Therefore, it is thought that the tracks which make up the peak of PIDA values at $6 \text{ MeV}\cdot\text{cm}^{-0.58}$, are partially reconstructed muon tracks. Effectively removing these partially reconstructed tracks is difficult, though if some degree of supplemental track stitching is performed during the analysis stage, they may be able to be removed. Evidence that this may work is presented in Figure 4.22, where through-going particles from Monte Carlo truth have been removed.

A feature of Figures 4.18b, and 4.18d, which was not present in Figure 4.16b, is the presence of a large peak at PIDA values of $3 \text{ MeV}\cdot\text{cm}^{-0.58}$. It is thought that this is due to very short delta rays coming off the high energy muon track as it passes through the detector. These tracks are considered to be due to the muon in the current framework, as the delta rays are not saved by GEANT4, and so LArSoft assigns any tracks which they produce to the parent of the electron, which in this case is the muon. Figures 4.19c, 4.20a, and 4.20b support this assessment. From Figure 4.19c, it can be seen that many of the tracks with very low values of PIDA have track lengths below 10 cm, which is much shorter than one would expect for a cosmic ray muon. This is conclusively shown by Figures 4.20a, and 4.20b, as it can be seen that the muon tracks which have low PIDA values have very short reconstructed track lengths, compared to the track length in the detector from Monte Carlo truth. Figure 4.20b then shows that many of the shortest reconstructed tracks are due to particles which have significantly longer true track lengths than that which has been reconstructed. Placing a cut at a minimum reconstructed track length of 10 cm will also remove some of the muon tracks which contaminate the range of PIDA values expected for protons. However, it can be seen from Figure 4.19b that removing all reconstructed tracks which are less than 10 cm in length will remove some of the reconstructed proton tracks which are within the expected PIDA value ranges. It may therefore be necessary to develop a more sophisticated cut to remove these delta rays if the identification of protons is attempted in the cosmic ray data collected by the 35 ton detector. Neither the development of such a cut, or the attempt to identify protons from the cosmic ray data collected by the 35 ton detector is presented in this thesis though.



(a) The calculated PIDA value for protons as a function of the reconstructed track length, filled for all tracks.

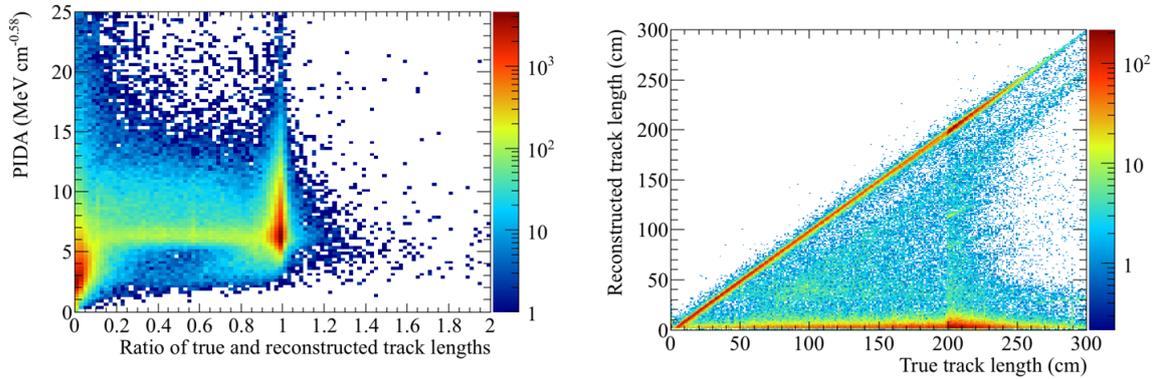
(b) The calculated PIDA value for protons as a function of the reconstructed track length, filled for tracks which pass all cuts.



(c) The calculated PIDA value for muons as a function of the reconstructed track length, filled for all tracks.

(d) The calculated PIDA value for muons as a function of the reconstructed track length, filled for tracks which pass all cuts.

Fig. 4.19 The calculated PIDA values, as a function of the reconstructed track length, for the simulated proton enriched sample in the 35 ton detector. Top left: the PIDA values, as a function of the reconstructed track length, when filled for all proton tracks. Top right: the PIDA values, as a function of the reconstructed track, filled for proton tracks which pass all cuts. Bottom left: the PIDA values, as a function of the reconstructed track length, when filled for all muon tracks. Top right: the PIDA values, as a function of the reconstructed track, filled for muon tracks which pass all cuts. All plots are made for tracks which have been reconstructed using a cheated interaction time determination.



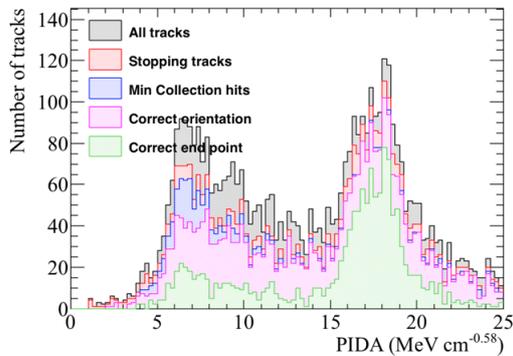
(a) The calculated PIDA value for muons, as a function of the ratio between reconstructed track length and the track length from Monte Carlo truth, filled for all tracks.

(b) The relationship between reconstructed track length, and the track length from Monte Carlo truth, filled for all tracks.

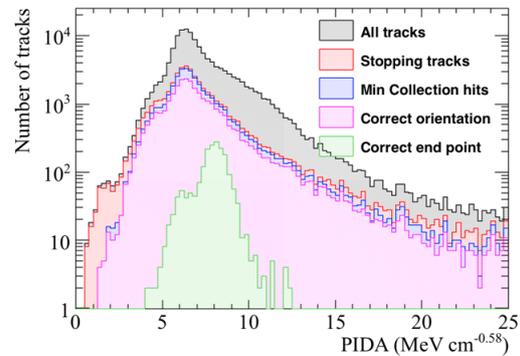
Fig. 4.20 The calculated PIDA values, as a function of the reconstructed track length, for muons in the simulated proton enriched sample in the 35 ton detector. Left: the PIDA values, as a function of the ratio between reconstructed track length, and the track length from Monte Carlo truth. Right, the relationship between reconstructed track length, and the track length from Monte Carlo truth. All plots are made for tracks which have been reconstructed using a cheated interaction time determination.

Figure 4.21, shows the distribution of calculated PIDA values for protons and muons in the proton enriched CRY sample, after a cut on the minimum reconstructed track length of 10 cm is applied.

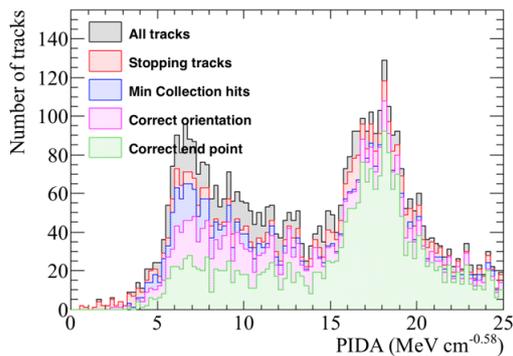
When comparing Figures 4.18, and 4.21, the removal of muon tracks with PIDA values of $3 \text{ MeV}\cdot\text{cm}^{-0.58}$ is very apparent, and shows that the cut on the minimum reconstructed track length was successful. It can also be seen that the number of muon tracks which have PIDA values consistent with a proton has also decreased by around a factor of 2. However, this has come at the cost of some of the proton tracks which could have been correctly identified as protons being removed, due to their short track lengths. This is unfortunate, but some loss is unavoidable, as the significantly more numerous muons would overwhelm any proton tracks, should the identification be attempted in cosmic ray data collected by the 35 ton detector. The peak around PIDA values of $8 \text{ MeV}\cdot\text{cm}^{-0.58}$ for muon tracks, is however, only identifiable once the requirement that the end point of the reconstructed track is close to the end point of the particle from Monte Carlo truth. Therefore, a more robust system of cuts is required in order to produce a sample of muon tracks which can be reliably identified as such.



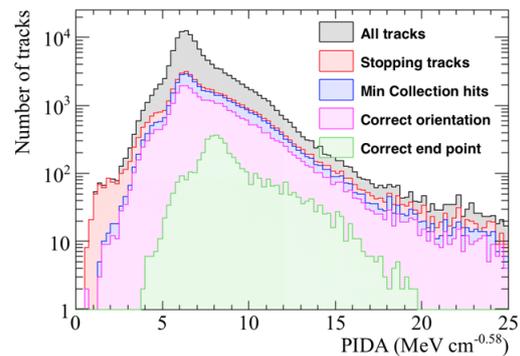
(a) The PIDA values calculated for protons, using the photon detectors to calculate an interaction time.



(b) The PIDA values calculated for muons, using the photon detectors to calculate an interaction time.



(c) The PIDA values calculated for protons, using a cheated interaction time determination.



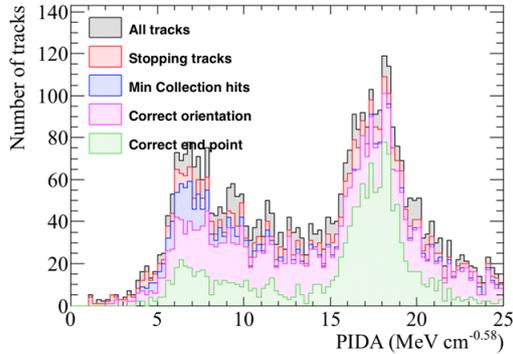
(d) The PIDA values calculated for muons, using a cheated interaction time determination.

Fig. 4.21 The calculated PIDA values for the simulated proton enriched sample in the 35 ton detector, after a cut on the minimum reconstructed track length is applied. Top left: the PIDA values calculated for protons using the photon detectors to measure the interaction time. Top right: the PIDA values calculated for muons using the photon detectors to measure the interaction time. Bottom left: PIDA values calculated for protons using a cheated interaction time determination. Bottom right: the PIDA values calculated for muons using a cheated interaction time determination. The cut on minimum track length is applied before any graphs are filled. A series of criteria designed to select only tracks due to stopping particles which have a required number of collection plane hits is applied. The tracks are then further refined using truth information such as the true end point of the particle.

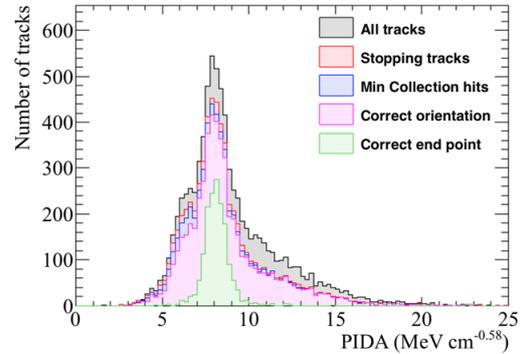
As outlined above, partially reconstructed muon tracks have PIDA values of $6 \text{ MeV}\cdot\text{cm}^{-0.58}$, and so in order to remove the large peak of these tracks still seen in Figure 4.21, some form of track stitching is required in order to remove these tracks. It is possible that a large number of these tracks are due to muons which pass through the detector, as many cosmic muons will do, but have been reconstructed as at least two separate tracks. Should this be the case, at least one of the tracks would be considered to “stop” in the detector. The effect of removing these split tracks can be seen in Figure 4.22, where all particles which do not stop in the detector, from Monte Carlo truth, have been removed. This has the effect of removing a large number of the reconstructed muon tracks, as most muons will not stop in the detector, but will instead pass through it. The application of this cut also effectively “cheats” the stitching which is required to remove the majority of split muon tracks, though any muons which stop in the detector, but have been split into two or more tracks by the reconstruction, will still be present in Figure 4.22. It can be seen that these tracks are still present, as there is still a peak in PIDA values of $6 \text{ MeV}\cdot\text{cm}^{-0.58}$, though these tracks are now overshadowed by tracks with PIDA values of $8 \text{ MeV}\cdot\text{cm}^{-0.58}$.

The number of muon tracks with high PIDA values in all plots is concerning, though this is particularly true for those shown in Figure 4.22, as these tracks are due to particles which stop in the detector, and are accurately reconstructed. This is because the contamination caused by these tracks, would mean that it would be difficult to ascertain whether a track with a PIDA value of around 18 is actually a proton.

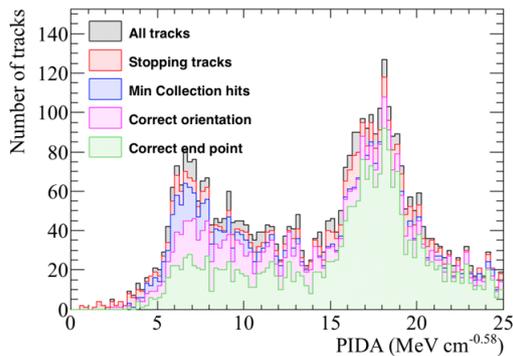
The basic framework by which particle identification can occur has been outlined here, and is shown to be very effective when particles are simulated in isolation. It has also been shown that a clear separation between muons and protons can be seen in cosmic rays, though there is a non-negligible contamination of the proton track PIDA range by muon tracks. This separation was only possible when using information from Monte Carlo truth though, and so a more sophisticated method of identifying through-going muons is required, should the analysis be performed on the cosmic ray data, such as that collected by the 35 ton detector.



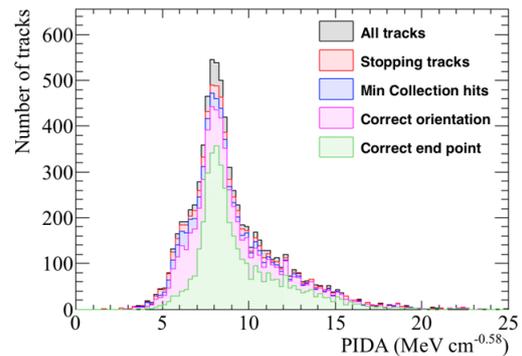
(a) The PIDA values calculated for protons, using the photon detectors to calculate an interaction time.



(b) The PIDA values calculated for muons, using the photon detectors to calculate an interaction time.



(c) The PIDA values calculated for protons, using a cheated interaction time determination.



(d) The PIDA values calculated for muons, using a cheated interaction time determination.

Fig. 4.22 The calculated PIDA values for the simulated proton enriched sample in the 35 ton detector, after tracks from particles which do not stop in the detector are removed. Top left: the PIDA values calculated for protons using the photon detectors to measure the interaction time. Top right: the PIDA values calculated for muons using the photon detectors to measure the interaction time. Bottom left: PIDA values calculated for protons using a cheated interaction time determination. Bottom right: the PIDA values calculated for muons using a cheated interaction time determination. The cut on minimum track length is applied before any graphs are filled, as is the cut on any reconstructed tracks which are associated with particles which do not stop in the detector, according to Monte Carlo truth. A series of criteria designed to select only tracks due to stopping particles which have a required number of collection plane hits is applied. The tracks are then further refined using truth information such as the true end point of the particle.

Chapter 5

The 35 ton data sample

The data taking period for the 35 ton prototype was from November 2015 until March 2016. This included an extensive commissioning period before the detector was filled with LAr, and the electric field was turned on. During this time many of the features of the data discussed below were first noticed, and attempts to rectify these were pursued. A long commissioning period was also required because many of the DAQ sub-systems were still under active development in November.

A total of 22 days worth of data was collected with the electric field set at $250 \text{ V}\cdot\text{cm}^{-1}$. The breakdown of when these periods occurred is shown in Figure 5.1. It is clear that the analysable data is interspersed with data where the electric field was not turned on, this is both due to extenuating circumstances such as a site wide power outage in early March, and a dedicated two week noise hunting exercise in February. The physics data taking period ended at 3am on 19th March 2016, when a filtration pump broke causing an unrecoverable loss of purity, as air was pumped into the detector. Following this, studies to understand the electronics noise, and to test the high voltage systems continued, but it was deemed too costly to acquire any more physics data. During this time the electric field was raised to the nominal value of $500 \text{ V}\cdot\text{cm}^{-1}$, and some of the causes of the higher than expected noise levels were discerned.

5.1 Organisation of the data structure

The 35 ton detector consisted of three detector sub-systems: Reconfigurable Computing Elements (RCEs) collecting Time Projection Chamber (TPC) data, SiPM Signal Processors (SSPs) collecting photon detector data, and Cosmic Ray Counters (CRCs) tagging cosmic

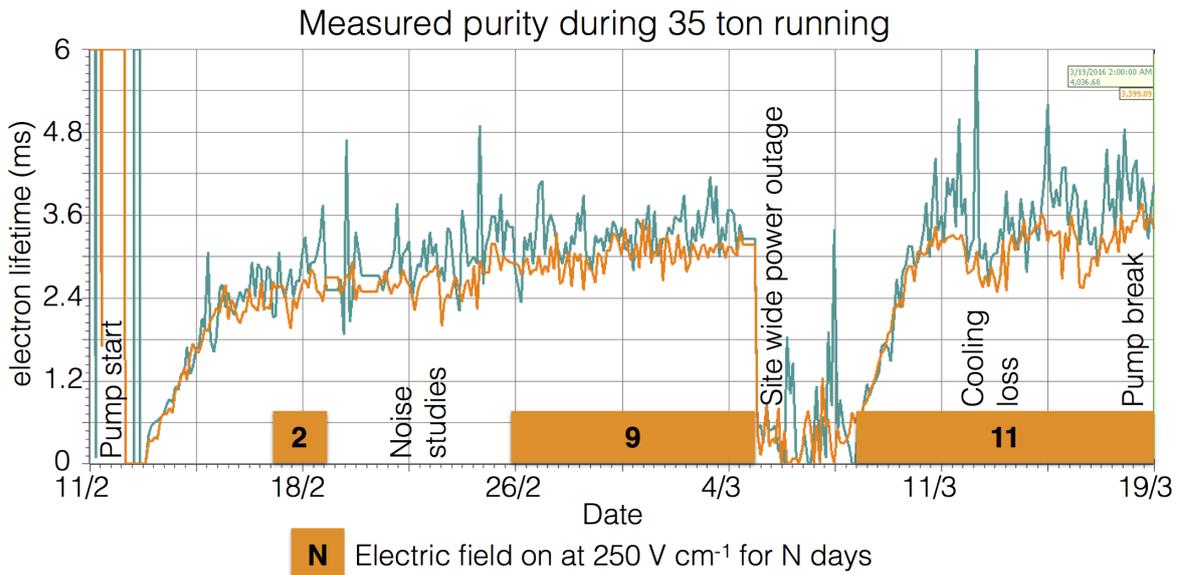


Fig. 5.1 Timeline showing the data collected during the 35 ton Phase II run once the purification pumps were turned on. The purity, as measured by two purity monitors at different heights in the detector, are shown by the green and orange curves.

rays. The DAQ combined these three data streams into synchronous events in time and saved them as LArSoft data objects. These data objects would later have to be converted to the offline data products, which the reconstruction tools developed on simulation used, this is discussed in Section 5.2. This section describes the structure of the data objects in the raw form.

During operations the DAQ was configured to maximise data throughput, and physics potential. This meant recording different lengths of times for each of the three sub-systems, as the data volumes and length of physics information were significantly different. For example, due to the emission of prompt light, the physics information from the SSPs is of a much shorter length of time than the physics information from the RCEs, where data has to be recorded whilst the electrons drift through the LAr. During the running period the recorded data was triggered by through-going muons which produced coincidences on the CRCs on opposite side of the cryostat. A coincidence is defined as two CRC modules recording a hit within 32 ns of each other. The system used to collect the CRC data was also responsible for generating the triggers, and so this meant that the trigger rate could be suppressed to approximately 1 Hz, by only producing triggers every N times a coincidence occurred, where N was a tuneable variable. A trigger rate of 1 Hz was used as the maximum speed at which data could be written to disk was approximately $60 \text{ MB}\cdot\text{s}^{-1}$, which is roughly equal to the

size of each triggered event when the entire detector is read-out in the configuration discussed below. The rate at which events were recorded could have been increased if zero-suppression of the TPC data had been used, however the noise level meant that this was not feasible.

With an electric field of $250 \text{ V}\cdot\text{cm}^{-1}$, and a drift of 223 cm, the drift time for electrons at the long drift Cathode Plane Assembly (CPA) was roughly 2.6 ms or 5200 ticks (where 1 tick is 500 ns). It was decided that in order for a track causing a counter coincidence to be separated from other tracks in the detector, it was necessary to have roughly one drift window both before, and after, the drift window around the coincidence. This means that data was recorded for 7.5 ms, or 15,000 ticks, around each coincidence. The SSPs only collected the prompt light from through-going particles, and so only $200 \mu\text{s}$ of SSP data was recorded for each event. The CRCs produced the least volume of data, and so were able to be read out constantly.

As the run mode required accessing buffered data, it had to be discretised inside the components before being sent to the event builders in the DAQ. In the discussion of how this worked, focus will be given on the RCE data, where some new terms need to be introduced. The smallest unit of data, called a nanoslice, is the data from one RCE for one tick, where each RCE controls 128 channels. A channel is defined as the signal which is read out from a single wire in the detector, this includes multiple wrapped wire segments for the induction wires. There were a total of 16 RCEs in the 35 ton detector, reading out 2048 channels. A microslice is then made by combining $1000 \times N$ nanoslices such that it contains 0.5 ms (1,000 ticks) of data across all channels, where N is the number of RCEs that are recorded in the run. Microslices are then combined to make millislices, the length of which was configurable. Once produced, these millislices were sent by the DAQ to the event builders, to be stored as time synchronous LArSoft data objects.

The time synchronous events produced by the DAQ, did not, however, correspond to the physics events. This is because the DAQ was originally designed to produce a continuous data stream. This meant that the DAQ was configured to pad events with headers when a sub-system provided no physics information, such as nanoslices in the case of the RCEs. Removing these padded header objects was a remit of the online to offline converter discussed in Section 5.2. The length of the millislices was configurable, and was chosen to be 10 ms (20,000 ticks) in order to best attempt to fully contain physics events, and reduce the need for the online to offline converter to stitch DAQ events together. The padding of millislices with headers between physics events introduced some peculiarities in the recorded data, such as

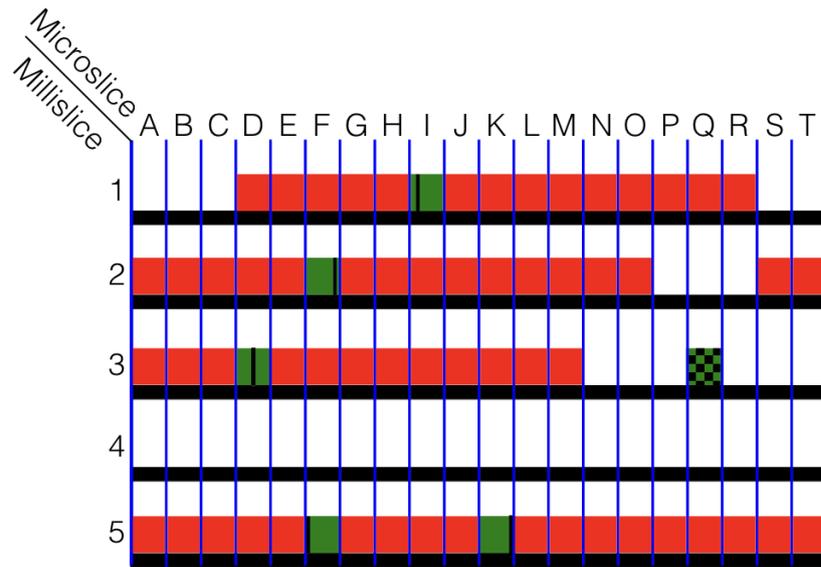


Fig. 5.2 A diagram of possible millislice structures for the TPC data recorded by the 35 ton detector. Each row represents a millislice (numbered 1 to 5), whilst each box represents a microslice (labelled A to T). The vertical blue lines delineate each microslice, giving 20 microslices per millislice. Solid red and green boxes represent microslices with TPC data in them. A group of 15 continuous red and green boxes are the recorded “physics events”. Green boxes represent triggers which were used, with the black lines showing the time in the millislice at which the trigger occurred. The green and black patterned box in microslice 3Q represents a coincidence which was not issued as a trigger. A possible reason for this trigger not being issued, is its proximity to a previous coincidence trigger which was issued from the same co-incidence.

millislices containing two parts of non-continuous data. This is shown in Figure 5.2, where the second millislice has no information for the time between the end of physics event 2 (2O), and the start of physics event 3 (2S).

During normal data taking microslices were buffered in the RCEs, so that if a trigger was issued they could be accessed before being deleted. As the data was buffered in the form of microslices, previous microslices could only be accessed as a whole. This meant that a whole number of microslices had to be loaded before the trigger, so when a trigger was issued part way through a microslice, the previous X microslices were sent to the event builders, where X was typically 5. As a result, there are always a minimum number of ticks both before (5,000 ticks) and after (9,000 ticks) the trigger, but the exact numbers can change by up to 1,000 ticks for a given event, depending on where in the microslice the trigger came. The

result of this is that it is impossible to know the number of ticks before/after a given counter coincidence. This is shown in Figure 5.2 where the black lines representing triggers, are seen to occur at different points within the microslices. For example, physics event 1 (1D to 1R) will have more data after the trigger than physics event 2 (2A to 2O), as the trigger occurred earlier in the triggered microslice.

5.2 Reformatting the data to the offline structure

Conversion of the data objects stored in the raw data to the data objects used in simulation required a suite of unpacking services to be written, the specifics of which are not discussed here. These all required a common interface through which to access the data, and check that the timing of each component was consistent, so that a final LArSoft file for downstream use could be produced. This interface had the added role of producing complete physics events, meaning that it had to be able to combine multiple millislices, and extract only the data containing the continuous physics events.

Following the unpacking of each of the sub-systems, the data reformatter would loop through the TPC ticks to see if a user defined set of conditions could be satisfied at that time. These conditions were usually whether an east-west or north-south counter coincidence (see Figure 2.16) occurred at that time, or if this millislice contained TPC data whilst the previous one did not. The latter was the default configuration, as this gave the option of preserving all of the data gathered, for reasons discussed at the end of Section 5.1. Other conditions were available, though rarely used, such as if the SSPs observed a large flash of light, or if there was a large change in the average TPC Analogue-to-Digital Converter (ADC) value. Once a set of conditions are satisfied, a user defined number of pre-condition ticks are gathered. No pre-condition ticks are gathered when the previous millislice contains no TPC data, as there is no previous data to load which would not have a gap in time, see Figure 5.2, millislice 5. In the case of using a counter coincidence to make an event, a value of 300 pre-condition ticks is normally used, with a maximum of 5000 ticks being able to reliably collected. Once the pre-condition ticks are gathered, a further N post-condition ticks are gathered, where N is defined by the user. Usually 15,000 ticks are gathered when the previous millislice is empty, and 5,200 ticks are gathered when there is a coincidence, though a maximum of 9,000 ticks could be reliably gathered. Data from the other components is added to the event if its timestamp is within the timestamps of the first and last ticks in the event. This is done either, when no more TPC data is required, or at the end of a millislice if stitching is required. All timestamps are corrected such that the event began at $t = 0$, as the reconstruction assumes

this, and the timestamp of the start of the event is stored in the event record so that it can be accessed later if required.

It is important to integrate flexibility at all points in this process, so that the user can choose the length of events, which sub-systems are in the events, and what the conditions are for making events. It is also important for users to be able to run the service on already formatted events, as the unpacking services are the major overhead in running the interface. It is also conceivable that users would want to reformat Monte Carlo events to centre them around simulated counter coincidences, and so the use of the unpacking algorithms was determined by the interface depending on the format of the input file.

5.3 Observations on data quality and noise mitigation

Reformatting the online data to the offline format was an important step in maintaining data quality, as subsequently there was no access to the raw data due to the framework of the 35 ton software. Some of the important checks which are performed are outlined in Figure 5.3. If any of these issues are present in a given physics event it is discarded as the integrity of the data cannot be guaranteed. It was decided that these events would be discarded as non-synchronous events would lead to hits in the detector being at incorrect times, and padding empty events with pedestals could mean that tracks seem to disappear as they travel through the detector, only to reappear at a later time in a different detector location.

Another example of an inconsistent event is when the sub-systems are not synchronised with each other. This is normally caused by one of the sub-systems missing a clock increment from the master timing unit, due to the data trigger being issued close to an increment from the master unit. This misalignment causes an incorrect time sample being read out, and so the data from each sub-system within a millisecond is not consistent. The result of this is that the event will fail the timestamp check, and so won't be added to the event record. To avoid incomplete events, these physics events are also discarded when observed.

The electronic noise in the 35 ton was higher than anticipated, with the RMS of the RCE ADC being approximately 30 counts, compared to an expected thermal noise of around 2.5 ADC counts. As the electronics noise was so large it was not possible to ascertain the levels of backgrounds from other sources such as radiological decays. Many sources contributed to this elevated electronics noise, some of which are explained below.

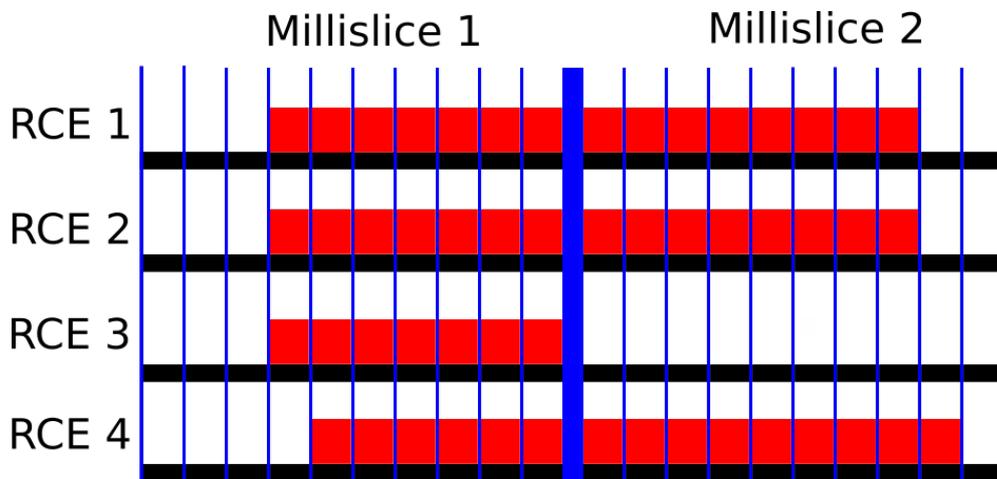


Fig. 5.3 A diagram of TPC microslices within millislices in the 35 ton data stream. Two millislices are shown, each containing 10 microslices. One physics event straddling the millislice boundaries is shown, and 4 RCEs representing each row are read out. The vertical blue lines delineate each microslice (0.5 ms, 1,000 ticks), with the thick blue line showing the millislice boundary. Solid red boxes represent microslices with TPC data in them. It can be seen that RCEs 1 and 2 contain data for the same interval, whilst the data from RCE 3 in millislice 2 has been “Dropped,” and the data from RCE 4 is shifted by 1 microslice from RCEs 1 and 2 and is thus “Inconsistent.” As a result of these issues this physics event would be discarded, as data integrity cannot be guaranteed.

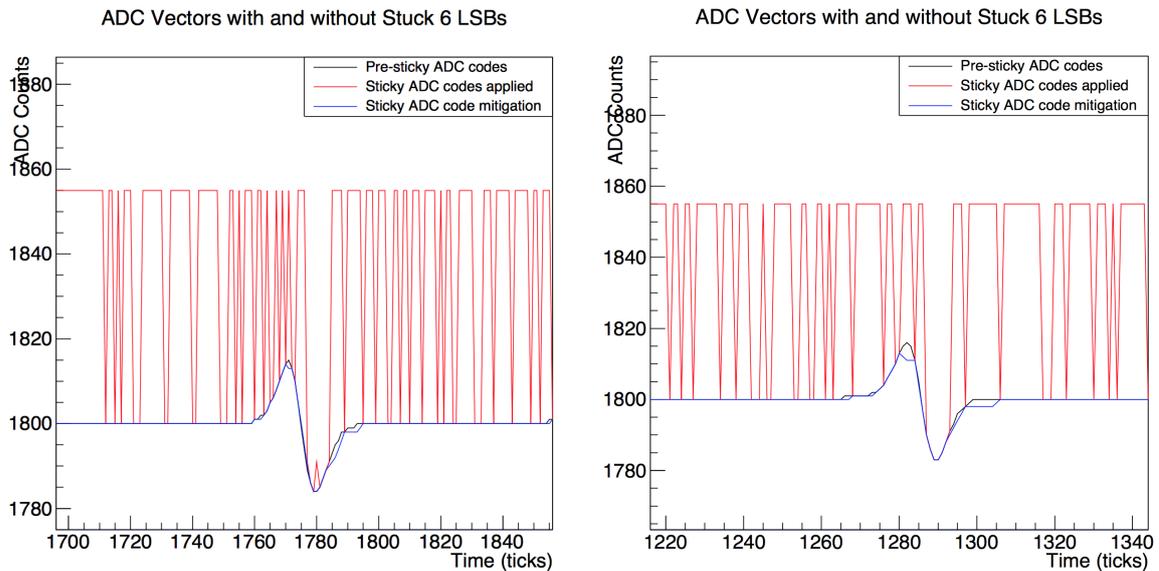


Fig. 5.4 Two Monte Carlo spectra showing the effect of the introduction and removal of stuck bits on a simulated signal. The black line shows the simulated signal on a wire, which is then modified by adding the effects of “stuck ADC codes,” shown by the red line. The “stuck ADC codes” are then removed, and the resulting signal is given by the blue line. It can be seen that the signal loss is minimal after the “stuck ADC codes” are removed. The figures were taken from [141].

Though not directly affecting the noise issues “stuck ADC codes” were a feature of the data which had to be removed. “Stuck ADC codes” were caused by bit level corruption where the lowest 6 bits in the ADC became frozen to either $0x0$ or $0x3f$. This was observed during the first stages of commissioning, and an algorithm to remove them was developed and tested on Monte Carlo [141]. In simulations it was observed that the signal could be recovered with minimal losses, as shown in Figure 5.4, where the signals after stuck code removal (blue lines) are seen to closely match the signals before stuck codes were added (black lines).

A significant portion of the noise was correlated between groups of 32 channels, where the ADCs would coherently oscillate. To remove these coherent shifts, ADC baselines were calculated for these groups of 32 channels at each tick, and then subtracted from the measured ADC values. This was found to be an effective method of removing coherent noise in MicroBooNE [142]. The effect of removing coherent noise is shown in Figure 5.5, where the signal peak becomes much easier to discern after noise removal, and a coherent noise peak around tick 6030 is removed. An issue with removing coherent noise in this way is that events which are parallel to the Anode Plane Assembly (APA) frames will produce signals at common times, and across adjacent wires, and these signals may be removed along with the

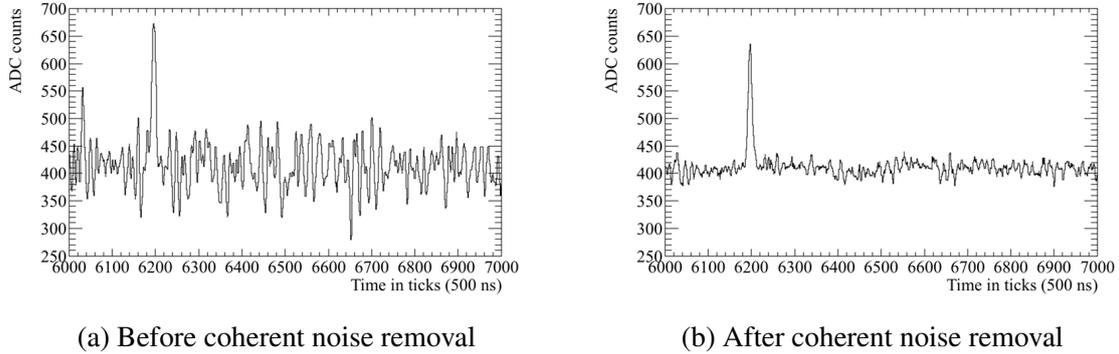


Fig. 5.5 The effect of coherent noise removal on a 35 ton signal event. Left: the signal before coherent noise is removed. Right: the signal after the coherent is removed. The signal peak around tick 6200 is much clearer after coherent noise removal, meaning that hit reconstruction becomes much simpler.

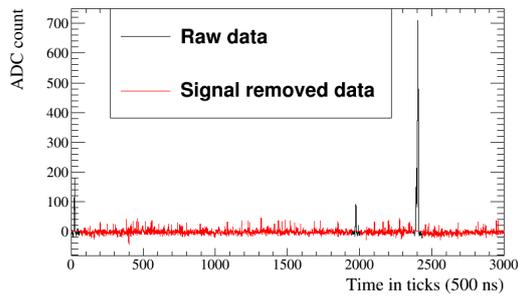
coherent noise. This will cause a reduction in the hit reconstruction efficiency. The only way to prevent this is to “protect” potential signal regions from the coherent noise removal, as is done in MicroBooNE [142].

After performing a Fast Fourier Transform (FFT) [143] on the coherent noise subtracted waveforms, it can be seen that signals occur with specific frequencies. Some of these frequencies are caused by real energy depositions, whilst others are due to the electronics noise. It is possible to remove the noise frequencies by applying Wiener filters [144]. Frequency spectra are taken for each of the three planes, and a clear signal is both preserved and suppressed. The raw signal spectra are then divided by the signal suppressed spectra, to produce *signal/noise* frequency spaces. The regions of frequency space to be conserved, given by regions of high *signal/noise*, can then be found by fitting a combination of sigmoid functions to the frequency spaces. Equations 5.1, and 5.2 show the sigmoid functions which are applied to the collection and induction planes respectively.

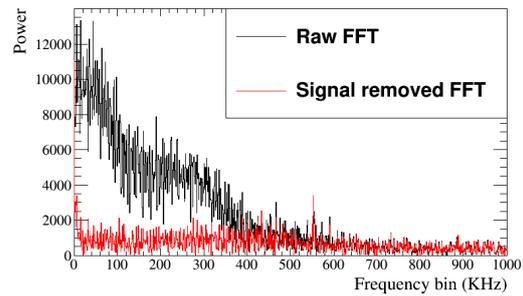
$$f(x)_{Col} = 1 - \frac{1}{1 + \exp(-(x - A_{Col})/B_{Col})} \quad (5.1)$$

$$f(x)_{Ind} = \left[1 - \frac{1}{1 + \exp(-(x - A_{Ind})/B_{Ind})} \right] \times \left[\frac{1}{1 + \exp(-(x - C_{Ind})/D_{Ind})} \right] \quad (5.2)$$

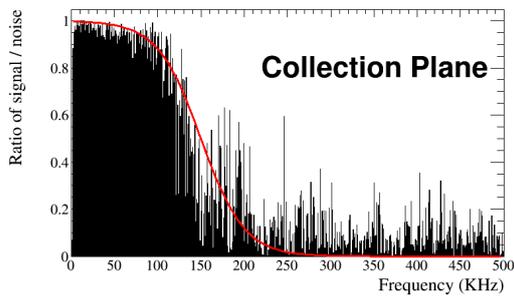
where A_{Col} , B_{Col} , A_{Ind} , B_{Ind} , C_{Ind} and D_{Ind} are tuned for each plane. A demonstration of how this was applied is shown in Figure 5.6. It is also possible to remove specific frequencies which are not removed by the filters, this was necessary for a 54 KHz noise component which was introduced by the fluorescent lights in the detector hall. After the run ended it was found



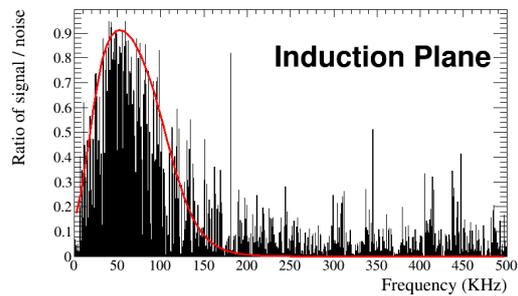
(a) A raw and signal subtracted waveform for a collection plane wire.



(b) The FFT of the raw and signal subtracted waveform for a collection plane wire.



(c) The *signal/noise* ratio for a collection plane wire, the red line shows the fraction of frequency power which passes the filter.



(d) The *signal/noise* ratio for an induction plane wire, the red line shows the fraction of frequency power which passes the filter.

Fig. 5.6 The application of Wiener filters to the 35 ton data. Top left: a waveform from a collection plane wire which is then signal suppressed. Top right: the FFT of both the raw and signal suppressed waveforms. Bottom left: the *signal/noise* ratio for the collection plane waveform, a sigmoid function has been overlaid to preserve the areas of high *signal/noise*, and suppress the areas of low *signal/noise*. Bottom right: the *signal/noise* frequency space ratio for an induction plane wire, a sigmoid function has been overlaid to preserve only the areas of high *signal/noise*.

that some of the high frequency noise components were introduced by a short on a warm power cable. The techniques used to find this cable will be used when commissioning future detectors [145].

An example of the effect of the noise mitigation steps is shown in Figure 5.7. Figure 5.7a shows the raw data, whilst Figure 5.7b shows the data after the stuck code unsticker, coherent noise removal and Wiener filter algorithms have been applied. The effect of noise removal is clear, as the signals from the tracks become much more pronounced, particularly on the bottom induction plane. However, it can also be seen that the noise removal algorithms also remove signals from tracks, as the depositions seen on the collection plane in Figure 5.7a at

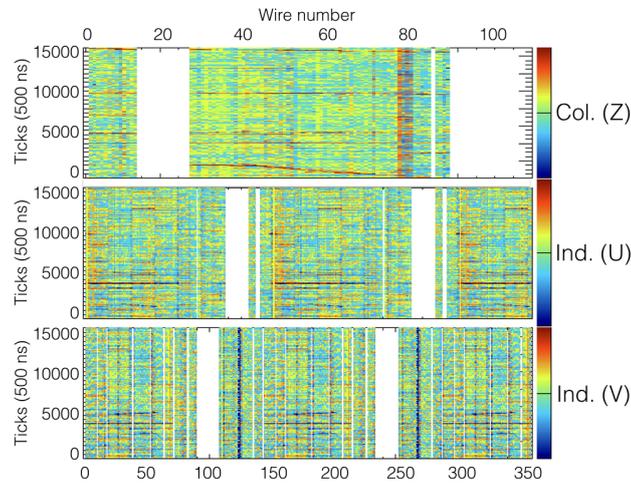
around tick 10,000 becomes much less pronounced in Figure 5.7b.

Transitions to a higher noise state were observed after cool down, this was associated with strong signals at frequencies between 400 and 650 KHz. The transitions would occur approximately every 2 hours, and were occasionally observed to happen shortly after a saturation event across the whole detector [145]. Once the state was induced, the only way to stop it was to power cycle the low voltage supplies. It was found that power cycling the short APA at the base of the detector (APA2) could both stop, and induce the higher noise state. Importantly, this was the only APA with electronics located at the base of the TPC. The data taken during the elevated noise state was unrecoverable, as the electronics noise was too large, and so upon the observation of a transition, the low voltage supplies were power cycled. It was observed that the transitions occurred much less frequently when APA2 was not powered, and so it was not used for significant portions of the data taking period. Despite efforts to study the transitions during warm testing they were unable to be induced, and have not been observed in other experiments such as MicroBooNE, despite the same low voltage supplies being used. It is thought that the cause of the transitions is a feedback loop in the low voltage cable, which was much longer in the 35 ton than in MicroBooNE. This would explain why APA2 was more susceptible to the feedback loop, as the cable is routed past its electronics [146].

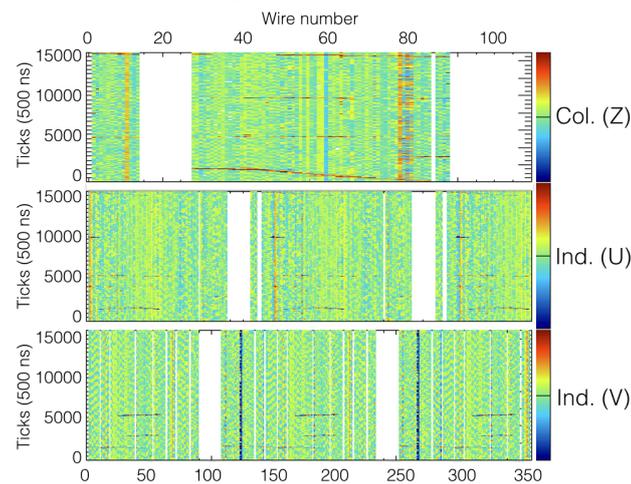
5.4 Performance of reconstruction algorithms

After performing the noise removal which was outlined in Section 5.3, hit and track finding was still more difficult than in simulations, due to the elevated noise level. In order for a reasonable number of hits to be reconstructed the hit finding threshold had to be substantially increased in data, as compared to Monte Carlo. This meant that many of the low energy hits would not be reconstructed, as will be shown in Figure 5.8.

A potential solution to not reconstructing the low energy hits, is to use the counter positions to select only hits which could have caused coincidences. When determining whether a reconstructed hit could have caused the counter coincidence, a two-dimensional window around the counter edges in the yz plane is constructed, and timing information is used to extend this to three dimensions. The x position of the hit can be calculated using the hit time, and electron drift velocity using Equation 4.1.



(a) Raw signal before noise removal



(b) Signal after noise removal

Fig. 5.7 Event displays showing the effect of the noise removal algorithms on data from the 35 ton detector. The event displays show the signals in the collection (Z), and induction planes (U and V). The plots show wire number, and time in ticks, on the x and y axes respectively. The charge is shown on the colour axis, where areas in red represent areas of high charge deposition, whilst blue represents areas of low charge deposition, areas of white show wires which did not read out any signals. The effect of the noise removal algorithms can clearly be seen, as large changes in charge due to the noise are no longer present after they have been applied. The application of the noise removal algorithms does however remove real signals, as the number of depositions across many channels around tick 10,000 on the collection plane are noticeably reduced after they are applied.

Determining whether collection plane hits are within the counter window is trivial as they have a constant z position, and either cover the full detector height (tall APAs), or roughly half of the detector height (short APAs). However, the wrapping of the induction planes, means that each wire segment has to be considered individually, and that multiple segments of a given wire could lie within the counter shadow. The 3-dimensional volume that is enclosed by connecting the edges of the counters which were hit in the counter coincidence, is called the “counter shadow.” Only those wires which lie within the 2-dimensional projection of this volume onto the yz plane, are considered here. Choosing between these potential wire segments is done by iterating through the following steps. If at any point only one segment satisfies a given condition, then that segment is chosen:

- Does the wire segment intersect any collection plane wires which record hits?
 - This is because when there is a signal on an induction plane there should also be signals on the collection wires.
- Are there adjacent wires which have hits at a similar time?
 - This is because one would expect a track to deposit energy on multiple adjacent wire segments.
- Which hit lies closest to the line defined by unique collection plane hits in the xz plane?
 - This follows identical logic to the first criterion, but selects the hit which best matches the collection plane hits, and attempts to remove the effect of noisy collection plane wires by only using wires which have one hit within the counter shadow. This would also hopefully improve the quality of the fit, as there will not be numerous outlying hits.
 - This can be changed to consider the line defined by previously selected hits in the given TPC and plane where the hit choices are.

Following a re-optimisation of the clustering algorithms, it was observed that the standard reconstruction could achieve track reconstruction to a similar efficiency as the counter shadowing method, and so the counter shadowing method has not been used in the discussions to follow [147]. There has since been an effort to improve the counter shadowing hit disambiguation to remove the outlying collection plane hits using the MLESAC method [148], whereby points which are far away from a best fit are ignored. These studies are still ongoing [149].

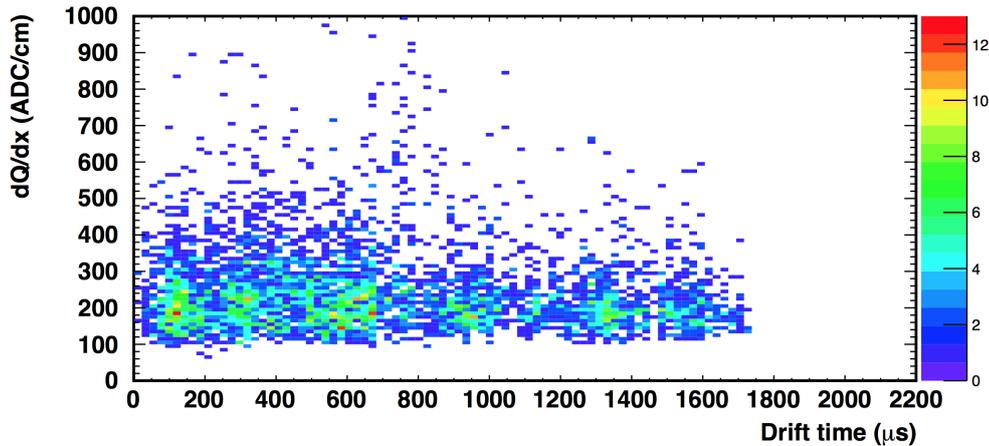


Fig. 5.8 The dQ/dx values for a sample of muon collection plane hits, note the cutoff at $100 \text{ ADC}\cdot\text{cm}^{-1}$ due to the hit finding threshold. The figure is taken from [151].

A symptom of the elevated noise state is that signals are often dropped on one of the induction planes, this means that the tracking algorithms often have to combine clusters in only two of the three planes. Reconstruction using two planes was shown to be effective by the ArgoNeuT collaboration [150], so the loss of signal in one of the three planes is not prohibitive to track reconstruction. Another consequence of the elevated noise level is that even when the counters are used to seed hit finding, the hit finding threshold is too high to reconstruct the very lowest hits. This causes the plot of dQ/dx for muons, shown in Figure 5.8, to look flat, due to a cutoff at $100 \text{ ADC}\cdot\text{cm}^{-1}$, below which no hits are reconstructed. The inability to reconstruct the lowest energy hits means that calorimetry is all but impossible on the 35 ton dataset, even though the tracking algorithms perform relatively well. The inability to perform reliable calorimetry en masse means that the only particles which can be assuredly identified are the muons which triggered the counter coincidences. This means that performing the analysis that was presented in Section 4.4 on the 35 ton dataset is extremely difficult, if not impossible.

The muons in the triggered sample will all traverse the detector, but their orientations can be carefully selected by the user. For example, one could easily select a sample of muons which cross the APAs at increasing angles, or are parallel to the wire planes at increasing drift distances. This is done by matching through-going muons with counter coincidences. The process by which this is done is identical for both north-south and east-west coincidences, though more focus will be given to the later, as it is with muons of this orientation that the analysis presented in Section 5.5 was performed. The same matching technique would also have been applied to vertical muons had the telescope trigger been utilised. For a reference as

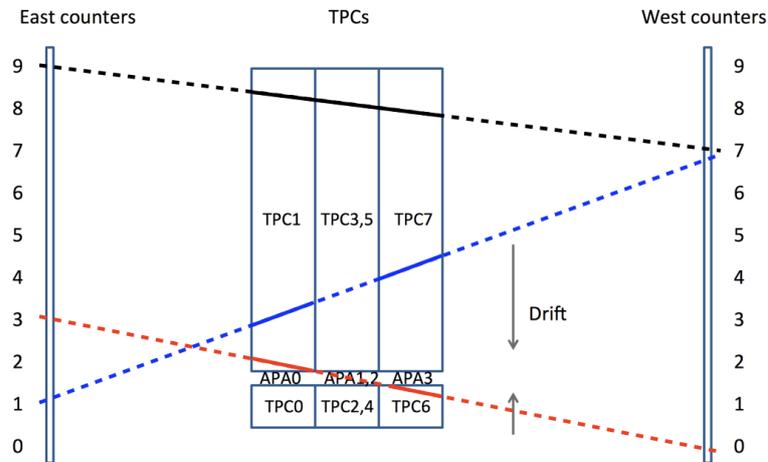


Fig. 5.9 The numbering scheme for the east - west counters in the 35 ton. The counters have been numbered from 0 to 9 depending on their position from the end of the short drift volume. This is different to the LArSoft numbering scheme shown in Figure 2.16 where they go from 6-15 and 28-37 for the east and west counters respectively. Three hypothetical muons which would have caused coincidence triggers are shown as dashed lines, and the hypothetical reconstructed tracks they produce are shown as solid lines. The red particle is an APA crossing event, and would produce tracks in TPCs 1 and 6. The black particle is fully reconstructed as one continuous track, however the blue particle is not reconstructed in the middle TPCs and so is reconstructed as two separate tracks.

to the locations of the counters around the cryostat, see Figure 2.16, and for a representation of only the east-west counters, see Figure 5.9.

It is possible to construct a line in the yz plane joining the centres of the two counters which were hit when a coincidence occurred. This can then be compared with the trajectory of a track in the yz plane, and the dot product of the two vectors calculated. A reconstructed track is assigned to a given counter coincidence if the dot product of the track and the coincidence is more than 0.98, and the hit times are consistent with the x positions of the counters. The results of the dot product calculation are shown in Figure 5.10. Matching only tracks which are well aligned with a counter coincidence should produce a pure sample of tracks, as parallel muons are unlikely to be highly correlated in both time and space, and any tracks reconstructed from the noise will have random directions. This is clearly apparent in Figure 5.11a where the track that is parallel with the coincidence does not have consistent z positions. Tracks which are parallel with an East-West coincidence can be excluded by considering whether they are consistent with the x positions of the counters which were hit.

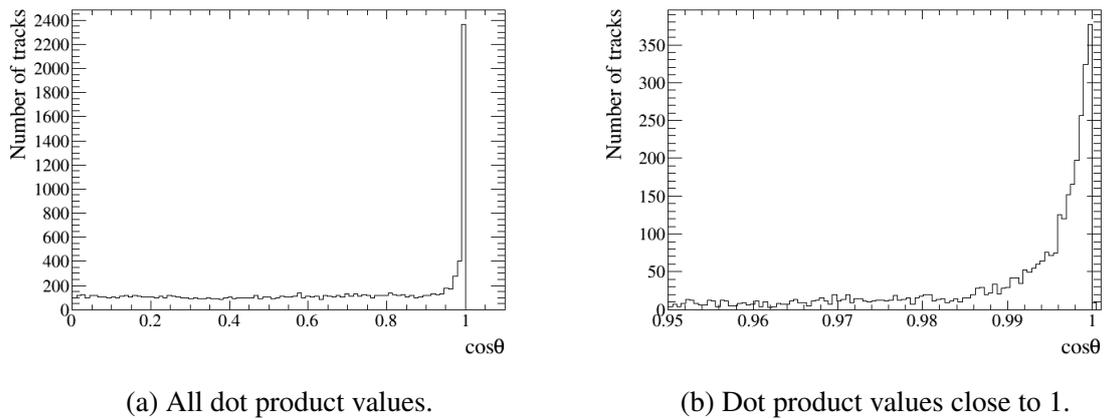


Fig. 5.10 The dot product of the track and vector joining the centres of the coincidence counters in the yz plane. The number of tracks with a given dot product is plotted on the y axis. A threshold value of 0.98 is required for a track to be considered to be due to the counter coincidence. It can be seen that many tracks are well aligned with counter coincidences, having dot products of more than 0.99.

This means one of the tracks which are parallel with the dashed line in Figure 5.11 b may be removed when the x positions of the tracks are considered.

By matching tracks in this way it is possible to evaluate the reconstruction efficiencies for these muons, at increasing drift distances and track angles. If multiple tracks are aligned with the coincidence, and are within the expected time region, then their track lengths are summed when calculating reconstruction efficiencies. When this occurs, it is expected that the track was split by a region of the detector either being turned off, or being too noisy to reconstruct a track. If these tracks have a combined track length of more than 50 cm, then the coincidence is identified as having been successfully reconstructed. This threshold is much lower than the true track length which should be reconstructed (more than 150 cm), but few particles are fully reconstructed in the data, and so a compromise is made to achieve a large enough sample of tracks upon which analyses can be performed. A reconstructed track that is 50 cm long is likely to have a large number of hits on collection plane wires that are not noisy, and it is these hits which are required when calculating purity or measuring the effect of diffusion, as discussed in Section 5.5. A track with length more than 50 cm is also likely to have been stitched between TPCs, due to the geometry of the 35 ton and track trajectories. The demonstration of stitching tracks between TPCs was a design goal of the 35 ton, and so identifying tracks where this was achieved satisfies that goal.

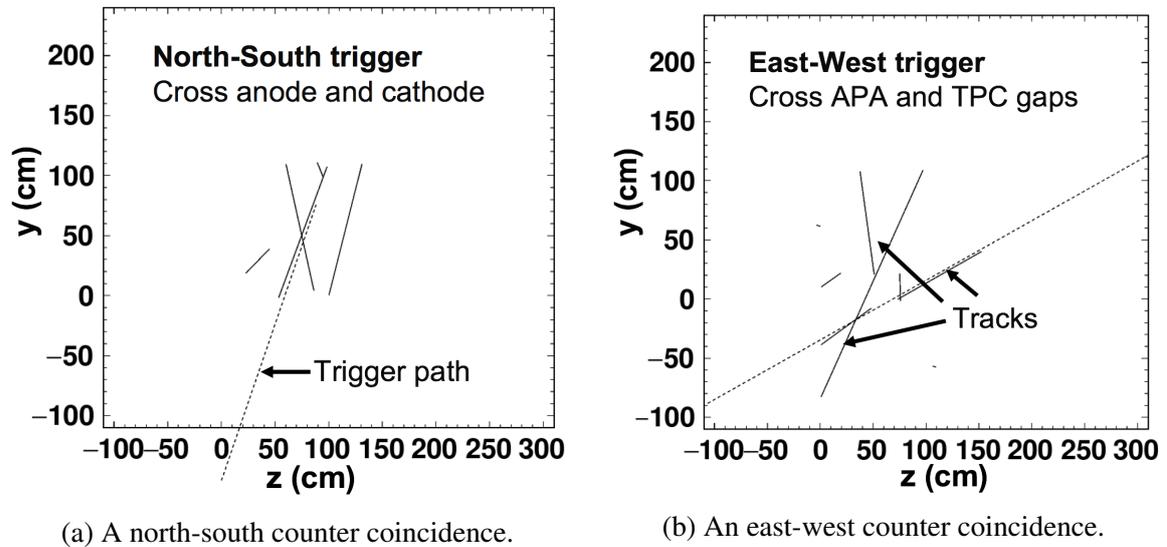


Fig. 5.11 The alignment of reconstructed tracks with the vectors joining the centres of the coincidence counters. The dashed lines show the vectors joining the centres of counters hit in the coincidence, whilst the solid lines show the reconstructed tracks. Left: the alignment of tracks with a north-south coincidence. Right: the alignment of tracks with an east-west coincidence. The z positions of the tracks are shown on the x axis, and the y positions of the tracks are shown on the y axis. The figures were taken from [147].

An important concept that must be introduced before these reconstruction efficiencies can be described is that of a “counter difference.” The “counter difference” of a coincidence and its associated tracks, is defined as the absolute difference between the counter numbers of the east and west counters that were hit. As such, the “counter differences” of the coincidences shown in Figure 5.9, are 2, 3 and 6 for the black, red and blue coincidences respectively. Given the orientation of the counters, the rarest counter difference will be 9, as only particles which hit counters (E_0 and W_9) and (E_9 and W_0) will have a counter difference of 9. In contrast to this, the most common value for the counter difference is 1, as there are many possible combinations of east and west counters being hit to give this counter difference. In the discussions below “counter difference” is occasionally referred to as “delta counter” or “ Δ counter.” Table 5.1 shows the approximate angles which tracks, with given counter differences, have relative to the APA frames.

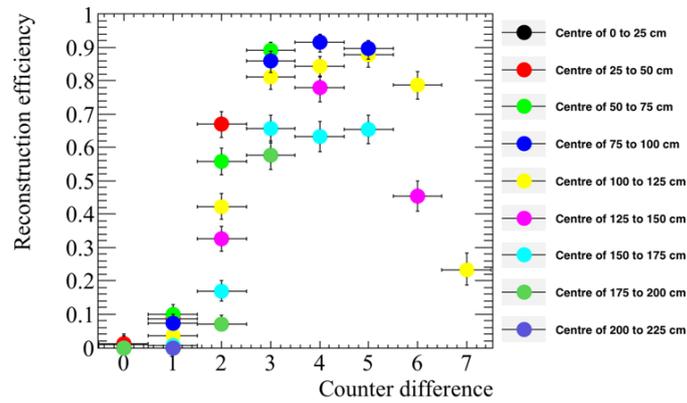
Figure 5.12 shows a range of reconstruction efficiency plots for combinations of different counter differences, and different drift distances. As the counter coincidences with large counter differences will have large variations in drift positions, the drift distance plotted here is the average x position of the counter centres that were hit. For example, if the two counters

Table 5.1 The angles which tracks, with given counter differences, have relative to the APA frames. Angles are calculated using the difference in the centres of the counters in the x direction, divided by the separation of the east and west counters in z . The depth of the counters, their extent in x , is assumed to be negligible.

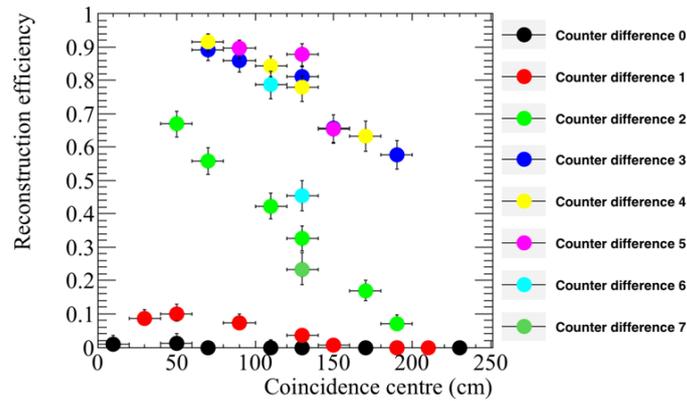
Absolute counter difference	Approximate angle ($^{\circ}$)
0	0 ± 2.1
1	4.2 ± 2.1
2	8.4 ± 2.0
3	12.5 ± 2.0
4	16.5 ± 2.0
5	20.3 ± 1.9
6	23.9 ± 1.8
7	27.3 ± 1.7
8	30.7 ± 1.6
9	33.5 ± 1.5

that produced the coincidence are at $x = 10$ cm and $x = 230$ cm respectively, then the drift distance plotted would be 120 cm. This distance is called the “coincidence centre” in the following discussion. Only coincidences which would produce tracks that are contained within the long drift volume are considered here, hence there being no negative x positions.

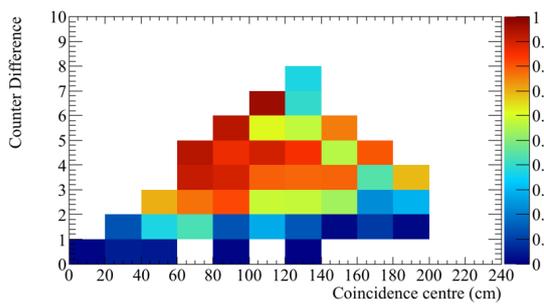
From Figure 5.12a, it is evident that the reconstruction efficiency for tracks with shallow angles relative to the APAs is extremely poor, with the efficiency for tracks aligned with counter differences of 0 or 1 never rising above 10%. This is due to the coherent noise removal, where hits which are correlated in time will be removed as they will be perceived as being due to the coherent noise, as opposed to being due to real signals. As the difference in counter number increases, the efficiency is seen to increase, though the rate of this increase is seen to depend on the “coincidence centre”. The effect of increasing “coincidence centre” can be seen more clearly in Figure 5.12b, where the efficiency for each counter difference as a function of “coincidence centre” is plotted. Here, it can be seen that the reconstruction efficiency decreases for coincidences that are centred further away from the APAs. This is due to the fact that when an energy deposition has further to drift, it will induce a smaller pulse on the wires, meaning that it is more likely to be below the hit threshold. Figure 5.12c combines Figures 5.12a and 5.12b, to show how the reconstruction efficiency for increasing “coincidence centre” changes, with increasing counter difference. It can be seen that tracks with counter differences of between 3 and 5, where the “coincidence centre” is between 60 cm and 140 cm away from the APAs, are the best reconstructed coincidences. Finally,



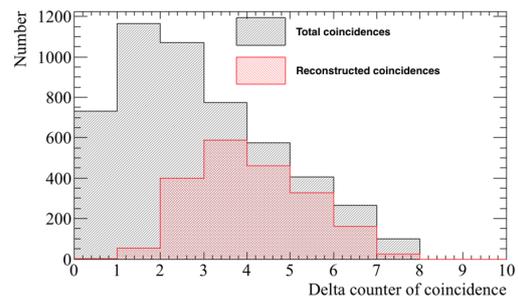
(a) The reconstruction efficiency as a function of counter difference for different coincidence centres.



(b) The reconstruction efficiency as a function of coincidence centres for different counter differences.



(c) The reconstruction efficiency as a function of coincidence centre against the counter differences.



(d) The number of events for each counter difference that were recorded in the data, and the number of those which were successfully reconstructed.

Fig. 5.12 The reconstruction efficiencies for coincidences that trigger an east-west coincidence in the 35 ton data over a 2 day running period.

Figure 5.12d shows how the frequency of coincidences of a given counter difference occurs, compared to how many events contain reconstructed tracks which are aligned with the coincidence. It can be seen that, as stated earlier, the most common counter difference is 1, with the least common being a counter difference of 9. However, given the low reconstruction efficiency seen for the lowest counter differences, few tracks are reconstructed. This means that when considering the reconstructed tracks, most are due to coincidences with counter differences of either 3, 4 or 5.

5.5 Measuring interaction times using electron diffusion

As electrons drift from the interaction point to the wire planes they become spread out in both time and space, this effect is known as diffusion, and is an important property of electron transport in LAr, which must be well understood. The mechanism by which diffusion occurs in LAr was first discussed in [152–155], and has since been developed to consist of a complete set of measurements for electric fields between 100 and 2000 V·cm⁻¹ [156]. The diffusion of electrons is rarely isotropic, and so the component that is transverse to the drift field, and the component that is parallel to the drift field, are normally measured separately. Diffusion parallel to the drift field is called longitudinal diffusion, and is generally smaller than the component of diffusion that is transverse to the drift field. Figure 5.13 shows how diffusion can smear the electrons collected on a set of wires when the electrons are initially highly correlated in time and space.

Longitudinal diffusion has the effect of spreading the drifting electrons out in time, causing signals to become wider in time, and smaller in height, as the total charge is conserved. The increasing hit width can be measured for increasing drift times (distances), provided the hits do not fall below a hit finding threshold. Transverse diffusion causes drifting electrons to spread out in space, changing the amount of charge deposited on a wire, and reducing the charge resolution of the detector. Transverse diffusion is measured by discerning how the FWHM of the hit charge distribution changes for increasing drift distances [156].

Through-going particles make ideal tracks to study diffusion as they are minimally ionising, and so have roughly constant energy depositions along their tracks. The tracks that they produce can also cover a wide range of drift distances, if they are not parallel to the APAs. The drift distances of hits within a track can be determined by matching the track with a counter coincidence as discussed at the end of Section 5.4. The x positions of the hits can

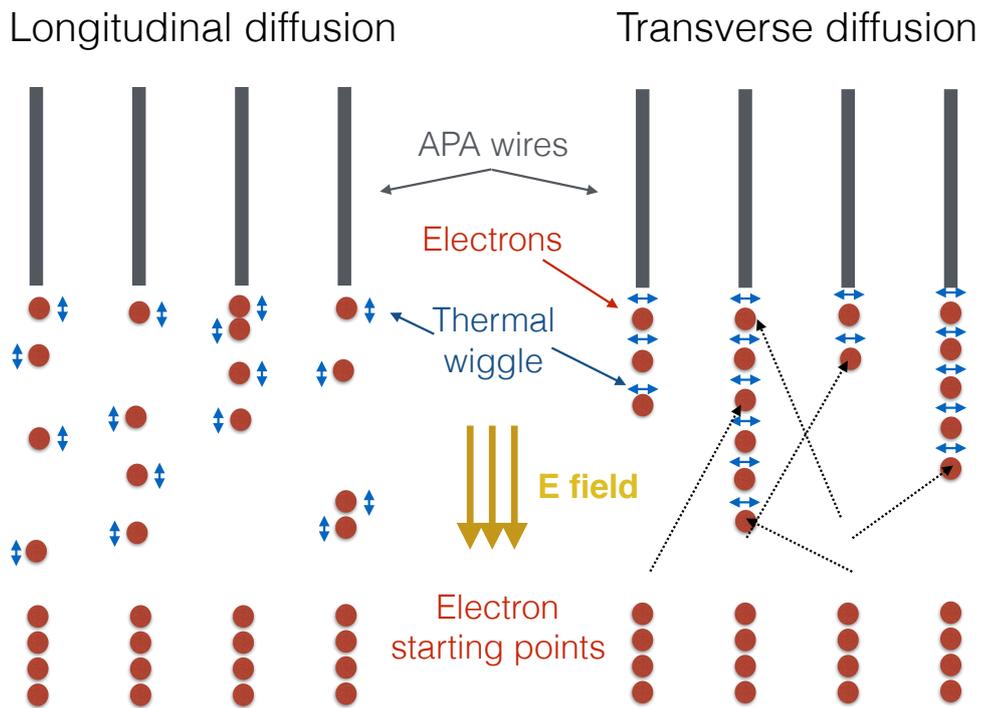


Fig. 5.13 A schematic showing the longitudinal diffusion (left), and the transverse diffusion (right), of electrons. In both cases, four electrons are initially shown below four wires, and are allowed to diffuse in either the drift direction, or perpendicular to the drift direction, in the longitudinal and transverse cases respectively. It can be seen that the effect of longitudinal diffusion is to make the electrons spread out in time, whilst the effect of transverse diffusion is to make the electrons spread out in space. The figure is taken from [157].

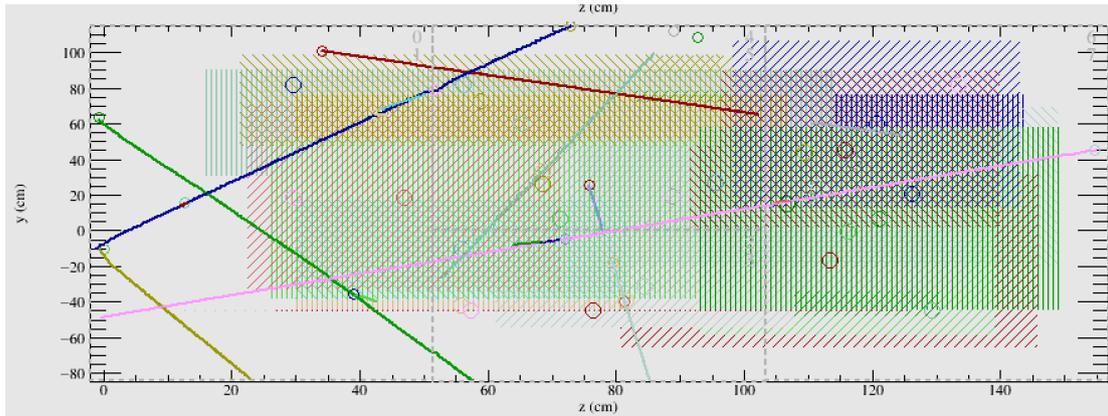


Fig. 5.14 A simulated event display showing multiple tracks and flashes to be assigned to each other in the 35 ton detector, in the yz plane. The coloured lines represent reconstructed tracks, whilst the coloured dashed boxes represent flashes.

then be corrected using the result of Equation 4.2, in Equation 4.1.

Traditionally the only way to determine an interaction time for a track is to either match it to an external calibration source, such as whether it aligns with an external counter coincidence, or to match it to a flash of scintillation light, as in Section 4.1. These techniques are particularly crucial for neutrino detectors on the Earth's surface, such as MicroBooNE, where each neutrino interaction usually has a background of at least one cosmic muon. The reconstructed tracks from this muon background have to be distinguished from those due to the neutrino interactions, in order correctly assign a scintillation flash to the reconstructed tracks. Figure 5.14 shows an event where scintillation flashes and cosmic muons need to be distinguished. However, it may be possible that the change in hit width due to diffusion as a particle travels through the detector could be used to determine the interaction time; though this has not been attempted before. To study whether this is possible, the effects of diffusion would have to be measured for a sample of tracks with known interaction times and orientations.

The 35 ton dataset is ideal for testing this hypothesis, as the counters are able to provide a sample of tracks with known angles and interaction times, which can be used to tune interaction time determination metrics. These metrics can then be applied to another sample of tracks, where the interaction time is known but not used, so that the accuracy of the calculated interaction times can be found. As longitudinal diffusion is the dominant effect that increases the hit time width, transverse diffusion will not be directly considered further. The noise level in the 35 ton data causes reconstruction issues, and so it is also useful to

compare the method against a low noise detector. Monte Carlo can provide this sample, and this comparison is shown in Section 5.5.2. It is also useful to observe the effects that different detector conditions such as, the electric field, the electron lifetime, the noise level, and the magnitude of diffusion, have on the method. This is shown in Section 5.5.3. First though, the method is performed on the 35 ton dataset.

5.5.1 Determining interaction times in 35 ton data

When calculating the determination metrics, only hits on wires which are not noisy are considered. This is because wires with a high level of correlated noise observe hits with a wider RMS. This is shown in Figure 5.15, where, when a baseline noise of 10 ADC counts is added to a simulated hit, with a peak value of 50 ADC counts, and an RMS of 10 ticks, the width increases by over 10%. Hits with delta rays also need to be removed, as the deposited energy will be larger and over a longer period of time than hits from the main track. This will make the RMS of the individual hit wider, and also increase the width of the charge distribution for the track. To remove these hits only hits which satisfy the following cuts are used:

- No hit on the same wire within 50 ticks of the hit in question. This removes delta rays.
- No more than 10 hits on the same wire in the whole 15,000 tick data sample. This removes clearly noisy wires.

These cuts will clearly become much more restrictive in a detector with high levels of electronics noise, but they are essential in order to produce a dataset which is not overpowered by noise. Only collection plane hits are used, as the charge resolution is better, and the signals are unipolar as opposed to bipolar, meaning that a Gaussian function can be easily fitted to the signals. Additionally, the *signal/noise* ratio on the collection planes was much higher than on the induction planes for the 35 ton dataset, and so the hits could be much more reliably reconstructed.

Diffusion is a track angle dependent property, and so track angle ranges have to be considered independently. To minimise the number of figures presented, when graphs are made for all counter differences separately, only graphs made for tracks which have a counter difference of 4 are shown. However, the procedure for predicting interaction times is identical for tracks of all counter differences. Tracks with a counter difference of 4 were chosen as they were one of the angles for which tracks were well reconstructed in the data, see Figure 5.12. Though hits are considered individually before being combined when considering the trends

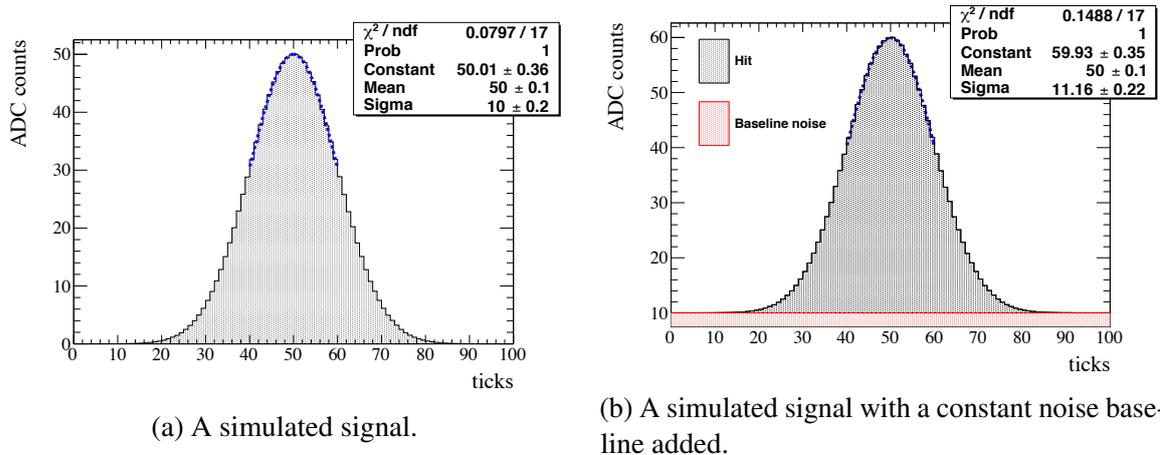


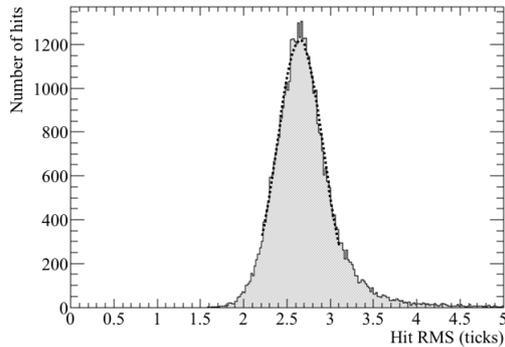
Fig. 5.15 ADC counts as a function of time for a simulated signal with a width of 10 ticks, and an amplitude of 50 ADC counts, both before and after, a constant noise baseline of 10 ADC counts is added, representing either correlated noise, or an incorrect pedestal reduction. The simulated ADC value is shown on the y axis, and the time, in ticks, is shown on the x axis. In reality electronics noise would fluctuate with time, though the effect is analogous. When a Gaussian function is fitted to each signal, it is seen to be more than 10% larger for the signal where the noise baseline is added. This shows that electronics noise, or an improperly subtracted pedestal can cause the measured width of a hit to increase. The figure is taken from [157].

for the entire track, tracks are considered en masse, and so the individual hits for each track are separated into 10 cm regions of increasing drift distance from the APAs. The following quantities are calculated for each 10 cm drift region:

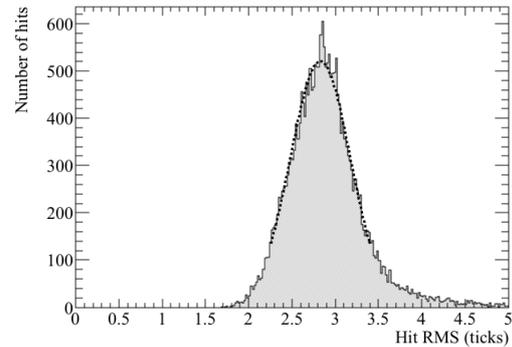
- The hit *RMS* - the most direct way to measure longitudinal diffusion.
- The hit *RMS/Charge* - an attempt to incorporate the effect of impurities in the LAr for relatively low purity data, as this will have a drift distance dependence.
 - The charge of a hit is calculated by integrating the ADCs of the reconstructed hit over time.

Fitting Gaussian functions around the peaks of the distributions should yield the most probable values for the drift regions, as is shown in Figure 5.16. Though it is observed that this is not the case for Figure 5.16c due to the large shoulder at high values of hit *RMS/Charge*, it can be seen that a Gaussian fit yields the correct value for the other fits.

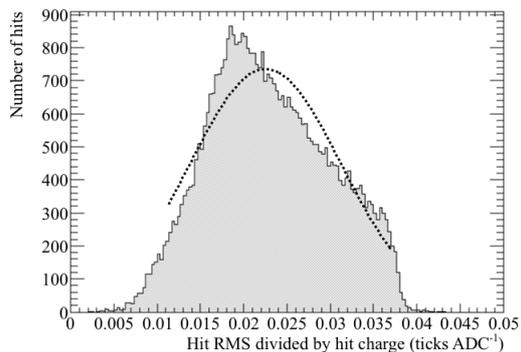
From Figure 5.16 it is clear the the width of the hit *RMS* distribution increases for hits which are further from the APAs. However, the width of the hit *RMS/Charge* is seen



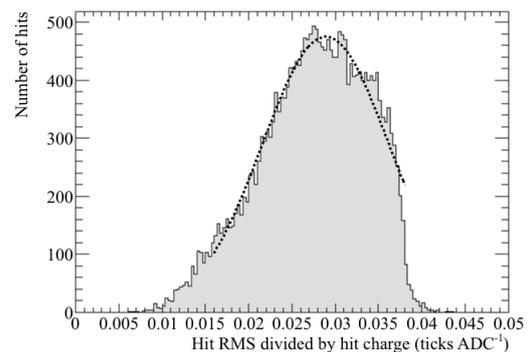
(a) The distribution of hit *RMS* value for hits between $x = 20$ cm and $x = 30$ cm.



(b) The distribution of hit *RMS* value for hits between $x = 140$ cm and $x = 150$ cm.



(c) The distribution of hit *RMS/Charge* value for hits between $x = 20$ cm and $x = 30$ cm.



(d) The distribution of hit *RMS/Charge* value for hits between $x = 140$ cm and $x = 150$ cm.

Fig. 5.16 The distributions of hit *RMS* (top), and hit *RMS/Charge* (bottom), for points between 20 and 30 cm from the APAs (left), and points between 140 and 150 cm from the APAs (right), for tracks associated with coincidences that have a counter differences of 4. The most probable values of hit *RMS* and hit *RMS/Charge* are determined by fitting Gaussian functions around the peaks of the distributions. These fits are shown as dashed lines.

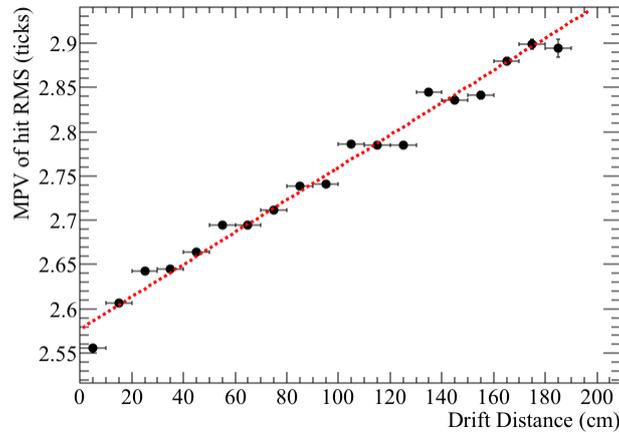


Fig. 5.17 The most probable values of hit *RMS* as a function of drift distance, for the hits within a track, associated with a coincidence that had a counter difference of 4.

to decrease, though this is due to a sharp cut-off at a hit $RMS/Charge$ equal to roughly $0.038 \text{ ticks} \cdot \text{ADC}^{-1}$. The reason for this cut off was shown in Figure 5.8. It can also be seen that the most probable values of both the hit *RMS* and hit $RMS/Charge$ increases with drift distance.

This drift distance effect can be observed by plotting the most probable values of hit *RMS* and hit $RMS/Charge$, as drift distance increases, for fixed counter differences. This drift distance dependence of hit *RMS* is shown in Figure 5.17, for tracks that are associated with a coincidence which had a counter difference of 4. A drift distance dependence can clearly be seen in the data, as the most probable hit *RMS* is seen to increase for hits which originate further from the APAs.

The angular dependence can then be shown by observing how the most probable fit values at a drift distance of 0 cm changes for increasing angles, this is shown in Figure 5.18. It is clear that there is an angular dependence on the hit width, as the most probable hit widths next to the APAs is seen to rise for tracks associated with coincidences with large counter differences. This angular dependence, along with the drift distance dependence, shows that when considering a large sample, diffusion can be separated into distance and angular dependencies. However, whether this can be observed for individual tracks has not yet been considered.

To consider single tracks, the best line fits for the counter differences for a large sample of tracks, such as in Figure 5.17, are required. These best line fits can then be used to predict the position you would expect a hit to originate from, given values for the hit *RMS* and

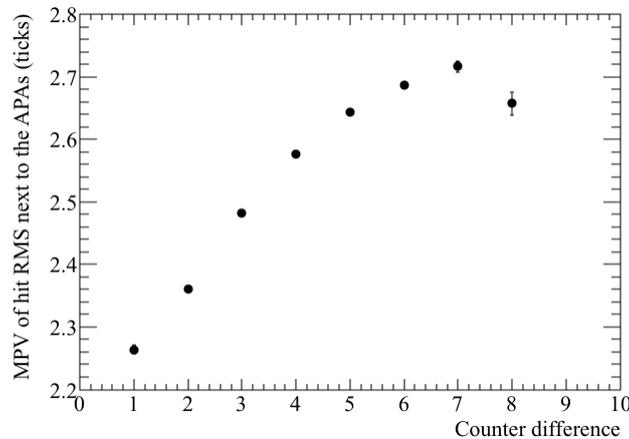


Fig. 5.18 The most probable values of hit *RMS* within 10 cm of the APAs, as a function of the counter difference of the coincidence, that the track, to which the hits belong, was associated with.

hit *RMS/Charge*, and the angle of the track to which it belongs. The predicted positions can then be compared to the known position from the counter coincidence to determine the accuracy of the prediction.

The distributions shown in Figure 5.16 are asymmetric due to some hits having large values of hit *RMS* or hit *Charge*. Asymmetry is also introduced by the threshold for hits. This comes about as a result of the elevated hit threshold, required to minimise the number of reconstructed noise hits, as shown in Figure 5.8. Whilst nothing can be done retrospectively concerning the omission of the lowest charge hits; the highest charge hits, which cause the tails at low hit *RMS/Charge*, can be removed. This can be done by not using hits which are in the tails of the hit *Charge* distribution. The tails of the distributions are removed by considering a plot of the normalised distribution of hit *Charge*, whereby the number of hits at a given value of hit *Charge* is divided by the number hits which have the most probable value of hit *Charge*. The tails of the distribution are removed by performing a cut on the normalised number of hits; whereby all values of hit *Charge* that have less than 0.25 normalised hits are removed. This is shown in Figure 5.19. Any hits with charges less than ~ 65 ADC, and any hits with charges more than ~ 170 ADC are not used. This will have the effect of removing some of the low charge hits which were reconstructed, but the main effect will be the removal of the significant number of high charge hits, which introduce the tail at low hit *RMS/Charge* (Figure 5.16).

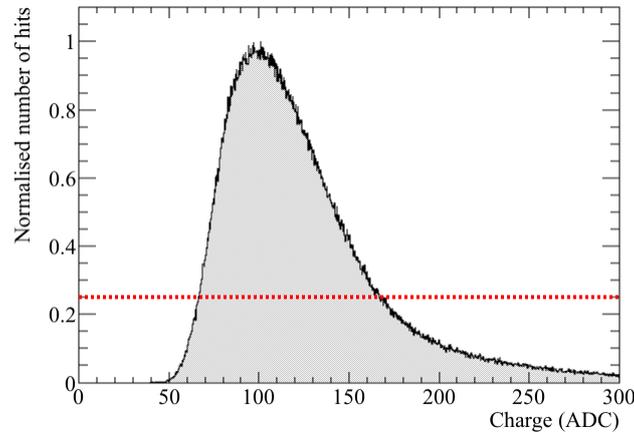


Fig. 5.19 The normalised distribution of hit *Charge*, shown in units of ADC, in the 35 ton dataset. The number of hits with the most probable hit *Charge* has been normalised to a value of 1. A cut on the normalised number of hits being greater than 0.25 is shown, the aim of this cut is to remove the tails of the hit *Charge* distribution.

After removing these hits, the difference in the hit times which are predicted from the parameterised fits, and the hit times which are reconstructed from the data, should be centred around the interaction time, and be distributed much more uniformly. This difference in hit times will be centred around the interaction time as the reconstructed hit times will be larger than the predicted hit times by an amount equal to the interaction time of the particle. In the following plots this interaction time has been calibrated out to be equal to 0, as the interaction time was a priori subtracted from the reconstructed hit times.

An intrinsic assumption in this method is that the track has a large number of collection plane hits, which do not contain delta rays, and are on wires which would not be identified as noisy. The tracks being considered here will have crossed all z values in the detector, meaning that a total of 336 collection hits could potentially be reconstructed. Given the reconstruction problems in the 35 ton detector, very few tracks will have hits on all of these collection wires. However, requiring at least 100 collection plane hits is not unreasonable, and would correspond to a reconstructed track length of at least 50 cm. The difference between the predicted and reconstructed hit time for each hit is shown in Figure 5.20, for both the hit *RMS* and hit *RMS/Charge* metrics.

Figure 5.20 shows that both distributions are centred around a time difference of $0 \mu\text{s}$ in the 35 ton dataset. This is encouraging as it shows that the method has potential. The width of the distribution for the *RMS/Charge* metric is smaller, and the peak higher, so it is

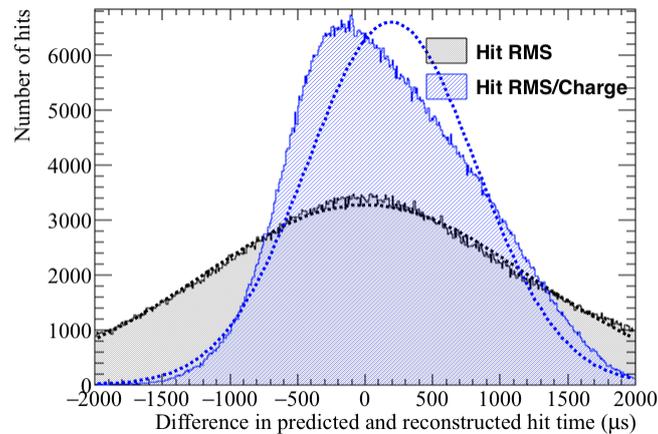


Fig. 5.20 The difference between the predicted and reconstructed hit times in the 35 ton dataset. The differences in time when the hit *RMS* metric is used are shown in black, whilst the differences in time when the hit *RMS/Charge* metric is used are shown in blue.

expected that this will provide the more robust metric. This is because these features show that the predicted hit times are likely to be close to the reconstructed hit times. The peaks are centred around a time difference of 0, as the hit times had previously been corrected using the measured interaction time from the counter coincidence. This was done so as to avoid the uncertainty which would arise from allowing the coincidences to remain at random times between ticks 5000 and 6000, in the 15000 tick event. For an explanation as to why this occurs, see the discussion concerning Figure 5.2.

It is interesting to observe the non-symmetric nature of the hit *RMS/Charge* distribution which is seen in Figure 5.20. This is caused by there being a tendency for the prediction metric to overestimate the x positions of the track. However, it is clearly evident that many of the predicted x positions are highly accurate, as there is a definite peak around a difference in predicted and reconstructed hit times of $0 \mu\text{s}$. When the difference in predicted and reconstructed hit times is plotted as a function of the central x position of the counter coincidence, it is seen that this bias towards overestimating the interaction time comes from hits which are close to the APA frames. This suggests that the fit which was made in Figure 5.16c has produced an incorrect MPV.

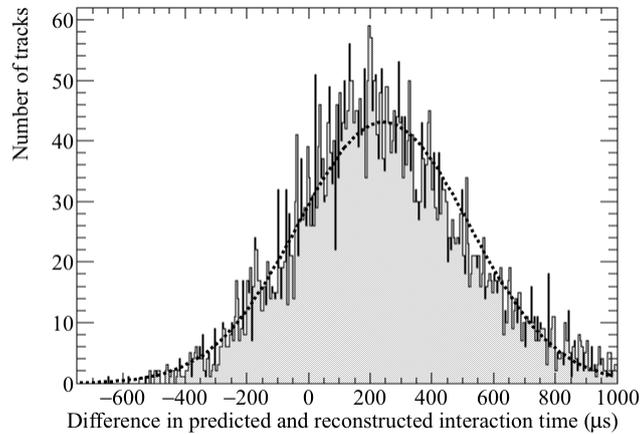
When evaluating interaction times, the average difference in reconstructed and predicted hit times across every hit on the track must be considered. This average difference in reconstructed and predicted hit times, is calculated by taking the sum of the individual differences in reconstructed and predicted hit times for every collection plane hit in the track, and dividing

this by the number of collection plane hits in the track. This is shown in Figures 5.21 and 5.22, where, as expected from Figure 5.20, the *RMS/Charge* metric provides a better estimation of the interaction time. The reason for this is that by utilising the charge information due to losses from impurities, this metric gains an extra handle on the drift distance, and hence the reconstructed time of the hits. The losses due to impurities may be difficult to measure in very high-purity LAr environments, as the decrease in collected charge with increasing drift distances is small [158]. The effect of increasing LAr purity is shown in Section 5.5.3. Using the change in hit charge in the 35 ton may have a drawback though, because, as shown in Figure 5.8, there is a threshold effect for hits with large drift times. However, as the same threshold effect is present in all 35 ton data samples, the limitation it introduces is mainly in the efficiency with which “good” collection plane hits will be reconstructed, and so this information can be confidently used.

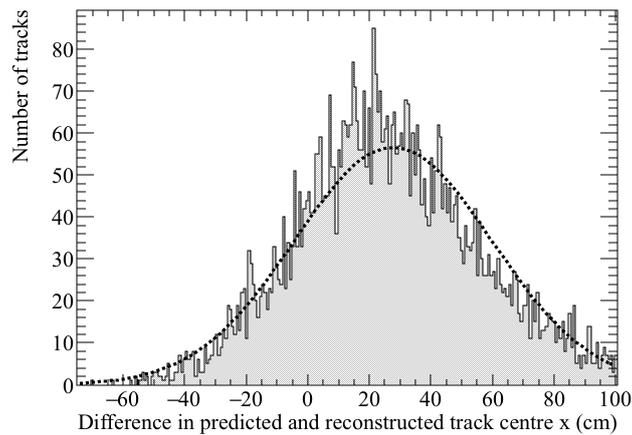
Figure 5.21 shows that using the effects of diffusion, and the hit *RMS*, the interaction time and central x position of a track, can be reliably predicted in the 35 ton dataset. It is found that there is a systematic offset in determining the interaction time of $240 \mu\text{s}$, where the distribution has a FWHM of $281 \mu\text{s}$. When this is converted into the difference in central x position of the track, the systematic offset is found to be 27.7 cm , with a FWHM of 32.1 cm . It should be noted that these systematic offsets, and the ones described below, could be calibrated out through the addition of a constant to the interaction time / central x position of the track which is calculated.

Figure 5.22 shows that using the effects of diffusion, and the hit *RMS/Charge*, the interaction time and central x position of a track, can be reliably predicted in the 35 ton dataset. It is found that there is a systematic offset in determining the interaction time of $171 \mu\text{s}$, where the distribution has a FWHM of $210 \mu\text{s}$. When this is converted into the difference in central x position of the track, the systematic offset is found to be 18.5 cm , with a FWHM of 23.0 cm .

The resolutions found are quite impressive, as given that the total drift time for electrons through the whole 35 ton detector volume of 223 cm is roughly 5200 ticks, it means that tracks can be cleanly distinguished throughout the detector volume. Though the resolution of the interaction time determination is impressive, it is concerning that there appears to be a systematic offset which has been introduced. This is seen by the distributions not being centred around a difference in predicted and reconstructed times of $0 \mu\text{s}$, as would be expected from the discussion concerning Figure 5.20. As discussed earlier, the issues with noise in the 35 ton dataset affect the accuracy with which tracking and calorimetry can

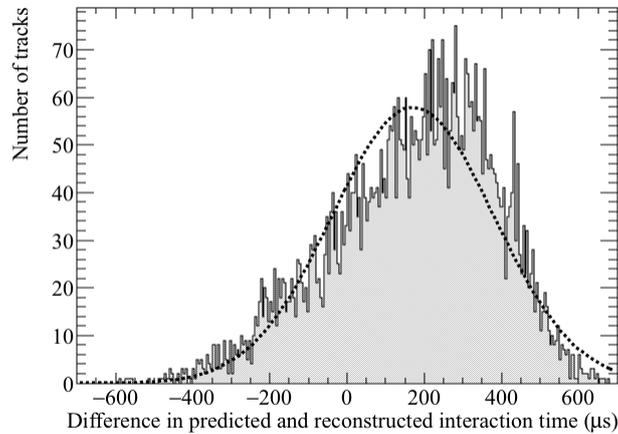


(a) The average difference in interaction times using the hit *RMS* metric.

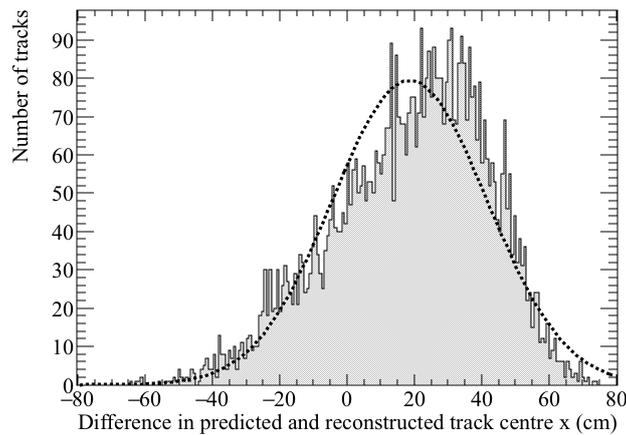


(b) The average difference in the central x position of a track using the hit *RMS* metric.

Fig. 5.21 The accuracy of the hit *RMS* method in the 35 ton dataset. Top: the accuracy to which interaction times can be determined in μs . Bottom: the accuracy to which the central x position of a track can be determined. The average time difference (x position) is calculated by taking the sum of individual hit differences for every hit in the track, and dividing this by the number of hits in the track. Gaussian functions are fitted to the distributions so that any offset in the predicted times or positions can be discerned.



(a) The average difference in interaction times using the hit *RMS/Charge* metric.



(b) The average difference in the central x position of a track using the hit *RMS/Charge* metric.

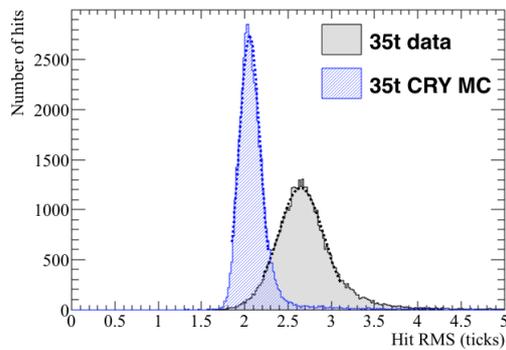
Fig. 5.22 The accuracy of the hit *RMS/Charge* method in the 35 ton dataset. Top: the accuracy to which interaction times can be determined in μs . Bottom: the accuracy to which the central x position of a track can be determined. The average time difference (x position) is calculated by taking the sum of individual hit differences for every hit in the track, and dividing this by the number of hits in the track. Gaussian functions are fitted to the distributions so that any offset in the predicted times or positions can be discerned.

be performed, and so it is reasonable to expect that the effectiveness of the interaction time determination was also affected. Therefore, it is prudent to repeat the study on a Monte Carlo sample, with the same detector conditions, but a much lower level of detector noise. This is presented in Section 5.5.2.

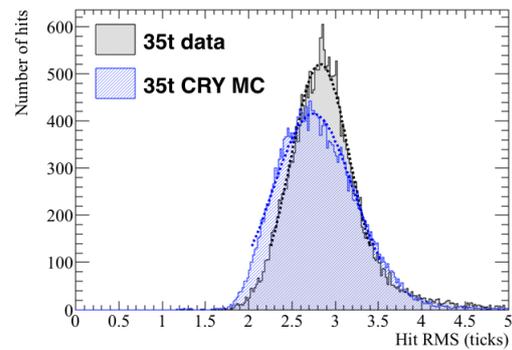
5.5.2 Determining interaction times in a low-noise detector using Monte Carlo, and differences with data

When determining interaction times in Monte Carlo simulations, exactly the same criteria are applied to the hits. This is because δ -rays would still change the measured hit width, and will be present in any sample. In a low noise detector it is expected that few wires would be removed due to being noisy, but for consistency there is no danger in applying this cut. Imposing a minimum number of collection plane hits is again important to ensure that the distribution of predicted hit times is centred on the interaction time. In addition to the same criteria being imposed on which wires are used, the same metrics are calculated. In all plots shown below the Monte Carlo dataset has been normalised to the size of the 35 ton dataset. This was done so that the area of the plots shown was the same, enabling easier comparison between the two datasets.

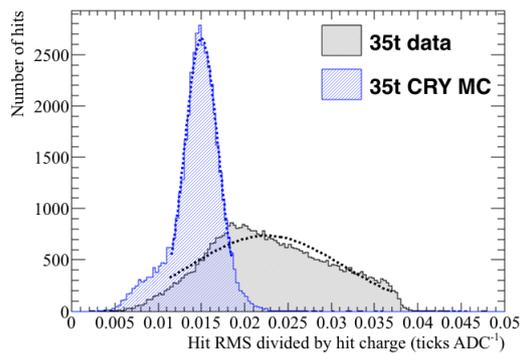
Figure 5.23 shows both the hit *RMS* and hit *RMS/Charge* distributions, for hits from tracks that are associated with a coincidence that has a counter difference of 4, and are between 20 cm and 30 cm away from the APAs, or between 140 cm and 150 cm from the APAs. It can be seen that there is a large difference in the distributions for hits which are relatively close to the APAs, at distances between 20 cm and 30 cm. The distributions are also much tighter in the Monte Carlo sample, showing that the variation between hits is much smaller. However, the difference between the Monte Carlo and 35 ton data samples are much smaller at large drift distances, showing that at large distances the distributions become much more varied. An important feature of the 35 ton data sample, which is not present in the Monte Carlo sample, is the sudden cut off in values of high hit *RMS/Charge*. This was briefly discussed in the consideration of Figure 5.16, but it is much clearer in Figure 5.23d, where the rapid decrease in the number of hits with values of hit *RMS/Charge* which are more than $0.038 \text{ ticks} \cdot \text{ADC}^{-1}$ seen in the 35 ton dataset, is not repeated in the Monte Carlo sample. This, along with Figure 5.8, shows clear evidence that the hits with low values of hit *charge* are not reconstructed in the 35 ton data sample.



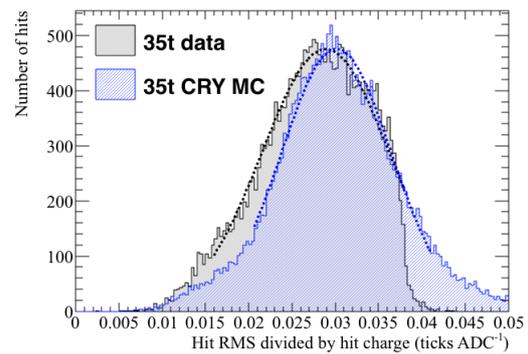
(a) The distribution of hit *RMS* value for hits between $x = 20$ cm and $x = 30$ cm.



(b) The distribution of hit *RMS* value for hits between $x = 140$ cm and $x = 150$ cm.



(c) The distribution of hit *RMS/Charge* value for hits between $x = 20$ cm and $x = 30$ cm.



(d) The distribution of hit *RMS/Charge* value for hits between $x = 140$ cm and $x = 150$ cm.

Fig. 5.23 The distributions of hit *RMS* (top), and hit *RMS/Charge* (bottom), for points between 20 and 30 cm from the APAs (left), and points between 140 and 150 cm from the APAs (right), for tracks associated with coincidences that have a counter differences of 4. The most probable values hit *RMS* and hit *RMS/Charge* are determined by fitting Gaussian functions around the peaks of the distributions. These fits are shown as dashed lines. The distributions for the 35 ton dataset are shown in black, whilst the distributions for the Monte Carlo simulation are shown in blue.

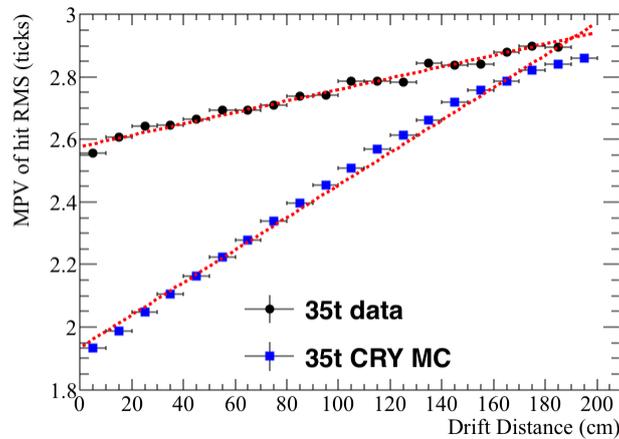


Fig. 5.24 The most probable values of hit RMS as a function of drift distance, for tracks associated with a coincidence that had a counter difference of 4. The distribution for the 35 ton dataset is shown in black, whilst the distribution for the Monte Carlo simulation is shown in blue.

The most probable values of hit RMS at increasing drift distance are shown in Figure 5.24, where the Monte Carlo simulation is again shown along with the values from the data. As was seen when considering the distributions at specific distances and counter differences, the most probable values of hit RMS in the Monte Carlo simulation are systematically lower than in the data. This is attributed to the elevated noise level seen in the data, because, as seen in Figure 5.15, when a noise base-line is added to a signal, the width of the signal increases. Another difference between the Monte Carlo and the data, is that the gradient of the most probable values of hit RMS in data is roughly half of that in the Monte Carlo. This could be due to an overestimation of the effect of longitudinal diffusion in the Monte Carlo sample, or, it could be due to the larger hit widths seen in the dataset at low drift distances, causing the effects of diffusion to be less apparent. Evidence for the latter can be seen in the stark differences between the figures in Figure 5.23 at relatively short drift distances, compared to the similarities seen at large drift distances, where it appears that the Monte Carlo accurately simulates the distributions seen in the 35 ton dataset.

The most probable value of hit RMS at a drift distance of 0 cm for a range of counter differences is shown in Figure 5.25. The change in MPV of hit RMS can be seen to increase for both the Monte Carlo and 35 ton data samples, which again shows that the effects of diffusion can be seen to be track angle dependant. However, the way in which the MPV of hit RMS increases is different in the two samples. The increased MPVs of hit RMS seen in the 35 ton dataset, is again ascribed to the increase in hit width caused by the elevated noise

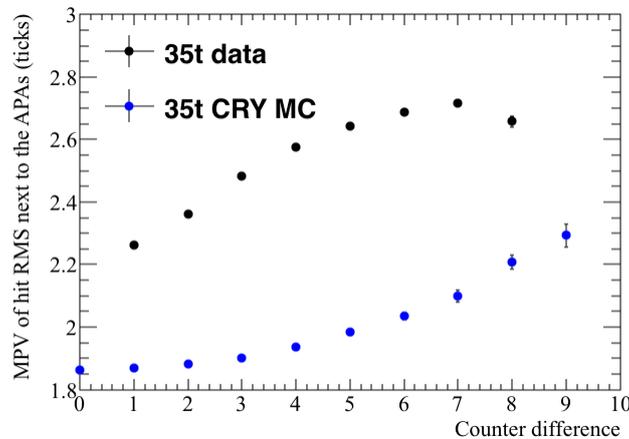


Fig. 5.25 The most probable values of hit *RMS* within 10 cm of the APAs, as a function of the counter difference of the coincidence, that the track was associated with. The distribution for the 35 ton dataset is shown in black, whilst the distribution for the Monte Carlo simulation is shown in blue.

level.

Upon calculating the fit metrics in the low-noise Monte Carlo dataset, it is then possible to use these to predict track interaction times. However, it is first necessary to calculate the normalised hit charge distributions, as was done for the 35 ton dataset. This is shown in Figure 5.26. It can be seen that there are hits with lower values of hit *charge* in the Monte Carlo sample, supporting the idea that there is a threshold effect in the 35 ton dataset. Importantly however, the aim of the cut to remove the tails of the hit *charge* distributions can be seen to be successful in both the 35 ton dataset and Monte Carlo sample, as hits with large value of hit *charge* are removed in both samples. The result of this will be that, the difference in predicted and reconstructed hit times, for a given track, will be centred on the interaction time of the track, as was presented in the discussion of Figure 5.19. The Monte Carlo hit times have been corrected using the time of the counter coincidence, as was the case for the 35 ton dataset.

Figure 5.27 compares how reliably the interaction time, and central *x* position of a track, can be predicted using the effect that diffusion has on the hit *RMS*, in the 35 ton dataset and a low-noise Monte Carlo sample. As was the case when considering the 35 ton dataset, the average time difference (*x* position) is calculated by taking the sum of individual hit differences for every hit in the track, and dividing this by the number of hits in the track. It is found that there is a systematic offset in determining the interaction time of 108 (240) μs in

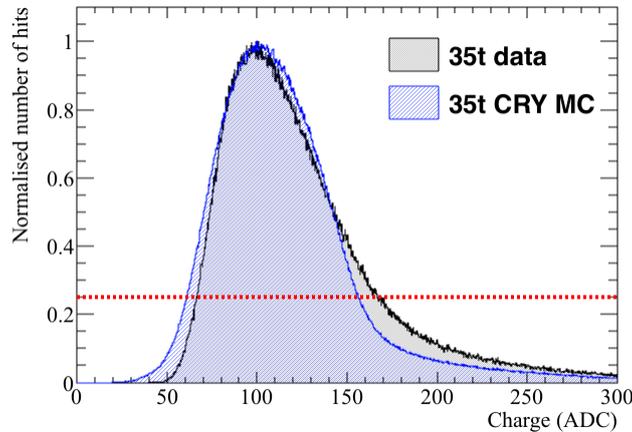
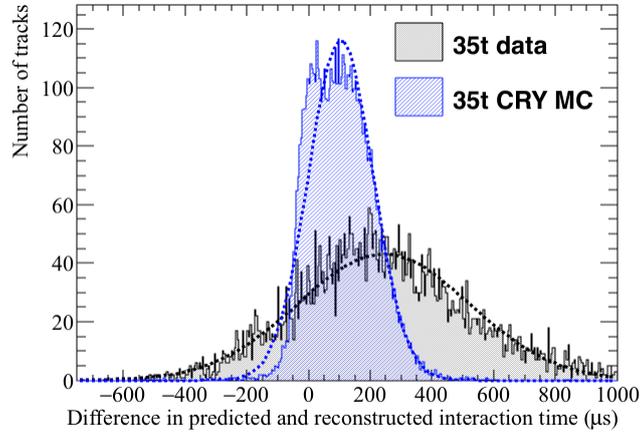


Fig. 5.26 The normalised distribution of hit charge, shown in units of ADC, in the 35 ton dataset and Monte Carlo sample. The number of hits with the most probable hit charge has been normalised to a value of 1. A cut on the normalised number of hits being greater than 0.25 is shown, the aim of this cut is to remove the tails of the hit charge distribution. The distribution for the 35 ton dataset is shown in black, whilst the distribution for the Monte Carlo simulation is shown in blue.

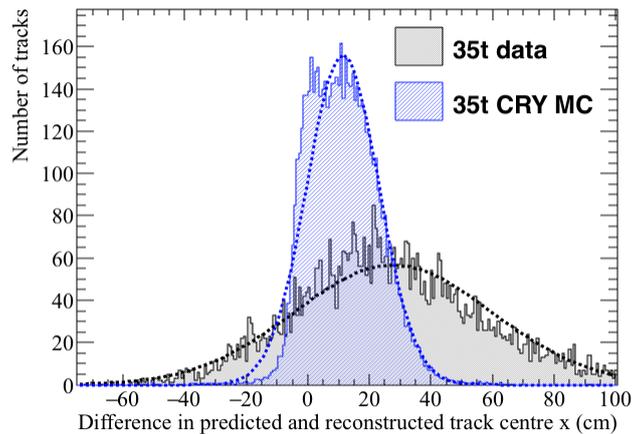
the Monte Carlo (data), where the distribution has a FWHM of 98 (281) μs . When this is converted into the difference in central x position of the track, the systematic offset is found to be, 11.8 (27.7) cm with a FWHM of 10.9 (32.1) cm.

Figure 5.28 compares how reliably the interaction time, and central x position, of a track can be predicted using the effect that diffusion has on the hit $RMS/Charge$, in the 35 ton dataset and a low-noise Monte Carlo sample. It is found that there is a systematic offset in determining the interaction time of 3 (171) μs in the Monte Carlo (data), where the distribution has a FWHM of 114 (210) μs . When this is converted into the difference in central x position of the track, the systematic offset is found to be, 0.4 (18.5) cm with a FWHM of 12.6 (23.0) cm.

The hit $RMS/Charge$ metric appears to be able to more accurately predict interaction times, as was seen when considering the 35 ton dataset. This is again due to the ability to incorporate information about losses due to impurities, which increase with drift distance. Also, as expected from the previous figures, and the lower noise state in the Monte Carlo sample, it is seen that the interaction times predicted in the Monte Carlo sample more closely match the true interaction times, than was the case with the predictions made in the 35 ton dataset. An important feature to observe is that, as well as more accurately predicting the

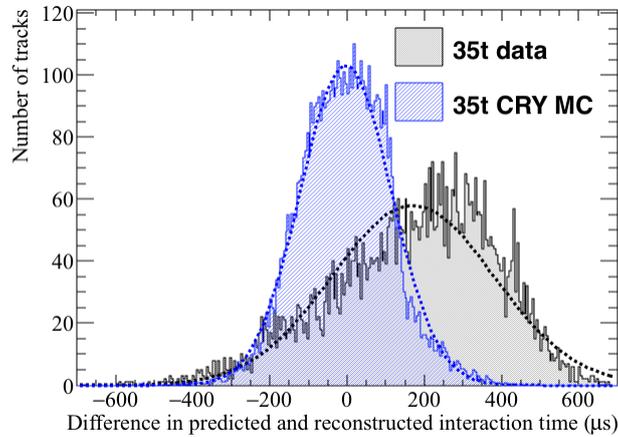


(a) The average difference in interaction times using the hit *RMS* metric.

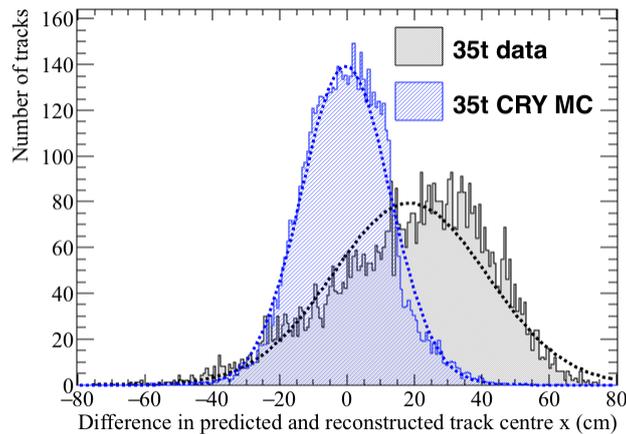


(b) The average difference in the central x position of a track using the hit *RMS* metric.

Fig. 5.27 The accuracy of the hit *RMS* method in the 35 ton dataset and a Monte Carlo simulation. Top: the accuracy to which interaction times can be determined in μs . Bottom: the accuracy to which the central x position of a track can be determined. The average time difference (x position) is calculated by taking the sum of individual hit differences for every hit in the track, and dividing this by the number of hits in the track. Gaussian functions are fitted to the distributions so that any offset in the predicted times or positions can be discerned. The distributions for the 35 ton dataset are shown in black, whilst the distributions for the Monte Carlo simulation are shown in blue



(a) The average difference in interaction times using the hit *RMS/Charge* metric.



(b) The average difference in the central x position of a track using the hit *RMS/Charge* metric.

Fig. 5.28 The accuracy of the hit *RMS/Charge* method in the 35 ton dataset and a Monte Carlo simulation. Top: the accuracy to which interaction times can be determined in μs . Bottom: the accuracy to which the central x position of a track can be determined. The average time difference (x position) is calculated by taking the sum of individual hit differences for every hit in the track, and dividing this by the number of hits in the track. Gaussian functions are fitted to the distributions so that any offset in the predicted times or positions can be discerned. The distributions for the 35 ton dataset are shown in black, whilst the distributions for the Monte Carlo simulation are shown in blue

interaction times, the widths of the distributions in Monte Carlo are less than half of that in the data, particularly when using the hit *RMS*, as shown in Figure 5.27. This means that the resolution with which tracks can be distinguished in the Monte Carlo sample is much better than in the 35 ton dataset, again this is attributed to the lower noise level in the Monte Carlo.

The calculation of interaction times is clearly much better in the low-noise Monte Carlo dataset, than in the 35 ton dataset. However, the distributions when using the hit *RMS* are still not centred around 0, implying that there is a systematic error in the method, which has not been removed when considering a low-noise environment. The cut applied on the normalised hit *charge* distribution was applied in order to remove the tails of the hit *RMS/Charge* distribution, as seen in Figure 5.26. Given that Figure 5.28 shows that when using this metric, the predicted interaction times are centred on the reconstructed interaction times, this appears to have been successful. However, the analogous cut is not performed on the hit *RMS* distribution, and so this could explain the decreased accuracy using this metric.

It is also possible that some δ -rays have not been removed. This is because the only way to remove hits containing unseparated δ -rays, is to look for the slight dip in the raw signal, which is associated with the δ -ray moving away from the main track. This would be almost impossible in the 35 ton dataset given the oscillatory nature of the noise. Were a cut on hit *RMS* to be applied, then these indistinguishable δ -rays would likely be removed. This is because hits with δ rays would lie in the high value tails of the hit *RMS* distribution, which the cut would remove.

The 35 ton dataset as a whole overestimates the interaction times though, and it is thought that this is due to elevated noise level. Justification for this assertion is discussed in Section 5.5.3 where the detector conditions of a simulated detector are varied. One of the detector conditions which is varied is the noise level in the detector.

5.5.3 Impact of changing detector properties using Monte Carlo samples

Much has been made of the difficulty that the noise level in the 35 ton dataset introduces, when performing reconstruction and analysis of the data. It is necessary to verify this claim, and so a sample of Monte Carlo events with increasing noise levels is produced, and analysed below. The noise level in the Monte Carlo samples is increased from the low-noise state used

in the previous section, to a level more similar to that which is seen in the 35 ton dataset. If the claim that the noise level made reconstruction more difficult is correct, then the accuracy with which the interaction time can be determined should be seen to anti-correlate with the noise level of the simulated detector. In addition to varying noise levels, the electron lifetime, the electric field, and the constant of longitudinal diffusion are varied. All samples have used the same initial muons, this is done so that the only difference between the different samples are the detector conditions. Only one detector condition is varied at a time, so that the effect of each detector condition can be studied in isolation. As only one detector property is changed between samples, there is one sample that is consistent to all sample sets. This is when the RMS of the noise is 2.5 ADCs, the electron lifetime is 3 ms, the electric field is $500 \text{ V}\cdot\text{cm}^{-1}$, and the coefficient of longitudinal diffusion is $6.2 \times 10^9 \text{ cm}^2\cdot\text{ns}^{-1}$ [155]. The coefficient of longitudinal diffusion (D_{long}) is defined as the amount of diffusion in the longitudinal direction, and is calculated using Equation 5.3 [155].

$$\begin{aligned}\mu(E) &= 1.1\mu(0)\sqrt{(E_h/E)} \\ \varepsilon_{long}(E) &= 0.4T(E_h/E) \\ D_{long}(E) &= \frac{\varepsilon_{long}\mu(E)}{e}\end{aligned}\tag{5.3}$$

Where T is the temperature of the LAr, E is the Electric field strength, E_h is the boundary field strength, or the point at which electrons begin to heat [155], and e is the absolute charge of an electron. When presenting the studies with changing detector conditions, only the accuracy with which the interaction time and central x position of a track can be predicted, is shown here. A more robust collection of figures can be seen in Appendix A.

Figures 5.29 and 5.30 show the effect that the increase in the noise level has on the accuracy to which the interaction time, and central x position of a track, can be determined.

Figures 5.31 and 5.32 show the effect that different electron lifetimes have on the accuracy to which the interaction time, and central x position of a track, can be determined.

Figures 5.33 and 5.34 show the effect that different electric fields have on the accuracy to which the interaction time, and central x position of a track, can be determined.

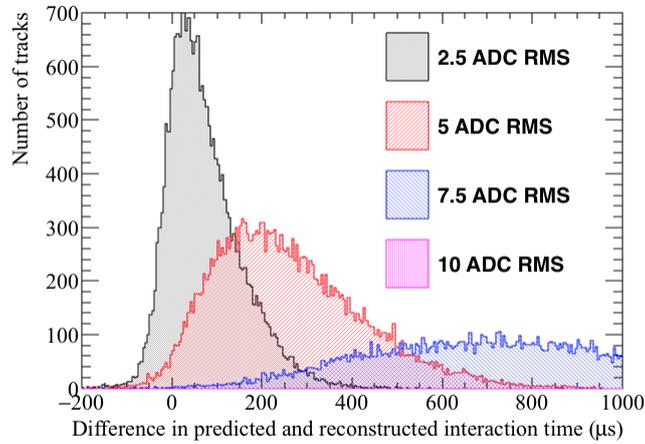
Figures 5.35 and 5.36 show the effect that different constants of longitudinal diffusion have on the accuracy to which the interaction time, and central x position of a track, can be

determined.

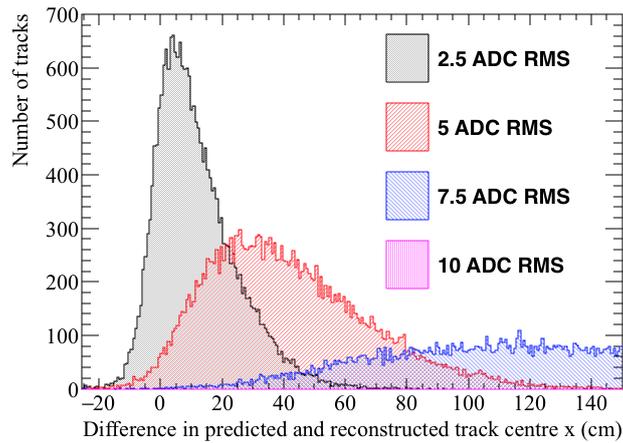
Figures 5.29 and 5.30 show the accuracy to which the interaction time, and central x position of a track, can be determined using the effect that diffusion has on the hit RMS and hit $RMS/Charge$, as the noise level in the detector changes. Figures 5.29 and 5.30 both show that the accuracy of the fits decrease with increasing noise levels, but the decrease in accuracy is manifested in different ways. As discussed in Section 5.3, the 35 ton data had significant amounts of coherent noise which was not expected, and so this had not previously been simulated. As this level of coherent noise is not expected in future detectors, coherent noise has not been simulated in these increased noise level samples. Instead, the electronics noise, or “thermal noise,” has been varied. The lowest noise level was the design noise level for the 35 ton, and is the noise level that is used in the other figures in this section. This level of thermal noise is very minimal, and so only noise levels which are more than this have been simulated when the effect of the noise level in the detector is observed. This is because the $signal/noise$ ratio which one gets with such a low ADC RMS is large, and so a decrease in this noise level is unlikely to make a significant difference in the accuracy of the method. However, as can be seen in the 35 ton data, and the following plots, increasing the noise level has serious consequences.

No noise mitigation algorithms have been applied to the increased noise samples shown here. Instead, the threshold that the hit finder uses has been increased to the level that was necessary for a reasonable number of hits to be reconstructed. A reasonable number of hits simply means, not reconstructing such a large number of noise hits that they outweigh the number of true signals from tracks. The required hit threshold was determined by looking at the deconvoluted signal, and choosing a threshold which was above the majority of the noise signals. The hit thresholds used for each noise level are summarised below:

- Noise level of 2.5 ADC RMS - hit threshold of 6 ADC, corresponding to an energy of $60 \text{ ADC}\cdot\text{cm}^{-1}$.
- Noise level of 5 ADC RMS - hit threshold of 10 ADC, corresponding to an energy of $100 \text{ ADC}\cdot\text{cm}^{-1}$.
- Noise level of 7.5 ADC RMS - hit threshold of 15 ADC, corresponding to an energy of $150 \text{ ADC}\cdot\text{cm}^{-1}$.
- Noise level of 10 ADC RMS - hit threshold of 20 ADC, corresponding to an energy of $200 \text{ ADC}\cdot\text{cm}^{-1}$.

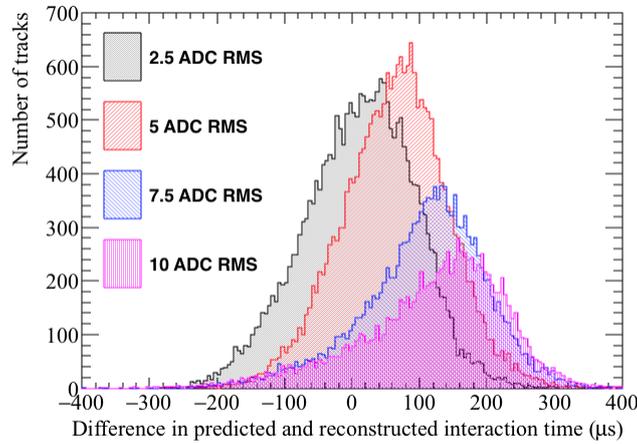


(a) The average difference in interaction times using the hit *RMS* metric.

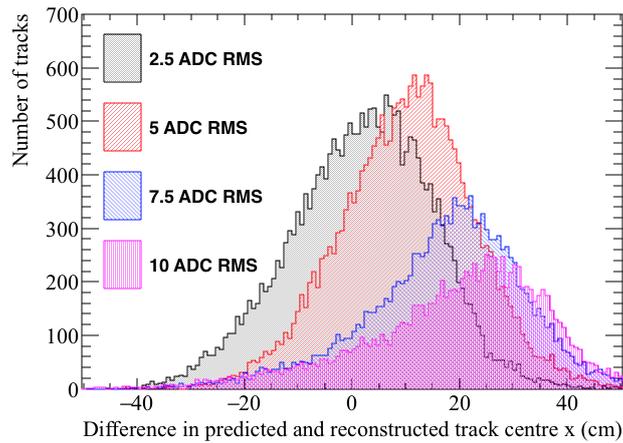


(b) The average difference in the central x position of a track using the hit *RMS* metric.

Fig. 5.29 The accuracy of the hit *RMS* method, for different electronic noise levels. Top: the accuracy to which interaction times can be determined in μs . Bottom: the accuracy to which the central x position of a track can be determined. The average time difference (x position) is calculated by taking the sum of individual hit differences for every hit in the track, and dividing this by the number of hits in the track. The 10 ADC RMS is not shown as the distribution is not contained within the graph.

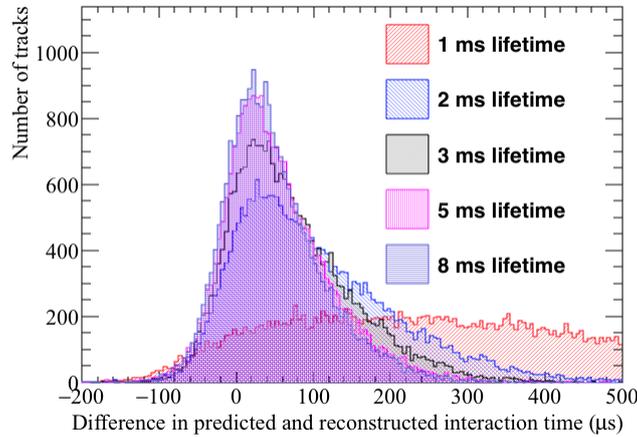


(a) The average difference in interaction times using the hit *RMS/Charge* metric.

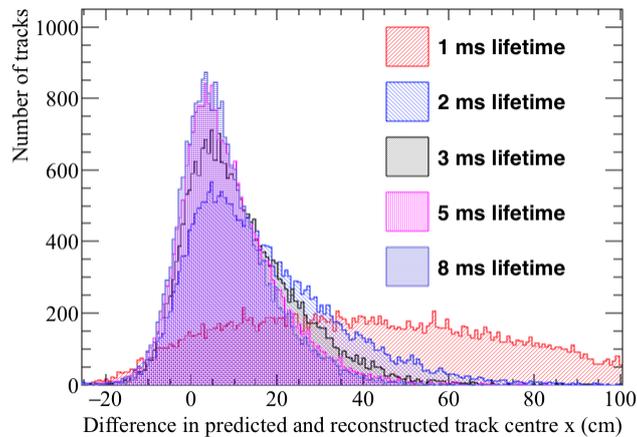


(b) The average difference in the central x position of a track using the hit *RMS/Charge* metric.

Fig. 5.30 The accuracy of the hit *RMS/Charge* method, for different electronic noise levels. Top: the accuracy to which interaction times can be determined in μs . Bottom: the accuracy to which the central x position of a track can be determined. The average time difference (x position) is calculated by taking the sum of individual hit differences for every hit in the track, and dividing this by the number of hits in the track. The 10 ADC RMS is not shown as the distribution is not contained within the graph.

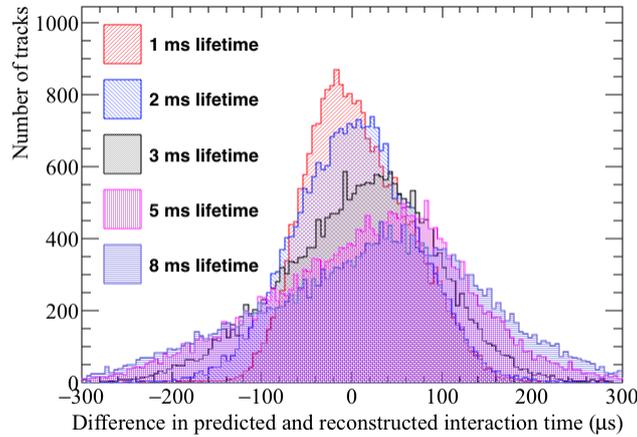


(a) The average difference in interaction times using the hit *RMS* metric.

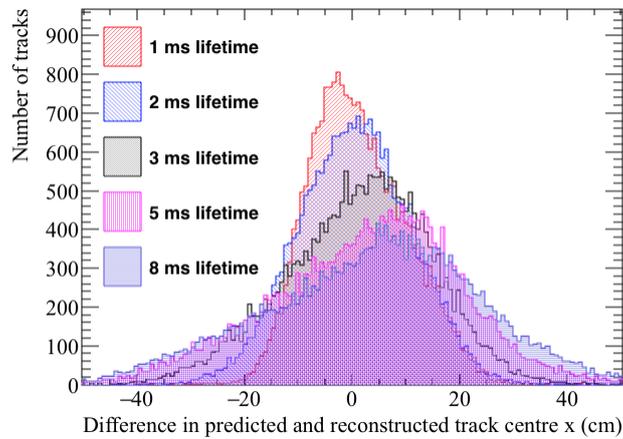


(b) The average difference in the central *x* position of a track using the hit *RMS* metric.

Fig. 5.31 The accuracy of the hit *RMS* method, for different values of the electron lifetime. Top: the accuracy to which interaction times can be determined in μs . Bottom: the accuracy to which the central *x* position of a track can be determined. The average time difference (*x* position) is calculated by taking the sum of individual hit differences for every hit in the track, and dividing this by the number of hits in the track.

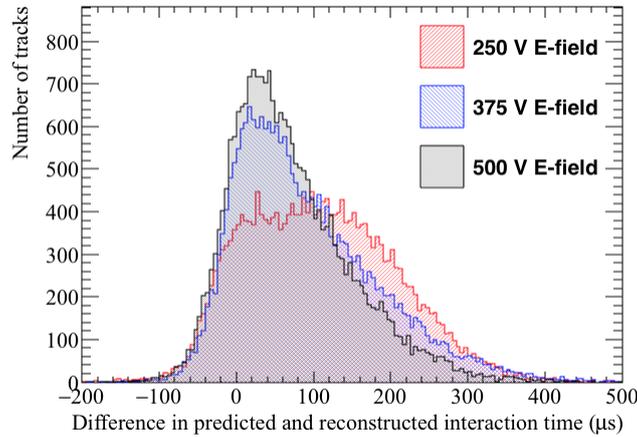


(a) The average difference in interaction times using the hit *RMS/Charge* metric.

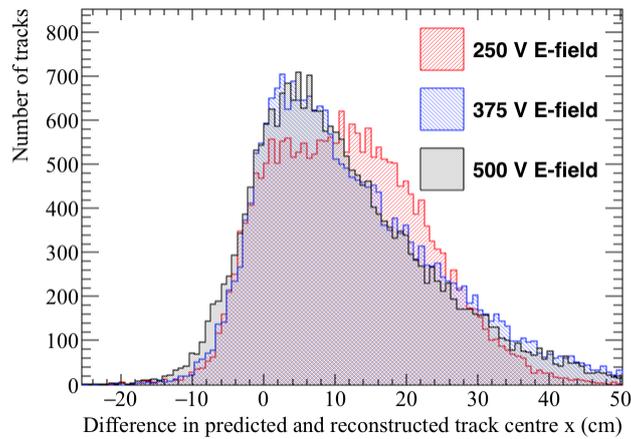


(b) The average difference in the central x position of a track using the hit *RMS/Charge* metric.

Fig. 5.32 The accuracy of the hit *RMS/Charge* method, for different values of the electron lifetime. Top: the accuracy to which interaction times can be determined in μs . Bottom: the accuracy to which the central x position of a track can be determined. The average time difference (x position) is calculated by taking the sum of individual hit differences for every hit in the track, and dividing this by the number of hits in the track.

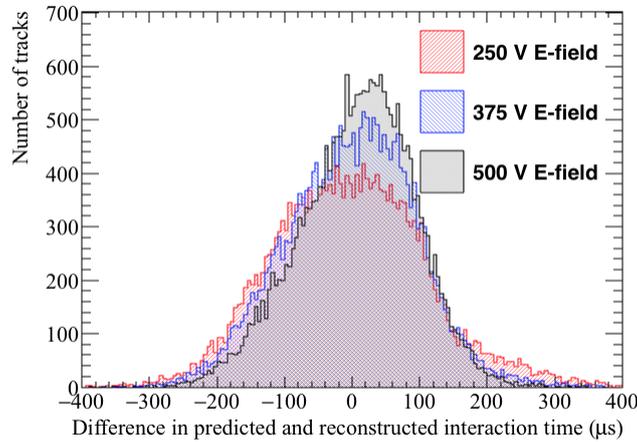


(a) The average difference in interaction times using the hit *RMS* metric.

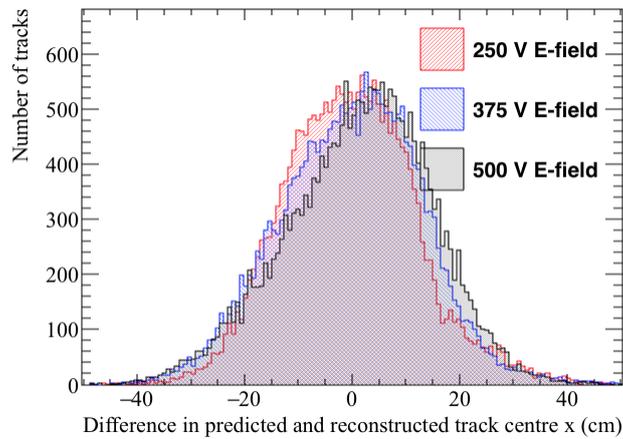


(b) The average difference in the central x position of a track using the hit *RMS* metric.

Fig. 5.33 The accuracy of the hit *RMS* method, for different values of the electric field. Top: the accuracy to which interaction times can be determined in μs . Bottom: the accuracy to which the central x position of a track can be determined. The average time difference (x position) is calculated by taking the sum of individual hit differences for every hit in the track, and dividing this by the number of hits in the track.

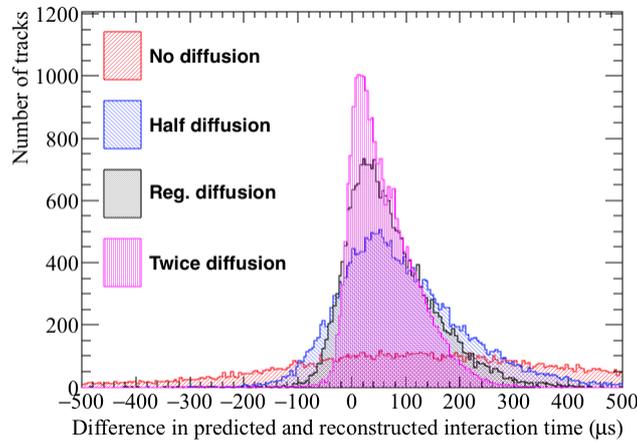


(a) The average difference in interaction times using the hit *RMS/Charge* metric.

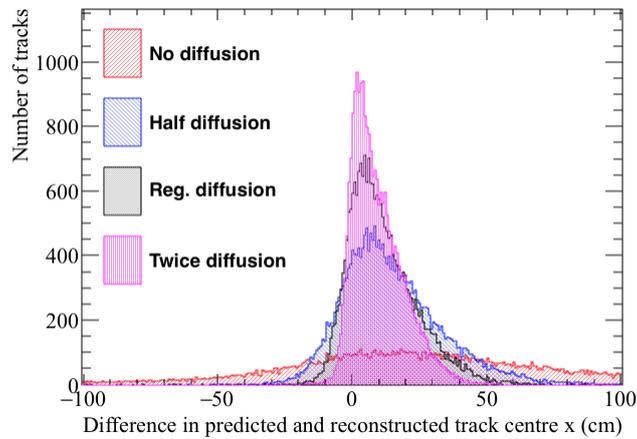


(b) The average difference in the central x position of a track using the hit *RMS/Charge* metric.

Fig. 5.34 The accuracy of the hit *RMS/Charge* method, for different values of the electric field. Top: the accuracy to which interaction times can be determined in μs . Bottom: the accuracy to which the central x position of a track can be determined. The average time difference (x position) is calculated by taking the sum of individual hit differences for every hit in the track, and dividing this by the number of hits in the track.

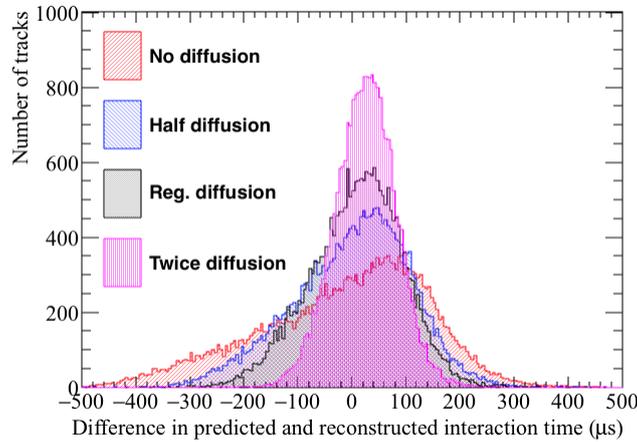


(a) The average difference in interaction times using the hit *RMS* metric.

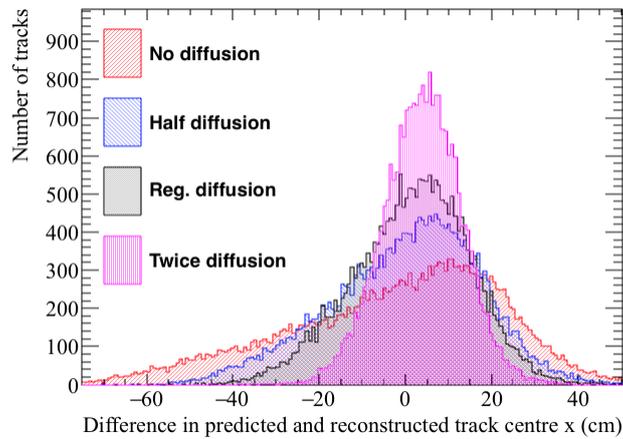


(b) The average difference in the central x position of a track using the hit *RMS* metric.

Fig. 5.35 The accuracy of the hit *RMS* method, for different values of the constant of longitudinal diffusion. Top: the accuracy to which interaction times can be determined in μs . Bottom: the accuracy to which the central x position of a track can be determined. The average time difference (x position) is calculated by taking the sum of individual hit differences for every hit in the track, and dividing this by the number of hits in the track.



(a) The average difference in interaction times using the hit $RMS/Charge$ metric.



(b) The average difference in the central x position of a track using the hit $RMS/Charge$ metric.

Fig. 5.36 The accuracy of the hit $RMS/Charge$ method, for different values of the constant of longitudinal diffusion. Top: the accuracy to which interaction times can be determined in μs . Bottom: the accuracy to which the central x position of a track can be determined. The average time difference (x position) is calculated by taking the sum of individual hit differences for every hit in the track, and dividing this by the number of hits in the track.

This means that the main effect of increasing the noise level is to remove the low charge hits, as they will fall below the hit threshold, as it is increased to compensate for the increased noise level. Given the uncertainty introduced into the calorimetry in the 35 ton detector, as shown in Figure 5.8, it can be seen that every effort should be made to ensure that the DUNE FD has noise levels as close to 2.5 ADC RMS as possible.

When considering Figure 5.29, it can be seen that the accuracy to which interaction times can be determined rapidly decreases as the noise level increases. This is partly due to the fits used to make the prediction metrics for the highest noise level not converging for counter differences of 1, 2, 3 and 4, as the MPV of hit *RMS* is not seen to increase for increasing drift distances. For evidence of this, see the Figures in Appendix A. Though this is the extreme case, it can be seen that the validity of the hit *RMS* for increasing drift distances becomes less predictable as the noise is increased. The result of this is a less accurate prediction metric, which leads to the large offsets and widths of the distributions that are shown in Figure 5.29. This is particularly true for the sample which has a noise level of 10 ADC RMS, where the accuracy of the time determination is so bad that it is not contained on the plot.

The most striking feature of Figure 5.30 is the decrease in statistics seen for the increasing noise levels. This shows the effect that increasing the noise level, and hence hit threshold has. This is because fewer tracks in total are reconstructed, and those that are reconstructed are less likely to meet the criteria about the number of collection hits required to make predictions.

Figure 5.31 shows that with an electron lifetime of 1 ms, the hit *RMS* metric is very inaccurate, this is likely due to hits which are a large distance away from the APAs being very difficult to reconstruct, because of the extremely poor lifetime. For this reason, the accuracy to which the hit *RMS* metric predicts the interaction time improves as the electron lifetime increases, though this increase is small between the 3 ms, 5 ms and 8 ms samples. Figure 5.32 shows the opposite effect, the accuracy to which the interaction time can be determined decreases with increasing electron lifetime for the hit *RMS/Charge* metric. This is shown by the widths of the distribution increasing as the electron lifetime increases. This happens because the decrease in hit charge is much greater when the electron lifetime is lower, and this dependence is the corner stone of this metric. The large decrease in hit charge for low electron lifetimes is why this metric performs so well for low electron lifetimes, and so the decrease in its accuracy is an unavoidable consequence of increasing electron lifetime.

Figure 5.33 shows that the accuracy to which the interaction time can be predicted, increases with increasing electric field. This is shown by the introduction of an offset in the predicted interaction time for lower values of electric field strength. The opposite is shown in Figure 5.34, as the accuracy to which interaction time can be predicted does not see the introduction of an offset, and is slightly better for the samples with lower electric field strengths. However, when these interaction times are converted to the central x position of a track, the accuracy is relatively unaffected by electric field strength. This is because the predicted central x positions are the same for all values of electric fields, and are peaked at the true central x position for both samples. However, there is a large sample of tracks in Figure 5.33 with an offset of about 10 cm, which makes the distribution of the average difference in predicted central x positions to not be Gaussian. The presence of this offset in both samples shows that the hit *RMS* and hit *RMS/Charge* metrics are both relatively unaffected by the electric field strength.

As would be expected, both Figures 5.35 and 5.36 show that the accuracy to which the interaction time and central x position can be predicted are highly dependent on the longitudinal diffusion constant. This is seen by the distributions becoming much narrower, and more closely centred around the true interaction time, or central x position, as the constant of longitudinal diffusion increases. It is interesting to note that the extremely poor resolution seen in Figure 5.35 when there is no longitudinal diffusion, is not present in Figure 5.36. It is thought that this is due to the effect of charge attenuation, which will still occur because of the finite electron lifetime.

5.5.4 The limitations of and future improvements to the method of interaction time determination using diffusion

The comparison of the 35 ton data and Monte Carlo samples, as well as the Monte Carlo samples with differing detector conditions, show that there is potential in the ability to determine interaction times using the effects of diffusion. However, there are still some issues which need to be overcome, some of these are briefly discussed below.

Many of the figures shown still have slight offsets even though the tails of the hit charge distributions have been removed. However, these offsets are generally confined to detector conditions which would not be considered optimal, such as very low electron lifetimes (1 ms) or high detector noise (5 ADC RMS). The latter is seen to be the case when considering the 35 ton dataset, where the high noise scale can be seen to affect the accuracy to which

the interaction time, and central x position, can be determined. A potential solution to reduce these offsets, and also to reduce the width of the distributions, is to perform the interaction time determination twice. The result of the first run, which is what is shown in this thesis, would then be used to select only hits which lie within the expected regions of hit RMS and hit $RMS/Charge$. This would be possible as the initial interaction time determination could be used to work out the rough x positions of hits, and then only hits which lie within a given region of the MPV would be used to determine the interaction time from the second pass. When performing the second pass of interaction time determination, the size of this region would be a user defined parameter, but would need to be small enough so as to exclude the beginnings of the tails of the hit RMS , and hit $RMS/Charge$, distributions.

An important improvement to the method would be to expand it to include the induction plane wires, as this will greatly increase both the number of wires which can be used, and the range of track angles whose interaction times can be predicted. The angular range of the method would increase, since, when using only collection plane wires, it is impossible to reconstruct enough hits for nearly vertical muons, as very few wires would be hit. This was discussed in Section 4.3. This was not attempted here, as the electronics noise in the 35 ton data was too large to be able to reliably reconstruct hits on the induction planes, without reconstructing many noise hits. This meant that the hit threshold on the induction planes was very high.

Chapter 6

Simulations of the DUNE Far Detector

Work presented in Chapters 3, 4 and 5 concerned the 35 ton prototype, however the following simulations have been performed with respect to the DUNE Far Detector (FD). Simulations in the FD have concentrated on cosmogenic background to neutrino oscillations, discussed in Section 6.1, and the muon background to nucleon decay, discussed in Section 6.3. The simulations shown in Section 6.1 are discussed in [159], and were performed for the Long Baseline Neutrino Experiment (LBNE), which along with the Long Baseline Neutrino Observatory (LBNO), formed the basis for DUNE, and so are included here for completeness. The author contributed by incorporating the *complex* detector geometry, and the accurate surface profile into the simulations, though the main body of work was performed by the author of [159]. The other work presented was performed for the DUNE collaboration, in conjunction with work done by Vitaly Kudryavtsev and Matthew Robinson, both at the University of Sheffield. This work was performed with the aim of assessing the muon-induced background to nucleon decay in the DUNE FD, at the 4850 ft level.

6.1 Simulations of the LBNE surface detector

A preliminary design of the LBNE experiment had a 10 kt LAr detector on the surface, with a 3 m rock overburden [160]. Due to the large flux of cosmic rays at such a shallow depth, and the relative scarcity of neutrino events, it is necessary to establish whether the cosmic background can be sufficiently removed to allow neutrino interactions to be discerned. In order to show this, a series of simulations using both a simplified, and a more detailed detector design are performed. The cosmogenic backgrounds due to muons, protons and neutrons, are considered. These simulations, unlike other work presented in this thesis, were performed using a stand-alone version of GEANT4, and not in LArSoft. For a more thorough

review of the simulations, see [159].

Simulations have been built using GEANT4 versions 9.4 and 9.6. Initial simulations were performed using version 9.4, before being rebuilt to use version 9.6. A study showing that the background rate did not change with the newer version of GEANT4 was performed. In all simulations the physics list “Shielding” was used [161]. Shielding allows particles to undergo Compton scattering, inelastic scattering with nuclei and nucleons, and includes all electromagnetic (EM) processes for all leptons, mesons baryons, and generic ions. Muons are able to undergo electron/positron pair production, and electron/positron annihilation is also possible.

The author’s work built on initial simulations which were performed using a simplified geometry, referred to as the *simple* geometry, which consisted of a single box of LAr measuring $30 \times 15 \times 16 \text{ m}^3$. This detector was enclosed in rock measuring $(5 \times 10^3) \times (5 \times 10^3) \times 22 \text{ m}^3$, in the x , y , and z coordinates respectively. The detector was positioned so as to have a 3 m overburden of rock. As simulations were not performed in LArSoft, the co-ordinate system used was defined as follows;

- x - parallel to the beam direction.
- y - perpendicular to the beam direction.
- z - vertical direction.
- The co-ordinate system is centred on the middle of the detector volume in x , y and z .

The results of these initial simulations, discussed in detail in [159], are not shown here, but will be summarised in Table 6.2.

The improvements which the author made to these initial simulations were two fold. Firstly a more detailed geometry, referred to as the *complex* geometry, was included, this had a much more realistic detector design [162]. The *complex* geometry had two identical cryostats, each containing 120 TPC cells measuring $2.52 \times 2.28 \times 7.00 \text{ m}^3$. These cells each contain an active volume of LAr measuring $2.27 \times 2.25 \times 6.30 \text{ m}^3$. The TPC cells are arranged such that are 10, 6 and 2 TPC cells in the x , y and z directions, respectively. This gives a total volume of LAr in each cryostat measuring $28.20 \times 13.95 \times 15.0 \text{ m}^3$, with a mass of 5.35 kton. This gives a total fiducial mass of 10.7 kton of active LAr in the two cryostats. Running vertically between the TPC cells are anode plane assemblies (APAs) and cathode plane assemblies (CPAs), all of which are embedded within the larger blocks of LAr.

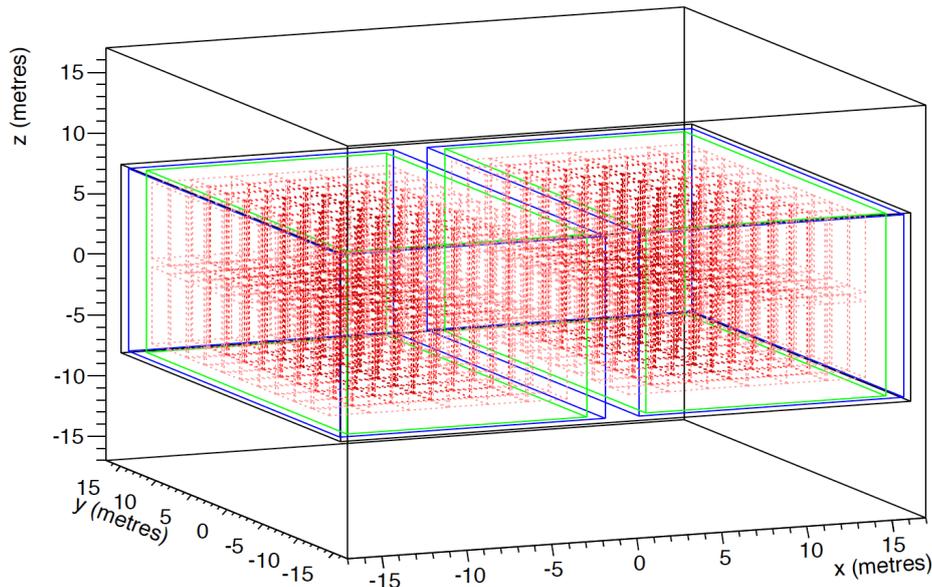


Fig. 6.1 The *complex* detector geometry used in the LBNE surface detector simulations. The TPC cells are shown in red, and are orientated such that there are 10, 6, and 2 TPCs in the x , y , z coordinates of each cryostat respectively. The steel walls of the cryostats are green, whilst the outer edges of polyurethane are blue, and the concrete enclosures are black. Figure taken from [159].

These blocks of LAr are further housed inside stainless steel containers, which are insulated by 0.8 m of polyurethane. The two cryostats are surrounded by 0.5 m of concrete on all sides, with 3 m of concrete separating them. A representation of the detector, produced by the GEANT4 visualisation tool, is shown in Figure 6.1.

Secondly, an accurate surface profile, including the hills surrounding the far detector site was incorporated into the simulations. The proposed location of the LBNE detector was the same as the proposed DUNE detector location, though it was on the surface, and not in the mine. A satellite generated map, centred on SURF, which encompasses an area of $20 \times 20 \text{ km}^2$ is used to generate the accurate surface profile. The map is sampled in bins measuring $5 \times 5 \text{ m}^2$. To include the effects that the surrounding hills would have on the muon flux, muons are initially generated 600 m above the surface, and are then stepped through the surface profile, until they reach a box measuring $80 \times 80 \times 36 \text{ m}^3$ surrounding the detector. The amount of rock that a muon passes through is calculated, and the energy losses which this will cause are then taken into account before subsequent simulations. By taking the surface profile into account in this way, simulation time can be reduced by not having to simulate the surface profile for every new set of muons. The surface profile is modified so as to have

a flat surface measuring $100 \times 100 \text{ m}^2$ above the detector location, as it is envisioned that there will be surface facilities around the detector site. The accurate surface profile is shown in Figure 6.3a, where it is presented in the context of producing muons for the DUNE FD location.

It should be emphasised that all work presented here has been done using Monte Carlo truth, and so no reconstruction has been performed. It also means that at all times only Monte Carlo truth information is used. Therefore, any positions or trajectories which are referred to are those which are recorded by GEANT4 [108]. This information saved by GEANT4 is “smeared” before analysis is performed, in an attempt to take into account of some of the detector effects which were not simulated. This smearing is particle type and energy dependant, and is performed on the particles position, trajectory, and energy [159, 163].

6.1.1 Classifying signal and background events

Before the simulations can be described in detail, it is necessary to outline what separates a signal event from a background event. In the LBNE and DUNE FD, a ν_e appearance signal event would occur when a ν_e undergoes a charged current interaction with either an electron or nucleon. This interaction will produce an electron, which will in turn produce an EM shower. It is this EM shower which will be identified as a neutrino appearance signal. The electron track which is produced in these interactions will not necessarily be isolated, and may be accompanied by hadronic debris.

Cosmic ray particles are able to produce signals which mimic this appearance signal. These signal mimicking events can come from a large variety of sources, including, but not limited to, knock-on electrons from muons, bremsstrahlung photons, meson decays to photons, and EM showers which originate outside of the active volume of the detector but then enter it. A background event is then defined as the initial photon in an EM shower. The first generation photon is used as this will contain the total energy of the shower, and removes the need to record every particle produced in the shower. The final position (point of pair production) of the photon is used in all calculations where the position is required, as the photon will not be observed in the actual detector, only the electrons and positrons it produces. Background showers which start with an electron from Compton scattering are not counted, as pair production is the major source of EM showers at the energies considered here.

6.1.2 Description of cuts used

As noted above, the event rate due to cosmic backgrounds will be much larger than the neutrino event rate, and so a series of cuts, designed to remove cosmic background whilst preserving signal events, has to be developed. This section will briefly outline the cuts which were developed to achieve this. A more rigorous definition of the cuts can be found in [159].

The simplest cut considers the energy of the electromagnetic (EM) cascade which is induced. As is the case for DUNE, the LBNE beamline was a broadband neutrino beam, where neutrino analyses would have been concentrated within the 0.25 - 5.0 GeV energy range. This means that any EM cascade which deposited more than 5 GeV of energy, or less than 0.25 GeV of energy into the detector would not be considered as a signal event.

As charged particles within the detector will produce tracks, it is possible to calculate the minimum separation between these charged tracks and the point at which a photon pair produces. This minimum distance of separation, called Point of Closest Approach (PoCA) in this thesis, is calculated by extrapolating the trajectory of the photon backwards from the location at which it pair produces, towards the charged particle tracks. The charged particle tracks are also extrapolated backwards, and so the smallest separations may be outside of the active volume of the detector. A photon is identified as being due to the charged particle track if the PoCA is below a certain threshold. Should a photon be due to a charged particle track, then it is identified as being a background event. Thresholds of 30 cm and 10 cm are used when considering the initial charged muon, and all charged particles respectively. The latter threshold is smaller to avoid removing neutrino interactions, and also has a lower limit of 2 cm, meaning that any photons which have PoCAs with respect to any charged particle of less than 2 cm are not identified as background events. This is because it is possible that charged hadrons are produced at the neutrino interaction vertex, and so the electron produced in the interaction could be found to have a charged particle very closely related to it, meaning that were this lower limit not in place, this electron would fail the all charged PoCA cut.

It has been found that the angle of a neutrino event with respect to the neutrino beam is highly correlated to the energy of the neutrino [164]. This means that it is possible to use the angle of a shower with respect to the beam, to distinguish between signal and background events. The effectiveness of this constraint is highly energy dependant, as any high energy photons which are not tightly correlated to the beam axis will be identified as backgrounds, though few low energy photons will be [165]. Figure 6.2 shows an example of how the PoCA

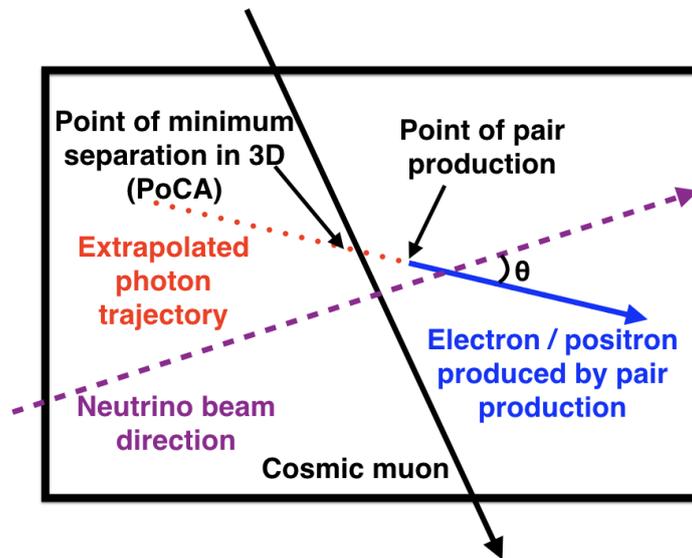


Fig. 6.2 A diagram of the PoCA, and the angle w.r.t the beam calculations in the LBNE surface detector simulations. A cosmic muon which passes through the detector is shown as a black line, whilst an electron/positron that is produced is shown as a blue line. The minimum distance between the muon and electron/positron is shown (d), as well as the point of closest approach (PoCA) when the electron/positron track is extrapolated backwards. The extrapolated track is shown as a dotted red line. The neutrino beam direction is shown as a dashed green line, and the angle that the electron/positron makes with respect to this is shown (θ). Figure is modified from [159].

and angle of a photon with respect to the beam axis are calculated.

It is envisioned that not all of the instrumented LAr will be able to be used to identify neutrino events. This is because a fiducial cut will be applied around the detector edges, in order to ensure that the entire track is contained within the active volume of the detector. For the *complex* detector geometry, this fiducial cut is only applied to the outward facing edges of the active volumes. It is assumed that tracks passing between cells, within a given cryostat, will be able to be stitched together, as is done in LArSoft.

Additionally, signal events should be able to be distinguished from background events based on the energy deposition measurements. These will come from identifying the start of an EM shower as coming from either an electron, or a photon, in the case of a signal, and background event, respectively. Studies have shown that the failure rate of this separation tails off at around 10% for showers with energies above 0.5 GeV [166]. Therefore, a flat reduction of 1/10 is applied to all surviving simulated background events. This cut will later

be referred to as “ $e - \gamma$ separation.”

Finally, it is envisioned that an efficient photon detection system will be used, which uses the scintillation light emitted by excited argon [167]. This efficient photon detection system should be able to provide information on individual events, and so identify whether a candidate event occurred within the beam spill. If this is used, then the effective drift time of the detector can be reduced from 1400 μs to only 10 μs , a reduction by a factor of 140. This is also applied as a flat reduction to all surviving background events.

6.1.3 Generating particles for background studies

When simulating the cosmogenic background for a detector on the surface it is necessary to simulate the backgrounds due to muons, protons, neutrons and photons. It was found that the background due to cosmic photons was negligible [159], and so they are not discussed here.

The muons used in the simulations for the LBNE surface detector are generated using a Gaisser’s parameterisation [114], which has been modified for large zenith angles, and muon decay [168]. The differential muon intensity at sea level ($\frac{dI_\mu}{dE_\mu d\Omega}(E_\mu, \theta)$), in units of $\text{cm}^{-2} \cdot \text{s}^{-1} \cdot \text{sr}^{-1} \cdot \text{GeV}^{-1}$, is shown in Equation 6.1.

$$\begin{aligned} \frac{dI_\mu}{dE_\mu d\Omega}(E_\mu, \theta) = & A \times 0.14 \times (E_\mu + \Delta)^{-\gamma} \times p_d \\ & \times \left[\frac{1}{1 + \frac{1.1(E_\mu + \Delta)\cos\theta^*}{115\text{GeV}}} + \frac{0.054}{1 + \frac{1.1(E_\mu + \Delta)\cos\theta^*}{850\text{GeV}}} + R_c \right] \end{aligned} \quad (6.1)$$

where E_μ is the muon energy in GeV at the surface, θ is the muon zenith angle at the surface, θ^* is the muon zenith angle at the height of muon production, $\cos\theta^* = \sqrt{1 - 0.99(1 - \cos^2\theta)}$, Δ is the muon energy loss in the atmosphere and is equal to $2.06 \times 10^{-3} \times (\frac{1030}{\cos\theta^*} - 120)$, R_c is the ratio of prompt muons to pions, and p_d is the probability for a muon to survive in the atmosphere [169]. There is also a normalisation constant A , and a spectral index γ , which can both be chosen to fit experimental data. When considering shallow depths, or surface locations, values of $A = 1$ and $\gamma = 2.7$ are used [114]. When considering larger depths, as will be done in Sections 6.2 and 6.3 where MUSUN [116, 117] is used, the typical values are $A = 1.84$ and $\gamma = 2.77$. These values are taken so as to match the “depth - vertical muon intensity” which was measured by the LVD experiment [168].

Table 6.1 The minimum energies of simulated particles, when determining the cosmogenic background for the LBNE surface detector. The percentage of background events which are caused by particles of these energies, or above, is shown. The percentage of the particle flux above this energy, is also shown. It can be seen that, with appropriate minimum energy constraints, it is possible to avoid simulating over 80% of the muon flux, 99% of the proton flux, and 95% of the neutron flux. Table taken from [159].

Primary particle	Min. energy simulated (GeV)	Background (%)	Particle flux (%)
Muons	10	92.3	19.6
Proton	10	92.7	0.76
Neutrons	1	95.6	6.5

The protons and neutrons used in the simulations are generated using CRY [111, 112], on a plane measuring $50 \times 50 \text{ m}^2$ in the xy plane. It is not necessary to generate protons and neutrons over a large area as their interaction lengths are much shorter than muons, and so protons and neutrons at large angles, which pass through large amounts of rock will not be a serious concern. The particles are generated at an altitude of 2100 m above sea level. However, the location of the LBNE surface detector would have been 1505 m above sea level, and so the particle fluxes generated by CRY have to be corrected to the particle fluxes expected at 1505 m above sea level [159]. The fluxes generated by CRY are also subject to a further correction, as CRY underestimates the cosmic ray flux by as much as 70% [170]. This is because CRY only considers protons striking the Earth's atmosphere, and not heavier nuclei such as deuterons and α -particles. As MUSUN calculate the muon flux for locations at sea level, this flux also has to be corrected so as to be the flux at 1505 m above sea level [159].

To reduce simulation time, only particles which will cause most background events are simulated. For example, it is found that 92.7% of the proton induced background events are due to protons with energies of 10 GeV or more, yet these particles represent only 0.76% of the total proton flux. This means that by not simulating protons with energies below 10 GeV, the vast majority of background events will be simulated, but this will require less than 1% of the simulation time. The background events which are not simulated, can be accounted for by correcting the background seen from simulations, by the proportion of background events which were not simulated [159]. The minimum energies of the simulated muons, protons, and neutrons, along with the percentage of background events which they cause, and the percentage of the particle fluxes above this energy, are shown in Table 6.1.

Table 6.2 The normalised background rate per calendar year for the *simple* detector geometry, separated by particle species, using the flat surface profile. The rates for muons entering the detector, muons missing the detector, proton induced events, and neutron induced events are shown. The annual background for muons entering the detector is only an approximate value, as initial simulations were performed without saving proton hit information, the inclusion of which greatly increases the accuracy of the all charge PoCA cut. Table is taken from [159].

Primary particle	Annual background rate
Muons entering	≈ 1.18
Muons missing	0.11 ± 0.02
Protons	2.57 ± 0.08
Neutrons	1.23 ± 0.07
Total	≈ 5

It can be seen from Table 6.1, that by only simulating protons above 10 GeV 7.3% of the background were not simulated, and so the background rate from simulations should be scaled by a factor of $1.0787 \left(\frac{1}{0.927}\right)$.

6.1.4 Results from background simulations

Before the results of the simulations involving the *complex* geometry, and accurate surface profile are discussed, it is useful to briefly summarise the results of the simulations using the *simple* geometry, and flat surface profile. This is shown in Table 6.2.

It can be seen from Table 6.2 that the overall background rate is dominated by hadronic components of the cosmic flux. It is expected that the additional shielding in the *complex* geometry will suppress these components significantly. The annual background induced by cosmic muons has been split into events where the muon enters the detector, and events where it does not enter the detector. This is because when a muon does not enter the detector, the PoCA with respect to the initial muon cannot be calculated. It will also be instructive to observe how the background from each case changes, as the detector geometry is made more accurate. The *complex* geometry, has both a larger surface area and a larger active volume of LAr, therefore it is expected that the muon fluxes will increase. This increase may be quite large for muons which miss the active volume of the detector, as muons may pass through the vertical gaps between TPC volumes and produce secondaries which enter the active volume of the detector. Events like this would, at first glance at least, appear very

Table 6.3 The normalised background rate per calendar year, for events where a primary muon enters the active volume of the detector, for the *complex* geometry and accurate surface profile. A total of 2×10^8 muons with energies greater than 10 GeV are generated, representing 0.1003 years worth of detector live time. The background rate is separated into different first generation photon ancestries. The application of the cuts outlined in Section 6.1.2 is shown, where E_γ is the 0.25 - 5.0 GeV cut, $PoCA_\mu$ is the PoCA w.r.t. the initial muon cut, $\theta_{beam}(E)$ is the energy dependent cut on the angle between the beam and photon trajectory, $PoCA_{all}$ is the PoCA w.r.t. all charged particles cut, $D > 30$ is the 30 cm fiducial cut. Following this, two scaling factors of 1/10 and 1/140, representing the e- γ separation, and the use of an efficient photon detection system respectively, are applied. The errors quoted are Gaussian, and are purely statistical, unless the simulated background rate per calendar year drops to 0, in which case an upper limit at 90% confidence level [171] is used, with any scaling factors being applied to this limit. No errors are quoted if the error is less than 1% of the simulated background rate per calendar year. Table is taken from [159].

	E_γ	$PoCA_\mu$	$\theta_{beam}(E)$	$PoCA_{all}$	$D > 30$ cm	$e - \gamma(E)$	γ detection
<i>Total</i>	1.32×10^7	$(6.38 \pm 0.09) \times 10^4$	$(2.87 \pm 0.06) \times 10^4$	3796 ± 229	2854 ± 199	285 ± 20	2.03 ± 0.14
$\pi^0 \rightarrow \gamma$	2.24×10^6	$(5.82 \pm 0.09) \times 10^4$	$(2.62 \pm 0.06) \times 10^4$	3339 ± 215	2743 ± 195	274 ± 20	1.96 ± 0.14
<i>Ext</i> $\rightarrow \gamma$	4.48×10^6	5237 ± 270	2425 ± 183	457 ± 80	111 ± 39	11.1 ± 3.9	0.08 ± 0.03
$\mu \rightarrow \gamma$	6.36×10^6	0 - 34	0 - 15.20	0 - 2.01	0 - 1.51	0 - 0.15	0 - 0.001
<i>Other</i> $\rightarrow \gamma$	7.87×10^4	333 ± 68	97 ± 37	0 - 15.20	0 - 11.43	0 - 0.11	0 - 0.002

similar to a neutrino event, as they will be isolated in the centre of the detector.

Only results for the *complex* geometry and accurate surface profile are shown. A separate set of simulations using the *complex* geometry and simple surface profile are shown in [159]. However, as discussed in [159], these background rates were found to be consistent with the background rates for the *complex* geometry and accurate surface profile.

Table 6.3 shows the background rate for only muons which enter the active volume of detector, as sequential cuts are applied, for the *complex* geometry and accurate surface profile. A total of 2×10^8 muons with energies greater than 10 GeV are generated, representing 0.1003 years worth of detector live time. The overall number of background mimicking events is seen to increase when using the *complex* detector geometry and the accurate surface profile (1.32×10^7), as opposed to using the *simple* detector geometry and simple surface profile (1.07×10^7) [159]. It is observed that the expected background rate per calendar year for the *complex* detector geometry in the flat (1.94 ± 0.23), and accurate surface profiles (2.03 ± 0.14), are consistent [159]. This means that including the accurate surface profile does not have a significant effect on the background rate.

The effectiveness of the PoCA cut with respect to the initial muon (2nd column) is obvious, as the annual background event rate is reduced by over 99%. The rejection rate is observed to be 100% when considering the $\mu \rightarrow \gamma$ photon ancestry (4th row). The rejection rate is also very high in the $Ext \rightarrow \gamma$ photon ancestry (3rd row) where it approaches 100%. This column delimitates photons whose parents did not enter the detector. Therefore, that the parent of the photon was a muon can only be discerned when the trajectory of the photon is extrapolated backwards, outside of the detector. When this is done it is found that many of the photons have very small PoCAs with respect to the muon, and so they are identified as background events. The remaining background is dominated by photons with π^0 parents (2nd row). Around half of the surviving events are removed by the application of the angular cut (3rd column), but more than 80% of these events are removed by the application of the all charged PoCA cut (4th column). This again shows how powerful the use of the PoCA cut is when rejecting background events. Whilst the fiducial cut (5th column) is seen to be effective at removing the remaining $Ext \rightarrow \gamma$ events, this is not the case for the $\pi^0 \rightarrow \gamma$ events. Therefore, the development of efficient methods of e- γ separation (6th column), and the use of an efficient photon detection system (7th column), are critical in reducing the annual background to a rate which would not be prohibitive to observing neutrino interactions.

The background rate for the *complex* detector geometry and accurate surface profile, for muons which miss the active volume of the detector, is shown in Table 6.4. The same sample of muons generated to create Table 6.3 are used here, meaning that a detector live time of 0.1003 years has been simulated, though only events where the muon does not enter the detector are used. The overall number of background mimicking events is seen to increase substantially when using the *complex* detector geometry and the accurate surface profile (19800 ± 500), as opposed to using the *simple* detector geometry and simple surface profile (10700 ± 300) [159]. This is to be expected, as there is a much larger external surface area, and the vertical gaps produce internal gaps between the TPCs. The expected background rate per calendar year for the *complex* detector geometry in the flat (0.58 ± 0.14), and accurate surface profiles (0.65 ± 0.08), are consistent, though the total number of background events is less when the accurate surface profile is used (19800 ± 500 compared to 22100 ± 900) [159].

Table 6.4 shows that muons which miss the active volume of the detector cause much fewer background events than those which strike the active volume of the detector. However, the events which they do cause, are much more likely to survive the application of all cuts. This is due to a combination of both the fiducial cut being less effective, and there being more external photons. Both of these differences are caused by muons which

Table 6.4 The normalised background rate per calendar year, for events where a primary muon misses the active volume of the detector, for the *complex* geometry and accurate surface profile. A total of 2×10^8 muons with energies greater than 10 GeV are generated, representing 0.1003 years worth of detector live time. The background rate is separated into different first generation photon ancestries. The application of the cuts outlined in Section 6.1.2 is shown. Table is taken from [159].

	E_γ	$D > 30$ cm	$\theta_{beam}(E)$	$PoCA_{all}$	$e - \gamma(E)$	γ detection
<i>Total</i>	19800 ± 500	4004 ± 235	1551 ± 147	914 ± 113	91 ± 11	0.65 ± 0.08
$\pi^0 \rightarrow \gamma$	3284 ± 213	1108 ± 123	526 ± 85	166 ± 48	16.63 ± 4.80	0.12 ± 0.03
<i>Ext</i> $\rightarrow \gamma$	16500 ± 500	2858 ± 199	998 ± 118	748 ± 102	75 ± 10	0.53 ± 0.07
<i>Other</i> $\rightarrow \gamma$	28 ± 5	28 ± 5	28 ± 5	$0 - 20$	$0 - 2.01$	$0 - 0.01$

pass through gaps between TPCs, and produce secondaries which are inside the detector walls. These events will produce photons which are not removed by the fiducial cut, and are identified as external photons, as the muon in these events did not produce a track. There will also be more external photons, as the surface area of the detector has substantially increased now that there are two identical cryostats, as opposed to a single block of LAr.

Table 6.5 shows the background rate for protons, as sequential cuts are applied, for the *complex* geometry and accurate surface profile. A total of 1×10^7 protons with energies greater than 10 GeV are generated, representing 2.482 years worth of detector live time. As the protons are generated on a plane which measures 50×50 m², and the accurate surface profile was modified to have a flat area above the detector measuring 100×100 m², it is not necessary to propagate protons through the accurate surface profile. As discussed earlier, protons are generated on a plane measuring 50×50 m² as protons will not induce background events at the high inclinations which muons do. The increased overburden causes the overall number of background mimicking events to decrease substantially when using the *complex* detector geometry (1.55×10^4), as opposed to using the *simple* detector geometry ($(1.05 \pm 0.06) \times 10^5$).

It can be seen from Table 6.5 that the cut with respect to the beam angle, and the PoCA calculation with respect to all charged particles, are very effective as they remove 67% and 90% of all remaining background events respectively. The all charged PoCA cut is effective, since many of the photons produced will be close to the initial proton that was simulated. After all cuts are applied, the annual number of background events is seen to decrease by over a factor of 10 when using the *complex* geometry. This decrease is attributed to the additional shielding which is present in the *complex* geometry, as concrete, insulation and

Table 6.5 The normalised background rate per calendar year, for proton induced events, for the *complex* geometry. A total of 1×10^7 protons with energies greater than 10 GeV are generated, representing 2.482 years worth of detector live time. The background rate is separated into different first generation photon ancestries. The application of the cuts outlined in Section 6.1.2 is shown. Table is taken from [159].

	E_γ	$D > 30$ cm	$\theta_{beam}(E)$	$PoCA_{all}$	$e - \gamma(E)$	γ detection
<i>Total</i>	1.55×10^4	10559	3475 ± 39	319 ± 12	31.9 ± 1.2	0.23 ± 0.01
$\pi^0 \rightarrow \gamma$	1.18×10^4	9277 ± 63	3098 ± 36	297 ± 11	29.7 ± 1.1	0.21 ± 0.01
<i>Ext</i> $\rightarrow \gamma$	3120 ± 37	858 ± 19	279 ± 11	22 ± 3	2.2 ± 0.3	0.016 ± 0.002
<i>Other</i> $\rightarrow \gamma$	524 ± 15	424 ± 15	97 ± 6	$0 - 1.04$	$0 - 0.10$	$0 - 0.001$

Table 6.6 The normalised background rate per calendar year, for neutron induced events, for the *complex* geometry. A total of 1.1×10^8 neutrons with energies greater than 1 GeV are generated, representing 0.653 years worth of detector live time. The background rate is separated into different first generation photon ancestries. The application of the cuts outlined in Section 6.1.2 is shown. Table is taken from [159].

	E_γ	$D > 30$ cm	$\theta_{beam}(E)$	$PoCA_{all}$	$e - \gamma(E)$	γ detection
<i>Total</i>	8405 ± 113	5697 ± 93	1949 ± 54	225 ± 18	22.5 ± 1.8	0.16 ± 0.01
$\pi^0 \rightarrow \gamma$	6397 ± 98	5050 ± 87	1744 ± 51	194 ± 17	19.4 ± 1.7	0.14 ± 0.01
<i>Ext</i> $\rightarrow \gamma$	1796 ± 52	470 ± 26	169 ± 16	30.1 ± 6.7	3.01 ± 0.67	0.021 ± 0.005
<i>Other</i> $\rightarrow \gamma$	209 ± 18	175 ± 16	36.2 ± 7.4	$0 - 3.68$	$0 - 0.37$	$0 - 0.003$

inactive regions of LAr have been added.

Table 6.6 shows the background rate for neutrons, as sequential cuts are applied, for the *complex* geometry and accurate surface profile. A total of 1.1×10^8 neutrons with energies greater than 1 GeV are generated, representing 0.653 years worth of detector live time. The neutrons are generated on a plane which measures 50×50 m², and so as seen with the protons, it is unnecessary to propagate neutrons through the accurate surface profile. The overall number of background mimicking events is seen to decrease substantially when using the *complex* detector geometry, as opposed to using the *simple* detector geometry.

From Table 6.6 it can be seen that as was the case with the protons, the cut with respect to the beam angle, and the PoCA calculation with respect to all charged particles, are very effective. Upon the application of all cuts, the annual expected background from neutrons in the *complex* geometry is seen to decrease by almost an order of magnitude. Whilst this is not as dramatic as the reduction seen when considering proton induced backgrounds, it still represents a significant reduction in the background rate. This reduction is again attributed to the extra shielding which is present in the *complex* geometry.

Table 6.7 The normalised background rate per calendar year for the *simple* and *complex* detector geometries, separated by particle species, using the flat and accurate surface profiles respectively. The rates for muons entering the detector, muons missing the detector, proton induced events, and neutron induced events are shown. The annual background for muons entering the detector when using the *simple* detector geometry and flat surface profile is only an approximate value, as initial simulations were performed without saving proton hit information. The inclusion of proton hits greatly increases the accuracy of the all charge PoCA cut, however this was only seen in a small sample. Table is taken from [159].

Primary particle	<i>simp.</i> geom. & flat surf. prof.	<i>comp.</i> geom. & acc. surf. prof.
Muons entering	≈ 1.18	2.03 ± 0.24
Muons missing	0.11 ± 0.02	0.65 ± 0.08
Protons	2.57 ± 0.08	0.23 ± 0.01
Neutrons	1.23 ± 0.07	0.16 ± 0.01
Total	≈ 5	3.07 ± 0.25

6.1.5 Summary of simulations for the LBNE surface detector

Taking the sum of the expected background rates from Tables 6.3, 6.4, 6.5 and 6.6 gives an expected cosmogenic background after cuts for the *complex* detector geometry, and accurate surface profile, of 3.07 ± 0.25 events per year. This compares with an expected background rate of ≈ 5 events per year, for the *simple* detector geometry, and flat surface profile. The reduction in background rate is due to the additional shielding present in the *complex* geometry which reduces the hadronic background component from 3.80 ± 0.11 to 0.39 ± 0.01 , a decrease of roughly an order of magnitude. However, the larger surface area of the detector is found to cause the number of background events caused by muons to increase. This increase is very substantial when considering muons which miss the active volume of the detector. This is summarised in Table 6.7.

Following the simulation of large samples of muons, protons, and neutrons, the expected background for the LBNE surface detector is found when considering a *complex* detector geometry, and an accurate surface profile. This is compared to the expected background rate when using a *simple* detector geometry and flat surface profile [159]. It is observed that the effect of the *complex* detector geometry is to reduce the overall background rate, primarily due to the additional shielding which it provides. This additional shielding causes the proton and neutron induced backgrounds to decrease substantially, as was shown in Table 6.7. It is however found that there will be a significant source of backgrounds from muons which do not enter the active volume of the detector. This is due to the presence of vertical gaps between the TPC cells. The effect of incorporating the accurate surface profile is found

to be negligible, as the hills only offer minimal amounts of shielding. These simulations provide a relatively accurate estimate of the expected background for the LBNE surface design, although it must be stressed that these studies have been performed using Monte Carlo truth information, and so have not used reconstruction.

6.2 MUSUN in LArSoft

The primary muons in the following discussions are all generated using MUSIC [116, 118, 119] and MUSUN [116, 117], and so a brief overview of them is required. MUSIC first propagates muons through a medium, defined by the user, for given initial energies. A range of energies between 10^2 GeV and 10^7 GeV is considered, and their energy distributions are stored at depths of 100 to 15,000 m w.e. Energy losses due to four processes are considered; ionisation, bremsstrahlung, electron-positron pair production and muon-nucleus inelastic scattering. The output of MUSIC, along with the surface muon spectrum parameterisation and the surface profile, is then used by MUSUN to generate a muon energy spectrum and angular distribution, for a given detector location.

The location of the DUNE far detector, near the Ross shaft at SURF, has global coordinates of $44^\circ 20' 45.21''$ North, $103^\circ 45' 16.13''$ West. The rock composition is assumed to be, $\langle Z \rangle = 12.09$ and $\langle A \rangle = 24.17$. The density is assumed to be $2.70 \text{ g}\cdot\text{cm}^{-3}$ [172], though some measurements suggest that it may be closer to $2.80\text{-}2.90 \text{ g}\cdot\text{cm}^{-3}$ [173, 174]. The vertical flux calculated by MUSIC/MUSUN for the Davis Cavern of $5.18 \times 10^{-9} \text{ cm}^{-2}\cdot\text{s}^{-1}\cdot\text{sr}^{-1}$, is well matched to the flux measured by the active veto system of the Davis' experiment, which was $(5.38 \pm 0.07) \times 10^{-9} \text{ cm}^{-2}\cdot\text{s}^{-1}\cdot\text{sr}^{-1}$ [175]. Given the small differences in these values, and another measurement by the Majorana demonstrator [176], the systematic uncertainty in calculating the muon flux is estimated to be 20% [177].

The surface profile around the proposed detector location is shown in Figure 6.3a, where the proposed location is in the centre of the map. Each quadrant on the map has been divided into 4 angles of 22° to help guide the eye when comparing it to Figure 6.3b, where the distribution of azimuth angles is plotted. The vertical lines in Figure 6.3b show the division of the quadrants when the angle is calculated from East to the North. When moving from East to North it is possible to discern how the peaks and troughs on the surface profile, correspond to troughs and peaks in the distribution of azimuthal angle.

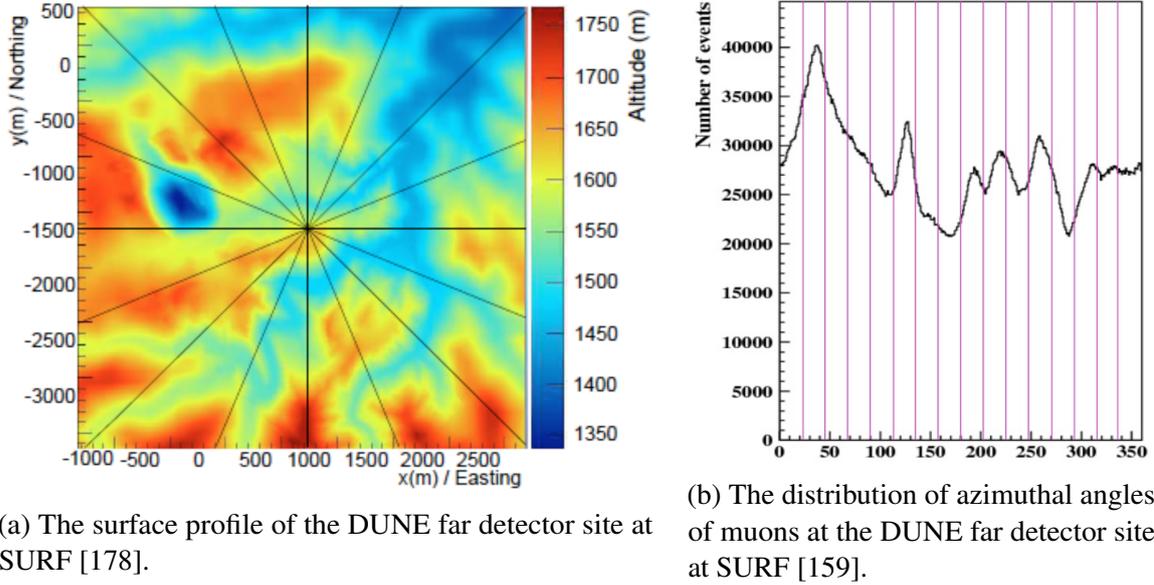


Fig. 6.3 The correlation between the surface profile, and the distribution of azimuthal angles at the DUNE far detector site. The quadrants have been divided into four angles of equal size. The azimuthal angle, calculated as the angle from East (pointing to the right in Fig. 6.3a), and increasing counterclockwise, is seen to follow the contours of the surface profile.

Table 6.8 Muon flux parameters as calculated with MUSIC/MUSUN.

Total flux ($\text{cm}^{-2}\cdot\text{s}^{-1}$)	Mean E_μ (GeV)	Mean slant depth (m w.e)	Mean θ ($^\circ$)
5.66×10^{-9}	283	4532	26

Given these parameters, the muon flux at the DUNE far detector location, when assuming a spherical detector geometry, and without simulating a detector cavern, is given in Table 6.8.

The muons simulated for DUNE are sampled on the surface of a box surrounding the detector hall, which also encompasses 7 m of rock above the cavern, and 5 m of rock on all other sides. This is to ensure that the simulated muons pass through a sufficient amount of rock to induce cascades, both above and around the detector hall. The secondaries produced in these cascades which enter the detector, in the absence of the initial muon, are of particular interest, as some of them could be mistaken for nucleon decay events. The study of these nucleon decay mimicking events is discussed in Section 6.3. The size of the box which the muons are sampled from is $74.43 \times 29.54 \times 30.18 \text{ m}^3$, compared to the simulated cryostat which has dimensions, $61.62 \times 14.94 \times 13.58 \text{ m}^3$. The dimensions are given as length \times width \times height, using the LArSoft coordinate system which was defined at the start of Section 2.5. The muons are sampled according to their energy spectrum, for a given zenith

and azimuthal angle, using the angular distribution obtained with MUSUN.

MUSUN has been incorporated into the DUNE software framework, as it had previously been maintained in FORTRAN as an external package. Before simulations in LArSoft were performed, it was ensured that the muon distributions produced by the ported LArSoft code were identical to the original distributions produced by the FORTRAN code. The distributions produced by the DUNE software framework are shown in Figure 6.4, and are consistent with the distributions made for the LBNE collaboration [179]. Figure 6.5 shows the initial positions of 10,000 muons around the simulated DUNE 10 kt module, as generated by LArSoft. The initial positions of the muons are shown as blue points, whilst the cryostat is a single black box, and each TPC is a red box.

It is found that the muon rate through the box upon which the muons are sampled is 0.1579 Hz. This rate is later used to normalise the background event rate in Section 6.3. Roughly a third of the muons which are generated, pass through the active volume, to give a muon rate through the active volume of 0.053 Hz.

The simulated far detector is contained in an excavated detector hall surrounded by rock. The cryostat is made of concrete, supported by a stainless steel structure. The inside of the cryostat is filled with LAr, which has a total mass of 17.1 kt. Contained within the cryostat is an active volume of LAr measuring $14.5 \times 12 \times 58 \text{ m}^3$, made up of 200 TPCs, to give a total mass of active LAr of 14.1 kt. This figure includes gaps between TPCs, when these are removed, the active mass becomes 13.9 kt of LAr. In Section 6.3 a fiducial cut of 2 cm is used, which requires that there are no energy depositions within 2 cm of the edge of the active mass. Following the application of this fiducial cut, the fiducial mass is reduced to 13.8 kt. When performing full energy reconstruction of signal events, this fiducial may have to be increased to ensure that events are fully contained.

6.3 Nucleon decay channels in DUNE

When searching for rare processes, where an experiment is unlikely to see more than a few signal events, an exhaustive study of the potential backgrounds is required. This is so that if a signal is observed, it could provide overwhelming evidence for the process. The search for nucleon decay in DUNE is one such process, and so an extensive study of the background to nucleon decay is required. As discussed in Section 1.2.2, cosmogenic muons cause

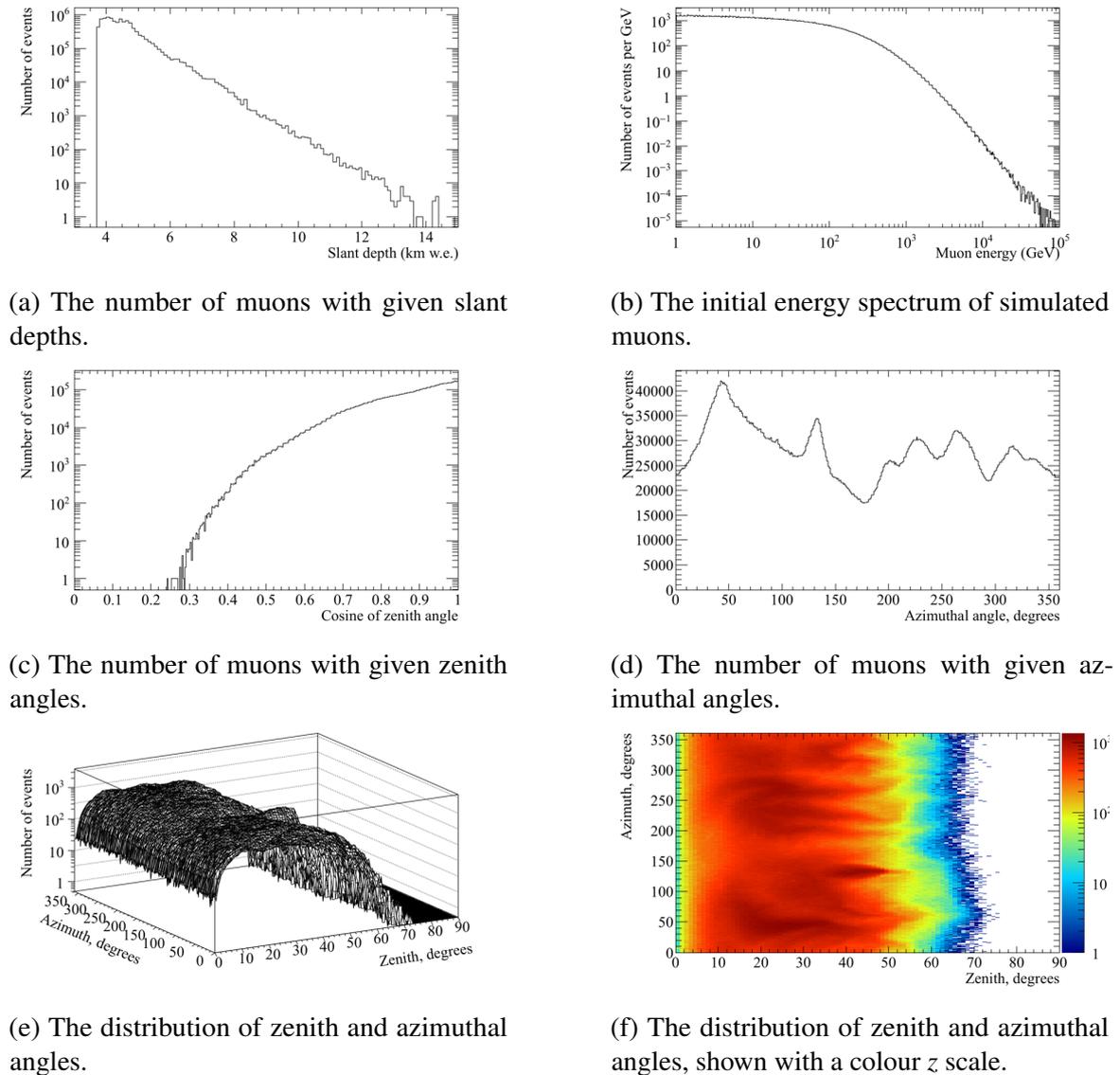


Fig. 6.4 The distributions of muon parameters for a sample of 10^7 muons generated by MUSUN in LArSoft. Top left: the slant depths of the simulated muons. Top Right: the energies of the simulated muons. Middle left: the zenith angles of simulated muons. Middle right: the azimuthal angles of simulated muons. Bottom left: the number of muons as a function of zenith and azimuthal angles. Bottom right: the number of muons as a function of zenith and azimuthal angles, as a coloured map.

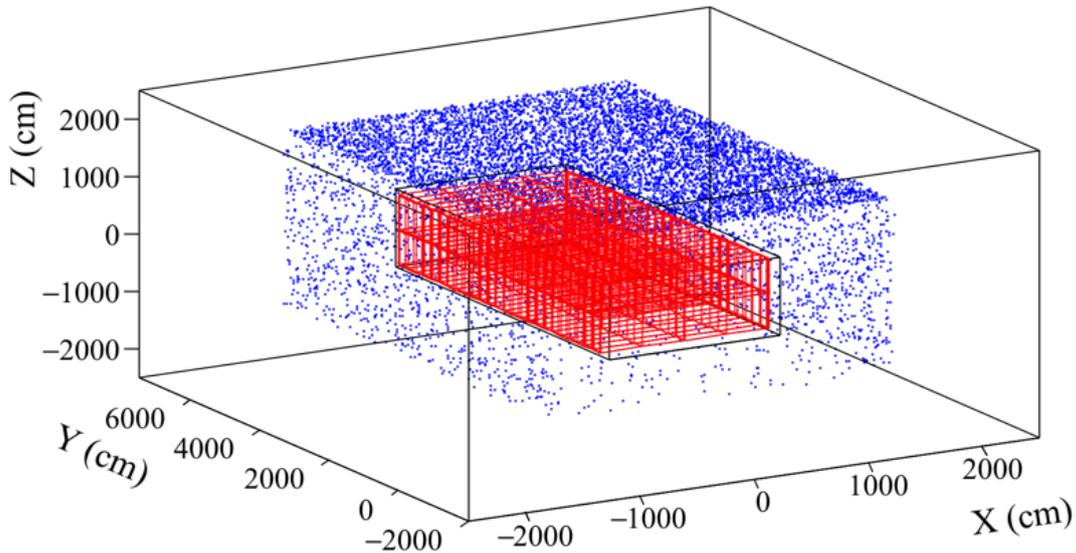


Fig. 6.5 The initial positions of 10^4 muons generated by MUSUN around a DUNE 10 kt module. The initial positions of the muons are shown as blue points, whilst the cryostat is a single black box and each TPC is a single red box.

backgrounds to nucleon decay events, as some of the secondary particles produced by their interactions are able to mimic the nucleon decay signatures. For this reason it is necessary to simulate this background, and to develop a series of cuts which can be applied to the energy depositions which they produce, to establish that they are not due to nucleon decays. When doing this, it is important to use a simulated cosmogenic flux that is as accurate as possible to the one which will be observed by the DUNE far detector. For this reason MUSUN was incorporated into LArSoft, as the muons that it generates correspond to the realistic surface profile, as described in Section 6.2. As discussed in Section 1.2.2, atmospheric neutrinos may also mimic nucleon decays, though they are not considered in this thesis as they were found to be negligible for the LBNE experiment [60].

The DUNE experiment will run for more than 20 years, and so the required statistics for background studies should be even higher than this, so as to ensure the significance of the results. Therefore, a sample of 2×10^9 muons has been generated. Given that the muon flux through the simulated box is 0.1579 Hz, this sample corresponds to 401.6 years of detector live time for a single DUNE FD module. This sample is thus equivalent to 100.4 years of detector live time for the full DUNE design consisting of four FD modules.

Producing samples of this size requires significant computer power, both in terms of running time, and storage space. As a result of this, many of the simulated events are discarded before being saved to disk. This is done through the application of a filter after GEANT4 [108], where events satisfying one of the following cuts are discarded;

- Contains a muon track of more than 1 m.
- There are no energy depositions in the entire detector volume.

This way, events which could mimic nucleon decay signals are not removed from the analyses performed on the large muon sample. In applying these filters it has been assumed that a muon track of more than a metre would definitely be reconstructed. It is also assumed that any signatures observed within one drift window of such a track would not be studied in a nucleon decay search, as there would be doubt as to the authenticity of the signal. Given that the total rate of muons through the active volume is 0.053 Hz, and that the drift time is a few ms, ignoring all times where any track from a cosmogenic muon is present results in less than 0.1% dead time. The dead time associated with ignoring events with muon tracks of more than 1 m is clearly less than this. This fraction of dead time is assumed to be acceptable. Filtering out events where there are no energy depositions in the detector is clearly acceptable, as there are no energy depositions which could mimic a nucleon decay signature.

After applying this series of cuts, 98.12% of the initial muon sample are removed, meaning that the initial sample of 2×10^9 muons is reduced to around 4×10^7 muons. This is a much more reasonable sample size to store on disk, and to perform analyses on. It is upon this reduced sample of muons that the cosmogenic background analyses are performed. As discussed in Section 2.3.2, the proton decay channel of $p \rightarrow K^+ + \bar{\nu}_e$ is referred to as the ‘‘Golden Channel’’ in LAr, this analysis is discussed in [177]. The related decay of a neutron in the decay $n \rightarrow K^+ + e^-$ is discussed here. The theoretical motivation for this channel was briefly discussed in Section 1.2.1.

6.3.1 Cosmogenic background to the $n \rightarrow K^+ + e^-$ decay channel

As was shown in Table 2.3, the predicted sensitivity that DUNE will have to this channel is much better than that of Super-K. As a result, it is an interesting decay mode to study. As discussed in Section 1.2.2, the cosmogenic background to nucleon decay is predominantly caused by neutral particles, such as a K^0 , entering the detector volume, and interacting

relatively far from the detector edges. This is particularly true for the “Golden Channel,” as shown in Figure 1.2, but it also holds for other channels. Events like this are the main cause for concern when eliminating all cosmogenic backgrounds. As mentioned in Section 1.2.2, it is difficult to separate a K^+ from a K^- in a LArTPC, and so any charged kaon is considered a background in the analysis which is presented in this thesis.

The analysis presented in this thesis has been performed on Monte Carlo truth information, and so does not contain any reconstructed quantities. Studies involving hit and track reconstruction are in progress [180], though they will not be discussed here. As Monte Carlo truth information has been used, perfect particle identification (PID) has been assumed. There was also no smearing of the energies, locations, or trajectories of any simulated particles, meaning that it is assumed that all deposited charge will be reconstructed, and that the detector characterisation is perfect. The energy cuts which are applied in Section 6.3.4 do allow for energy smearing to be taken into account though, as they can easily be made wider than the true distributions. The omission of smearing is something which will need to be refined in future analyses, and will be taken into account when the analysis progresses to use reconstructed quantities, as discussed in Section 6.3.5.

As is the case with the “Golden Channel,” the final state of the decay contains a single charged kaon, and so events which do not contain a kaon track can be immediately discounted. There is also an electron in the final state of the decay, and so this means that events which do not also contain an electron can be discounted. In a nucleon decay event, the kaon and electron produced in the final state will have a common vertex, and so the requirement that the two particles have a common vertex can also be applied. Other constraints that are applied to eliminate background events are; a cut on external muon track length in the active LAr, a cut on depositions near the detector edges, and criteria about the distribution of deposited energy. The criteria about the distribution of deposited energy are found by considering a sample of simulated neutron decay events, and are discussed in Section 6.3.4. These cuts, applied sequentially, are outlined below:

- The event contains energy depositions which are due to kaons or their daughters, and energy depositions which are due to electrons or their daughters.
- The event contains at least one kaon track, and at least one electron track/shower.
- The event contains a single kaon track, and a single electron track/shower.
- No muon which is externally produced travels more than 20 cm in the active detector volume.

- The event has no energy depositions within 2 cm of the detector edges.
 - This is changed to a maximum of 10 MeV of energy deposited within 2 cm of the detector edge, for reasons which will be discussed in Section 6.3.2.
- The kaon and electron share a common vertex, defined as:
 - The kaon and electron tracks being separated by no more than 5 cm.
 - If the kaon and electron tracks are separated by more than 5 cm, then the point of closest approach between the two extrapolated tracks is less than 2 cm.
- The energy depositions in the event are within the ranges expected from a nucleon decay event. This is explained in Section 6.3.4, but the energies considered are summarised below:
 - The energy directly deposited by the kaon and its secondaries, excluding its decay products.
 - The energy deposited by the kaon decay products and any of their secondaries.
 - The energy directly deposited by the electron and its secondaries.
 - The energy deposited near the shared kaon and electron vertex that is not associated with the kaon or electron.
 - The energy deposited in the detector which does not fit any of the above criteria.

A key limitation in the application of these cuts is the assumption that track and shower reconstruction, as well PID, can be performed with perfect accuracy. This is clearly unlikely to be true in the DUNE FD, as reconstruction failures such as missing track segments or the mis-identification of particles is unavoidable. One of the most likely forms of particle mis-identification will be mistaking pions for kaons. Studies to understand the extent of this failure given the current state of the reconstruction algorithms are only just beginning [181–184]. However, it is still useful to perform an entirely truth based study to determine whether the signal from a given decay channel will exist above the background signal. Future improvements to this study will incorporate the reconstruction algorithms into their method.

When performing the analysis it is important to be able to trace the particle ancestry. This is so that energy depositions in the detector can be properly assigned to the relevant particles. For example, a μ^+ is often produced when a K^+ decays at rest, and this muon may travel more than 20 cm. However, the cut on muon track length should not be applied to this muon as it was produced by the decay of the kaon. Similarly, if the kaon interacts in the detector, secondary particles will be produced which will be reconstructed as tracks coming off the

main kaon track. The initial kinetic energy of the kaon can only be determined by summing the energy depositions due to these secondary particles, and the energy depositions due to the kaon itself. Correctly calculating the initial kaon kinetic energy is critical when determining if an event is a nucleon decay event. The reason for this is that nucleon decay events have very specific energy spectra, and so being able to correctly assign the ancestry of energy depositions is vitally important. The same is true for the electron energy, which is calculated by tracing the ancestry of the particles produced in the EM shower which it produces back to the electron.

As no reconstruction has been performed, the tracks referred to here are different from those in Chapters 4 and 5. The definition of a track used here, is that the particle in question has energy depositions, on simulated wires, which are directly associated with it. These simulated wires are not the same as the wires which have been considered in Chapter 4, as the simulated signals have not been digitised. This distinction is important, as it allows the energy depositions directly from GEANT4 to be used, whilst also allowing for LArSoft methods concerning whether depositions are within TPC boundaries, to be utilised.

The simulated electrons may begin showering immediately, or they may produce a short “track like” segment before beginning to shower. To ensure that every electron shower can be identified, electrons are not required to produce a short “track like” segment. This means that all electrons are assumed to begin showering immediately, and it is also assumed that all of the energy in the shower can be identified as coming from a single electron. This definition of shower energy is the one that is used when the showering algorithms are developed in LArSoft. It is hoped that when DUNE begins taking data the showering algorithms will be able to achieve fully accurate showering reconstruction, though this has not yet been achieved.

When calculating the distance between the start of the kaon track, and the start of the electron track/shower, the energy depositions whose locations are closest to the Monte Carlo truth start points of the particles are used. For particles which are produced within the active volume, these locations generally correspond to the Monte Carlo truth start positions, though this is not always the case. For example, if a particle is created in the gap between two TPCs, then there will be no charge collected until it enters the active volume. This will result in the measured start position being shifted from the true generation point. This shift can prove troublesome when considering decay events, as if the decay occurred in the centre of an APA, then it is likely that the kaon and electron would deposit energy on opposite sides of the APA. This could cause the depositions to be separated by over 5 cm, as this is the width

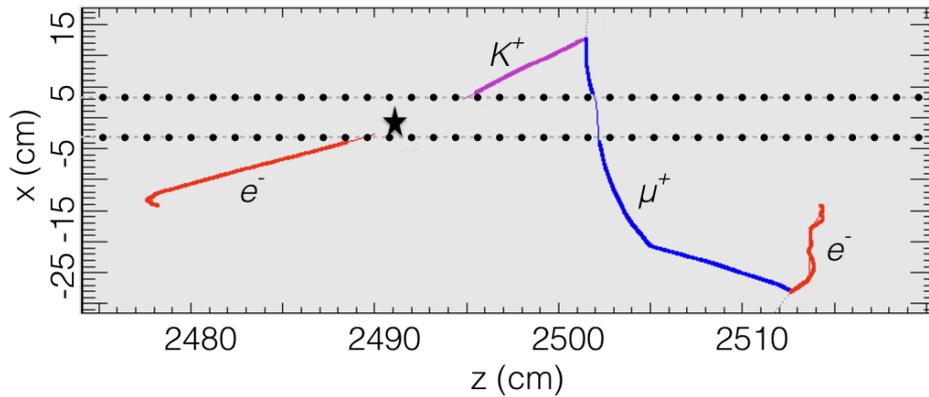


Fig. 6.6 A simulated $n \rightarrow K^+ + e^-$ decay which occurred in a gap between TPCs. The path of the kaon produced in the decay is shown as a purple line. The path of the muon, produced by the decay of the kaon, is shown as a blue line. The paths which the electrons in the event took are shown as red lines. The electron on the left of the figure is the electron produced in the neutron decay, whilst the one on the right of the figure, is produced by the decay of the muon. The thin coloured lines show track segments which were in uninstrumented parts of the detector, such as gaps between TPCs and APAs. The dotted black lines show the edges of the TPCs, and the black star shows the location at which the decay occurred. The distance between the first kaon energy deposition, and the first electron energy deposition, is found to be 10.7 cm. However, when the kaon and electrons tracks are extrapolated towards the true start position, the point of closest approach (PoCA) between the two tracks is found to be 0.67 cm. This shows that they do in fact have a common vertex, despite the large separation of the start points.

of the APAs. However, if the tracks are propagated backwards, towards their true start point, it should still be possible to determine that they had a common vertex. An example of a simulated decay event where this happens is shown in Figure 6.6.

In order for a kaon track and an electron shower to be considered to share a common vertex, the separation between the start of the kaon track and the start of the electron shower, must be no more than 5 cm. A maximum separation of 5 cm is used, as, if the two particles are produced in the centre of a TPC, a gap of 5 cm would require no energy depositions to be collected over approximately 10 collection wires. This is assumed to be unlikely during data taking, and cannot happen in the simulations considered here, as Monte Carlo truth information is used. However, as shown by Figure 6.6, it is possible for the kaon and electron produced in a signal event to be separated by more than 5 cm. To prevent events such as this being missed, a second criterion is applied to events with large separations. This criterion is that the “Point of Closest Approach” (PoCA) between the two particles, found by extrapolating the kaon track and the electron shower forwards and backwards,

is less than 2 cm. This is the same PoCA calculation which was made in Section 6.1.2, and it means that events such as the one shown in Figure 6.6 are still identified as signal events.

The fiducial cut is only applied to the outer edges of the cryostat, as if it were done with respect to the edge of every TPC in the far detector the loss of volume would be non-negligible. This means that the event shown in Figure 6.6 would not fail the fiducial cut, as the decay occurred over 6 m away from the edge of the detector, but happened to be in a gap between two TPCs. The need for a fiducial cut is two fold, firstly the vast majority of cosmically induced events in the detector will have a charged particle which enters the detector. Performing a fiducial cut will remove all of these events, and will then mean that the only cosmic background events which can mimic a signal event would involve either a significant amount of charge being missed, or a neutral particle entering the detector, and interacting relatively far from the detector walls. Secondly, in order to calculate the kinetic energies of the particles produced in the nucleon decay, and also to perform particle identification, they must be fully contained within the detector. As such, if one of the particles produced in the nucleon decay escapes the detector then its kinetic energy cannot be determined accurately, and if it is the kaon, or its decay products, then the particle cannot be identified using the method discussed in Section 4.4. This will also affect any particles which stop in the gaps between TPCs, as the end point will not be reconstructed, though PID may still be possible as the end point can be estimated. A fiducial cut of 2 cm is used, as the loss of active volume which this causes is negligible in the DUNE FD. A 2 cm fiducial cut also ensures that a significant amount of charge would have to be missed for a particle which enters/escapes the detector to be incorrectly identified as being contained within the detector.

Once both the ancestry of energy depositions in the simulation, and the initial kinetic energies, have been correctly accounted for and calculated, it is possible to observe the distribution of background events as the cuts outlined above are applied. The energy distribution of background events surviving the application of sequential cuts is shown in Figure 6.7. The energy distribution of background events per MeV of energy deposited, surviving the application of sequential cuts is shown in Figure 6.8. This distribution is obtained by dividing the number of events within an energy bin by the bin width.

From Figures 6.7 and 6.8, it can be seen that there are no background events which could mimic a decay signature, as no events survive the application of all cuts. This corresponds to a limit of the background rate of less than $0.44 \text{ events} \cdot \text{Mt}^{-1} \cdot \text{year}^{-1}$ at the 90% confidence level, using double sided errors [26] and a fiducial mass of 13.8 kt from an exposure of

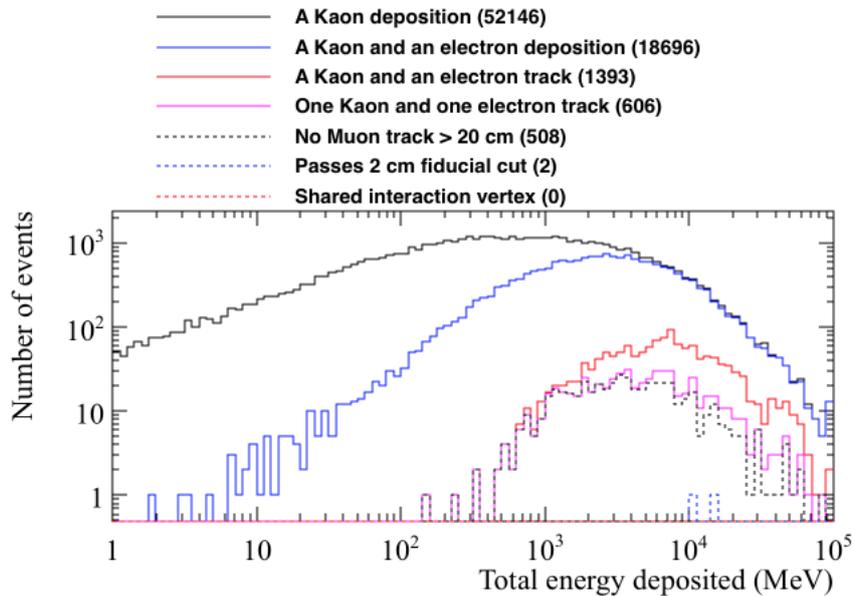


Fig. 6.7 The energy distribution of background events surviving the application of sequential cuts. The total energy deposited in the detector is plotted on the x axis. A sample of 2×10^9 muons, representing 401.6 years of detector for a single FD module is shown.

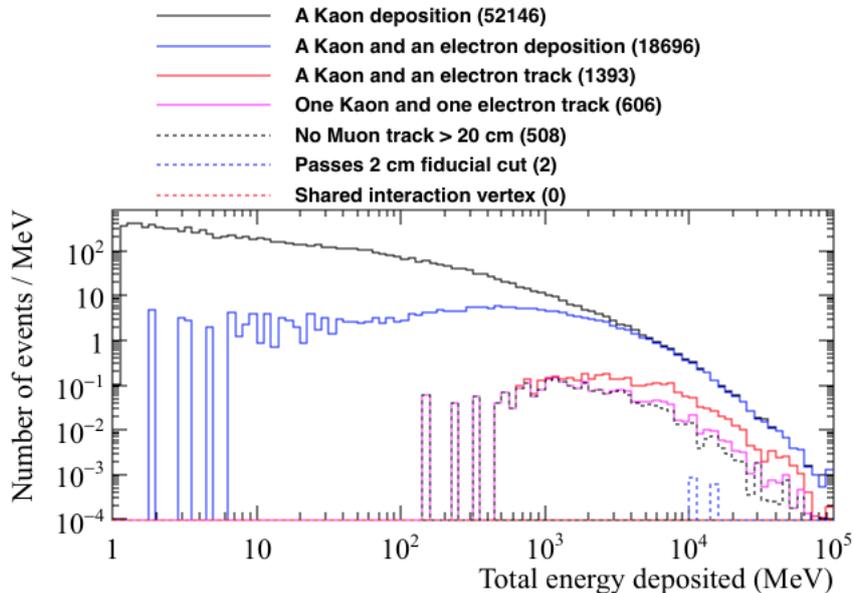


Fig. 6.8 The energy distribution of background events per MeV of deposited energy surviving the application of sequential cuts. The total energy deposited in the detector is plotted on the x axis. This distribution is obtained by dividing the number of events within an energy bin by the bin width. A sample of 2×10^9 muons, representing 401.6 years of detector for a single FD module is shown.

Table 6.9 The number of events which could mimic a $n \rightarrow K^+ + e^-$ decay, when cuts are applied in isolation. Cuts are applied after it is required that the event contains at least one kaon track and at least one electron shower. It is found that 1393 events have at least one kaon track and at least one electron shower, this is shown in the top row of the table. The fiducial cut of 2 cm is seen to remove almost all of the events considered.

Cut that is applied	Num. events surviving cut	% surviving
At least one kaon track, and electron shower	1393	100
Only one kaon track, and only one electron shower	606	43.5
No muon track that is longer than 20 cm in length	1223	87.8
No energy depositions within 2 cm of detector edge	5	0.359
The kaon and electron share a common vertex	64	4.59

5.542 Mt·year.

It is interesting to observe the effect that relaxing some of the cuts has on the background rate. For example, the cuts after the requirement that there be at least one kaon track and at least one electron shower in the event, could be relaxed. This is shown in Table 6.9, where the later cuts have been applied in isolation to observe their effectiveness.

The effectiveness of the fiducial cut is clearly apparent from Table 6.9, as it removes all but 5 of the 1393 events where there is both a kaon track and an electron shower. The requirement that the kaon track and electron shower share a common vertex is also seen to be very effective at removing background events, as only 64 of the 1393 meet this condition. When the cuts are no longer applied in isolation, but are instead applied in a different order to that used when producing Figures 6.7 and 6.8, it is found that in only 15 of the 606 events that have a single kaon track and a single electron shower, would the kaon and electron be considered to have a common vertex. When the additional constraint of there not being an external muon with a track length of more than 20 cm present in the detector is applied, 10 of the remaining 15 events background events would still not be removed. This shows that the only way to remove all of the background events is to apply all of the cuts which have been developed, including the fiducial cut.

6.3.2 Signal events in the $n \rightarrow K^+ + e^-$ decay channel

It is important to confirm that the cuts developed to reject cosmic backgrounds do not adversely affect the identification of nucleon decay events. For this reason, a sample of

10,000 neutron decay events in the DUNE far detector were generated using GENIE version 2.12.2 [109], and the so-called *hA* model of intranuclear effects. Studies have been done comparing the different models for intranuclear effects in GENIE [185], and future nucleon decay studies may be modified to use the so-called *hN2015* model. The *hA* model does not simulate any K^+ charge exchange or absorption, whereas the *hN2015* model does. As a result, the *hN2015* model predicts that more protons and neutrons will be emitted from the nucleus in final state interactions [185]. Neutron decays are generated at random positions within the detector volume, and so it is possible that the decay occurs in the gaps between TPC volumes, as shown in Figure 6.6, or near the edge of the detector, as is shown in Figure 6.9. However, many of the neutron decay events are fully contained within a single TPC, as shown by Figure 6.10.

The analysis performed on the cosmogenic background was primarily designed to reject background events, whilst also attempting to not use cuts which would also affect signal efficiency. Therefore, it is hoped that the loss of signal events will be minimal. When running the analysis on the simulated signal events, the same definitions for tracks, showers, and the ancestry of particles are used, as well as the same cuts that were outlined in Section 6.3.1.

The energy distribution of the signal events surviving the application of the sequential cuts is shown in Figure 6.11, this is the equivalent of Figure 6.7 for the cosmogenic background sample. The energy distribution of signal events per MeV of energy deposited, surviving the application of sequential cuts is shown in Figure 6.12, this is the equivalent of Figure 6.8 for the cosmogenic background sample. As before, this distribution is obtained by dividing the number of events within an energy bin by the bin width.

When comparing Figures 6.11 and 6.12 with Figures 6.7 and 6.8, the most obvious difference is that when considering the nucleon decay events, the total energy deposited in the detector never exceeds 1 GeV, whilst in the cosmogenic background sample, the energy deposited in the detector frequently exceeds 1 GeV. This is something which one would expect, as the simulated neutrons decay at rest, and so have a total energy of less than 1 GeV, meaning that there cannot be more than 1 GeV deposited in the detector. This is in stark contrast to the cosmogenic background, where the primary muons being generated have a mean energy of 283 GeV, as shown in Table 6.8. This means that many events will deposit significant amounts of energy in the detector, even if the primary muon misses the detector volume.

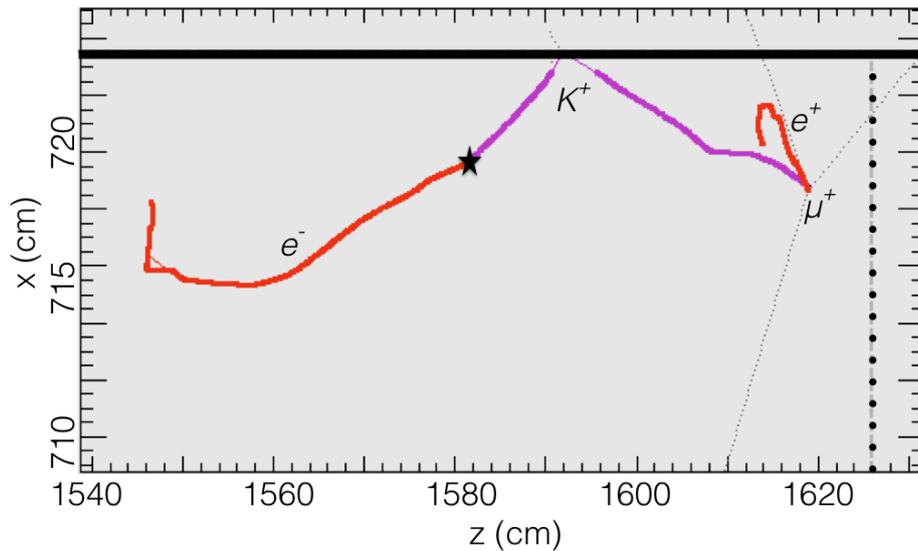


Fig. 6.9 A simulated $n \rightarrow K^+ + e^-$ decay which occurred near the edge of the detector volume. The path of the kaon produced in the decay is shown as a purple line. The path of the muon, produced by the decay of the kaon, is shown as a blue line, though it is very short and so barely visible. The paths which the electrons in the event took are shown as red lines. The electron on the left of the figure is the electron produced in the neutron decay, whilst the one on the right of the figure, is produced by the decay of the muon. The thin grey lines show neutral particles such as neutrons or neutrinos. The thin coloured lines show track segments which were in uninstrumented parts of the detector, such as the edge of the active volume. The solid black line shows the edge of the detector. The dotted black lines show the edges of the TPCs, and the black star shows the location at which the decay occurred. It can be seen that though most of the energy depositions are contained within the detector, the kaon passes very close to the edge of the detector, and so some of its charge is not reconstructed. The proximity of the decay to the detector walls causes there to be a large amount of charge deposited near the edge of the detector.

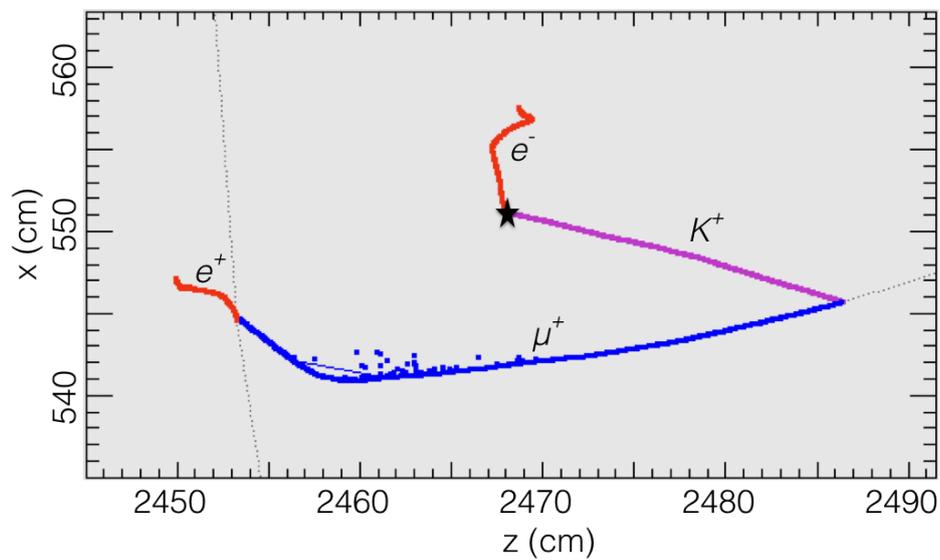


Fig. 6.10 A simulated $n \rightarrow K^+ + e^-$ decay which is fully contained in a single volume. The path of the kaon produced in the decay is shown as a purple line. The path of the muon, produced by the decay of the kaon, is shown as a blue line. The paths which the electrons in the event took are shown as red lines. The electron at the top of the figure is the electron produced in the neutron decay, whilst the one at the bottom left of the figure, is produced by the decay of the muon. The thin grey lines show neutral particles such as neutrons or neutrinos. The black star shows the location at which the decay occurred. It can be seen that all of the deposited charge is contained within a single TPC.

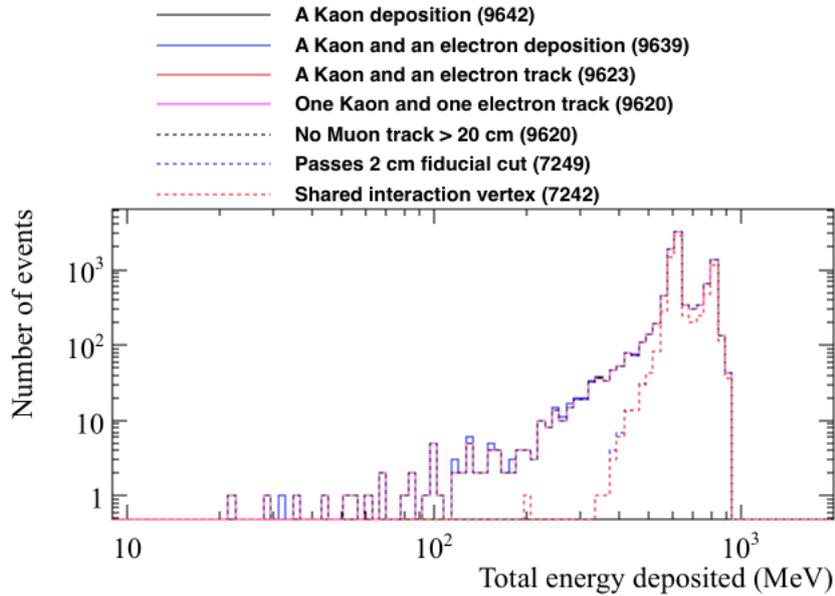


Fig. 6.11 The energy distribution of signal events surviving the application of sequential cuts in the $n \rightarrow K^+ + e^-$ channel. The total energy deposited in the detector is plotted on the x axis.

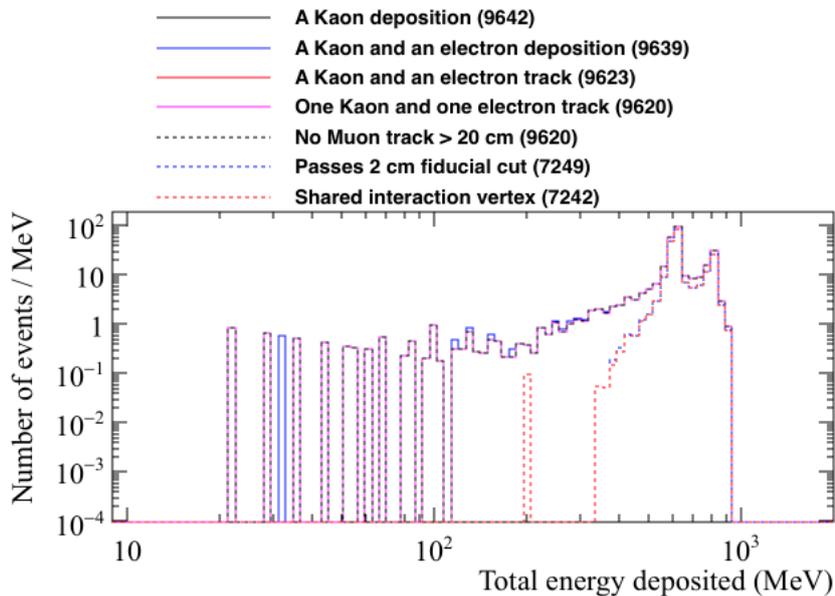


Fig. 6.12 The energy distribution of signal events per MeV of deposited energy surviving the application of sequential cuts. The total energy deposited in the detector is plotted on the x axis. This distribution is obtained by dividing the number of events within an energy bin by the bin width.

Table 6.10 The most common decay modes of charged kaons, and their probabilities [26]. The decay modes shown are with reference to K^+ , though the decays of K^- are the charge conjugations of these decays.

Decay mode	Measured probability (%)
$K^+ \rightarrow \mu^+ + \nu_\mu$	63.56 ± 0.11
$K^+ \rightarrow \pi^+ + \pi^0$	20.67 ± 0.08
$K^+ \rightarrow \pi^+ + \pi^+ + \pi^-$	5.583 ± 0.024
$K^+ \rightarrow \pi^0 + e^+ + \nu_e$	5.07 ± 0.04
$K^+ \rightarrow \pi^0 + \mu^+ + \nu_\mu$	3.352 ± 0.033
$K^+ \rightarrow \pi^+ + \pi^0 + \pi^0$	1.760 ± 0.023

A striking feature of Figures 6.11 and 6.12, is the double peaked structure of the total energy deposited in the detector. This can be attributed to the amount of energy from the kaon decay which is not reconstructed. The relative strengths of these structural features is due to the various probabilities of the kaon decay modes. Some of the more common decay modes are shown in Table 6.10, along with their probabilities. The amount of energy which is reconstructed from the particles produced by the decay of the kaon is shown in Figure 6.13.

As can be seen in Table 6.10, the most common decay mode for the kaon involves the production of a ν_μ , which will leave the detector without interacting. As the kaon is assumed to decay at rest, a significant proportion of the kaon rest mass energy will not be seen. This means that the most likely amount of energy deposited by the kaon decay products is around 200 MeV. The second most likely decay mode is also a 2-body decay, though some of the kaon rest mass energy will not be seen when the π^+ decays. This is why the second peak at 400 MeV is less than the rest mass of the kaon. Neutrinos are also present in many of the other decay modes, and so will also have large amounts of missing energies, whilst the 3-body decay modes will produce particles which have a continuous distribution of energies. These decay modes cause the energy deposited by the kaon decay products to be a continuous distribution, with two large peaks at roughly 200 MeV and 400 MeV, due to the 2 two body decays illustrated above.

The initial cuts, requiring that both a kaon track and an electron shower are observed in the decay, show that there are occasions when either the kaon, or the electron, do not deposit energy in the detector. This affects very few events, though an example of one such event is shown in Figure 6.14, where it can be seen that the kaon decayed before entering the active volume, and so no track was found for it. It would be very difficult to identify this

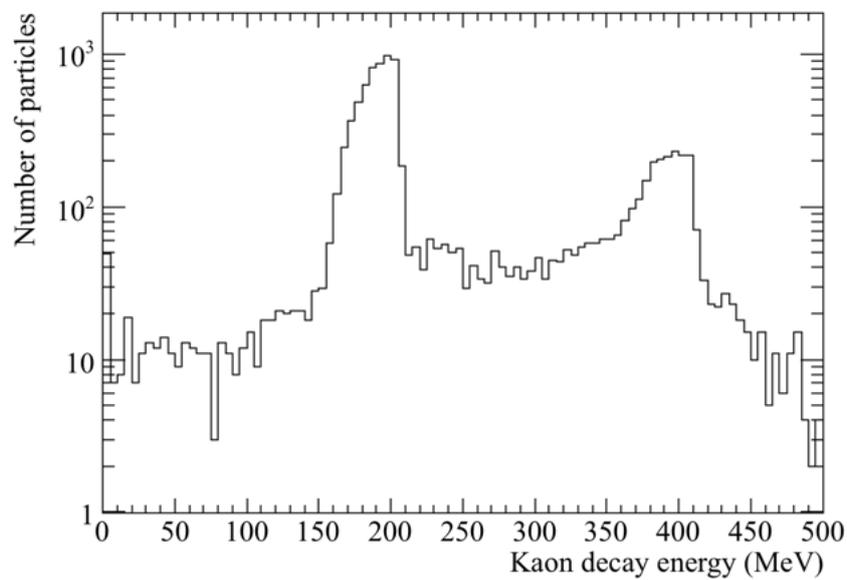


Fig. 6.13 The number of events as a function of the energy deposited by the kaon decay products. The energies deposited due to the $K^+ \rightarrow \mu^+ + \nu_\mu$ and $K^+ \rightarrow \pi^+ + \pi^0$ decay channels can be seen in the peaks at roughly 200 MeV, and 400 MeV, respectively. The underlying continuous distribution of depositions is caused by the 3-body decays shown in Table 6.10.

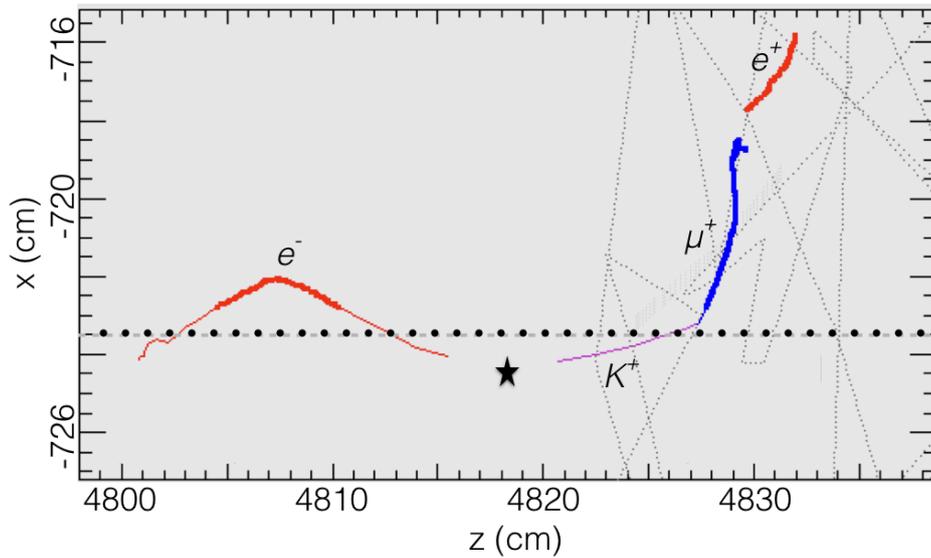


Fig. 6.14 A simulated $n \rightarrow K^+ + e^-$ decay where the kaon did not deposit any energy in the active volume. The path of the kaon produced in the decay is shown as a purple line. The path of the muon, produced by the decay of the kaon, is shown as a blue line. The paths which the electrons in the event took are shown as red lines. The electron on the left of the figure is the electron produced in the neutron decay, whilst the one on the right of the figure, is produced by the decay of the muon. The thin coloured lines show track segments which were in uninstrumented parts of the detector, such as gaps between TPCs and APAs. The dotted black lines show the edges of the TPCs, and the black star shows the location at which the decay occurred. It can be seen that the kaon decayed before it entered the active volume of the detector, and so no track was found for it. A significant portion of the distance which the electron from the decay travelled was also outside the active volume of the detector.

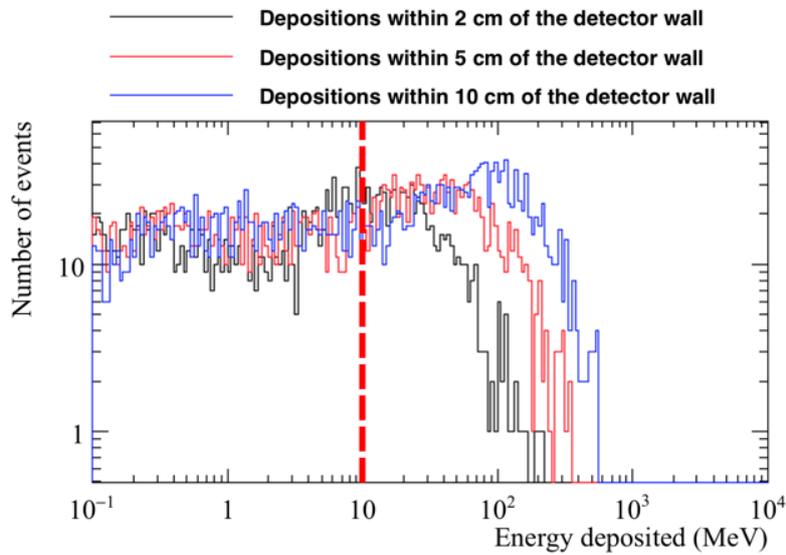
event as a signal event, as the presence of the kaon could only be inferred from the muon originating from the gap between the APAs, and the energy deposited by the kaon itself will not be reconstructed. Further compounding the identification of this event as a signal event, is the fact that the electron produced in the nucleon decay scatters back into the gap between the APAs. This means that a significant amount of the rest mass energy of the neutron would not be reconstructed.

The number of events which are removed by the fiducial cut is concerning, as when it is considered in conjunction with other cuts, it removes almost 25% of events. This suggests that the cut is too strict, as it would severely impair the prospect of identifying a real nucleon decay signal. The reason for this is two fold. Firstly, protons and neutrons are emitted from the nucleus in many of the simulated decays, and whilst the protons produced will create relatively short tracks, which are connected to the decay vertex, the neutrons will travel large

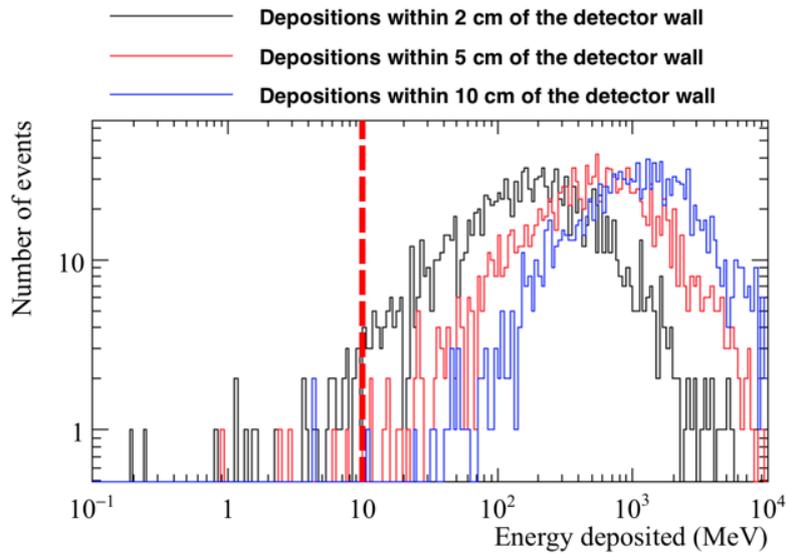
distances, and cause energy depositions which are far away from the decay vertex. The faint grey dashed lines, which can be seen in Figure 6.14, show neutral particles such as spallation neutrons or neutrinos which can be produced in the decay. None of the figures shown here contain spallation protons, but if they were present, they would be shown as additional purple lines originating from the decay vertex. In a large number of events, energy depositions from the spallation neutrons are causing events to fail the “no energy depositions within 2 cm of the detector edge” cut by depositing very small amounts of energy close to the detector walls. Secondly, though the decays are randomly distributed in the active volume of the detector, some of these decays will occur close to the detector walls. For example, over 5% (25%) of the neutron decays occur within 30 (100) cm of the detector walls. An example of one such event is shown in Figure 6.9, where the kaon produced in the decay deposits a significant amount of energy close to the edge of the detector. As a result, it is likely that this cut needs to be relaxed to instead be a cut on the amount of energy deposited within 2 cm of the detector edge. Figure 6.15 shows the amount of energy deposited within 2 cm, 5 cm, and 10 cm of the detector edges, for the simulated nucleon decay events (Figure 6.15a), and the cosmogenic background events (Figure 6.15b).

As can be seen from Figure 6.15, the amount of energy deposited near the detector edges is very different in the nucleon decay sample, and the cosmogenic background sample. As such, they can be relatively cleanly separated. In many of the signal events less than 0.1 MeV of energy is deposited near the detector edge, and so these events are not shown here. It can be seen that in some signal events there are significant energy depositions near the detector edge. Figure 6.9 shows an event where the kaon and electron shower deposit large amounts of energy near the detector edge. The energies deposited near the detector edge by the background events is very large, and is generally due to the presence of large showers which enter the detector.

A cut demanding that there should be no more than 10 MeV of energy deposited within 2 cm of the detector edge is used. This cut is designed to preserve as many of the simulated signal events as possible, whilst rejecting as many background events as possible. This cut removes only 683 of the 9,623 signal events, whilst only 39 out of the 1,393 cosmic background events meet this requirement. Whilst this does cause the fiducial cut to become less effective at removing background events from the cosmic sample, it does not cause the huge loss of signal events seen when using the hard cut of “no energy depositions within 2 cm of the detector edge.” It is also important to remember that the other cuts will still be applied to these 39 events, as the fiducial cut is not applied in isolation. As will be seen in



(a) The nucleon decay events.



(b) The cosmogenic background events.

Fig. 6.15 The number of events, as a function of the energy deposited within 2 cm, 5 cm, and 10 cm of the detector edges. Top: energy depositions for the simulated neutron decay events. Bottom: energy depositions for the simulated cosmogenic background events. The histograms are filled after the cut requiring that there is at least one kaon track and at least one electron shower in the event, is applied. As such, the histograms are filled with 9,623 and 1,393 events for the top and bottom histograms respectively. The dashed red line shows the effect of a cut on the energy deposited of 10 MeV. This cut will remove all events which are to the right of the line.

Figure 6.17b none of the muon-induced background events survive the application of all cuts.

The definition used to decide if the kaon track and the electron shower share a common vertex seems to be a reasonable requirement, as almost all of the signal events satisfy this definition. Figure 6.16 shows the distance between the start of the kaon track and the start of the electron shower in signal events. It can be seen that the separation between the two particles is very small. This is shown by the peak at events with separation of 0.1 cm, where events with separations smaller than this have been plotted in this bin. However, there are some events where the separations are very large (>10 cm). As discussed earlier, the decays in these events occur in the gaps between TPCs, such as that shown in Figure 6.6. When the requirement that the kaon track and the electron shower have a PoCA of less than 2 cm is then applied to these events, most of them are then seen to have a common interaction vertex. For example, this is found to be the case for the event shown in Figure 6.6. However, in some events the kaon track and the electron shower are still not found to have a common interaction vertex. Hand scanning shows that the particles in these events undergo scattering before entering the active volume, and so by the time that energy depositions occur, their trajectories are no longer closely aligned. This causes the PoCA to be larger than 2 cm, and so they are not identified as a signal event.

Figure 6.17 shows the energy distributions per MeV of deposited energy for the simulated decay (Figure 6.17a) and cosmogenic background (Figure 6.17b) events, after the fiducial cut is changed to require that there is less than 10 MeV of energy deposited within 2 cm of the detector edge. As before, these distributions are obtained by dividing the number of events within an energy bin by the bin width.

The number of signal events which are removed by the cuts can be seen to be much more reasonable after the fiducial cut is modified, as just 1,074 (10.74%) of the 10,000 signal events are removed as cuts are applied. It is seen that in many of the simulated decay events which fail the cuts, at least one of the kaon, or the electron, are not contained in the active volume of the detector. This means that, either no depositions are found for at least one of the particles (361 events), or there are large energy depositions close to the edge of the detector (683 events).

It can be seen that none of the cosmic background events pass the application of all cuts. Of the 24 events which pass the fiducial cut, only 9 of them deposit less than 1 GeV in the detector. This means that the total energy deposited in the detector is usually more

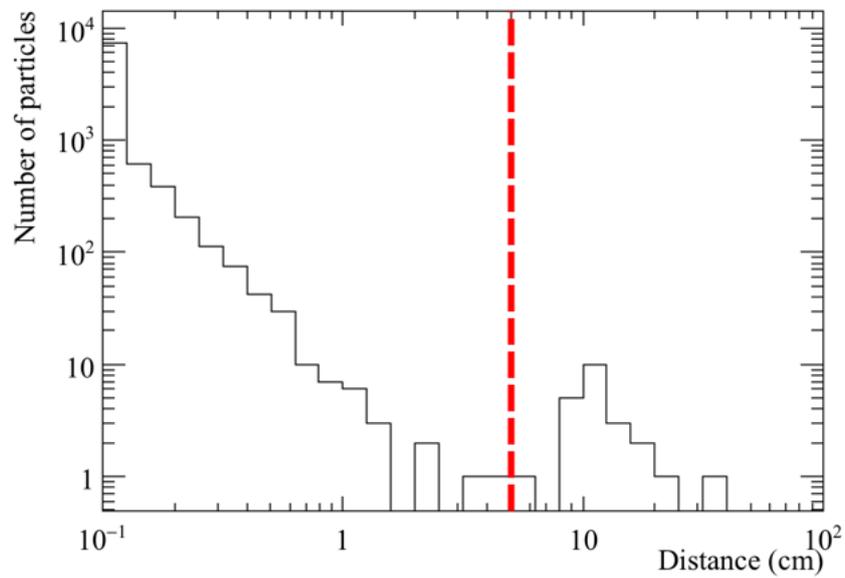
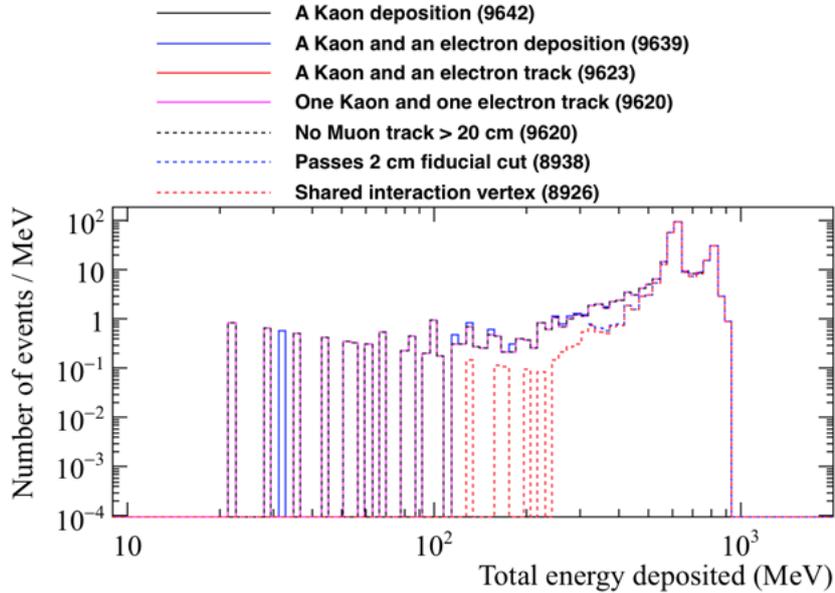
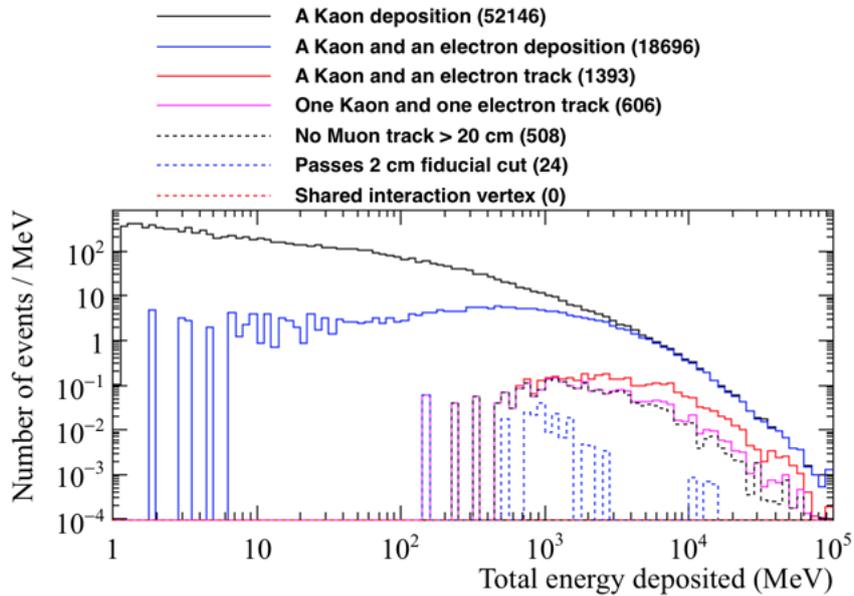


Fig. 6.16 The separation of the kaon and the electron produced in the simulated $n \rightarrow K^+ + e^-$ decays. The dashed red line, drawn at a separation of 5 cm, shows the maximum possible separation a kaon and an electron could have, and still be considered to have a common vertex. A separation of 5 cm is used as this is the separation of the TPCs in the FD. Events with separations of more than this, occur within the centre of an APA.



(a) The energy distribution of signal events.



(b) The energy distribution of cosmic background events.

Fig. 6.17 The energy distributions per MeV of deposited energy for the signal and cosmic background events, surviving the application of sequential cuts after the fiducial cut is modified. The total energy deposited in the detector is plotted on the x axis. These distributions are obtained by dividing the number of events within an energy bin by the bin width.

than the rest mass of a neutron, and so the event is not consistent with being from the decay of a neutron at rest. Section 6.3.4 will consider the energy deposition distributions of the simulated neutron decay events, though it is first necessary to calculate the expected values of these energy depositions, this is shown in Section 6.3.3.

6.3.3 Kinematics of nucleon decay

It is possible to apply strict energy criteria to the reconstructed energies of candidate nucleon decay events, as the distributions which one would expect are well defined. When determining the kinematics of the $n \rightarrow K^+ e^-$ decay channel, the neutron is assumed to decay at rest. The resulting kinematic equations are presented below:

$$E_n = m_n = E_K + E_e$$

$$E_e^2 = (m_n - E_K)^2$$

$$m_e^2 + p_e^2 = m_n^2 + (m_K^2 + p_K^2) - 2m_n E_K$$

using conservation of momentum $\vec{p}_K = -\vec{p}_e$

$$E_K = \frac{m_n^2 + m_K^2 - m_e^2}{2m_n} \quad (6.2)$$

$$E_e = \frac{m_n^2 + m_e^2 - m_K^2}{2m_n} \quad (6.3)$$

$$\begin{aligned} \sqrt{E_K^2 - m_K^2} = \vec{p}_K &= \sqrt{\left(\frac{m_n^2 + m_K^2 - m_e^2}{2m_n}\right)^2 - m_K^2 \left(\frac{4m_n^2}{4m_n^2}\right)} \\ \vec{p}_K &= \frac{\sqrt{(m_n^2 + m_K^2 - m_e^2)^2 - 4m_n^2 m_K^2}}{2m_n} \end{aligned} \quad (6.4)$$

$$\vec{p}_e = \frac{\sqrt{(m_n^2 + m_e^2 - m_K^2)^2 - 4m_n^2 m_e^2}}{2m_n} \quad (6.5)$$

Using $m_n = 939.565$ MeV, $m_K = 493.677$ MeV and $m_e = 0.511$ MeV [26], the total kaon and electron energies are equal to 599.479 MeV and 340.086 MeV respectively, whilst the momenta of the kaon and electron are both equal to 340.086 MeV·c⁻¹. A kaon (electron) with total energy $E_K = 599.479$ ($E_e = 340.086$) MeV will have a kinetic energy of 105.802 (339.575) MeV.

It is important to note that any nucleon which decays in LArTPCs will be contained in argon nuclei, this means that the energies of any particles which are produced will be

smearing by the Fermi motion of the decaying nucleon within the nucleus. Any kaon which is produced in the decay is also likely to scatter as it exits the nucleus, further smearing its energy and momentum [186]. This causes the true momenta and total energy of the particles produced in the decay to be different from the values which were calculated above. Thus, when searching for nucleon decay events, it is necessary to consider energy ranges of a few hundred MeV. This also approximates some of the uncertainties in measured energy associated with Fermi motion which has not been considered in this thesis.

6.3.4 Energy constraints on the cosmogenic background to the $n \rightarrow K^+ + e^-$ decay channel

As outlined in Section 6.3.1, it is possible to exclude background events from signal events using the distribution of energy depositions in the detector. As previously outlined, the energy depositions were split into several categories:

- The energy directly deposited by the kaon and its secondaries, excluding its decay products.
- The energy deposited by the kaon decay products and any of their secondaries.
- The energy directly deposited by the electron and its secondaries.
- The energy deposited near the shared kaon and electron vertex that is not associated with the kaon or electron.
- The energy deposited in the detector which does not fit any of the above criteria.

Following the earlier discussions in Section 6.3.1, it should be clear that the energy directly deposited by the kaon corresponds to the sum of all energy depositions, which are due to the kaon or its interaction products. Equivalently, the energy directly deposited by the electron, corresponds to the sum of all energy depositions which are due to the electron as it showers, and any particles created as a result of the shower. The energy deposited by the kaon decay products would include all depositions by the muon and subsequent electron, in the case that the kaon decayed via $K^+ \rightarrow \mu^+ + \nu_\mu$, and then the muon decayed via $\mu^+ \rightarrow e^+ + \nu_e + \bar{\nu}_\mu$. The energy deposited near the shared kaon and electron vertex, would primarily consist of energy depositions due to spallation products. These depositions would largely be due to protons, though may also include some depositions due to neutrons too, if they deposited energy near the interaction vertex. Any depositions within 5 cm of either the start of the kaon

track, or the start of the electron shower, are considered “near” to the interaction vertex. It is necessary to consider the energy depositions which are close to the start of both particles, because occasionally the particles are separated by the APA gaps, as shown in Figure 6.6. As can be seen from Figure 6.16, the separation between the start of the kaon track and the start of the electron shower, is normally very small. This means that the sum of the “near” energy depositions, can be considered to be the energy depositions “near” the common vertex between the kaon track and the electron shower. The final criteria, of any depositions which do not fit the above description, would largely consist of energy depositions due to the spallation neutrons in the decay sample. However, in the cosmic background sample, this would include all depositions by muons, pions, and any other particles in the detector which are not associated with either the kaon or the electron. In later discussions these depositions are generally referred to as “other” energy depositions.

In presenting the separation of simulated cosmic background events and simulated neutron decay events, the important distributions to consider are as follows:

- The energy directly deposited by the kaon, against the energy directly deposited by the electron. This is shown in Figure 6.18.
- The energy directly deposited by the kaon, plus the energy directly deposited by the electron, against the energy deposited near the shared kaon and electron vertex. This is shown in Figure 6.19.
- The energy deposited by the kaon, including decay products, against the energy deposited in the detector which does not fit any of the other criteria. This is shown in Figure 6.20.
- The energy deposited by the kaon, including decay products, plus the energy directly deposited by the electron, plus the energy deposited near the shared kaon and electron vertex, against the energy deposited in the detector which does not fit any of the other criteria. This is shown in Figure 6.21.

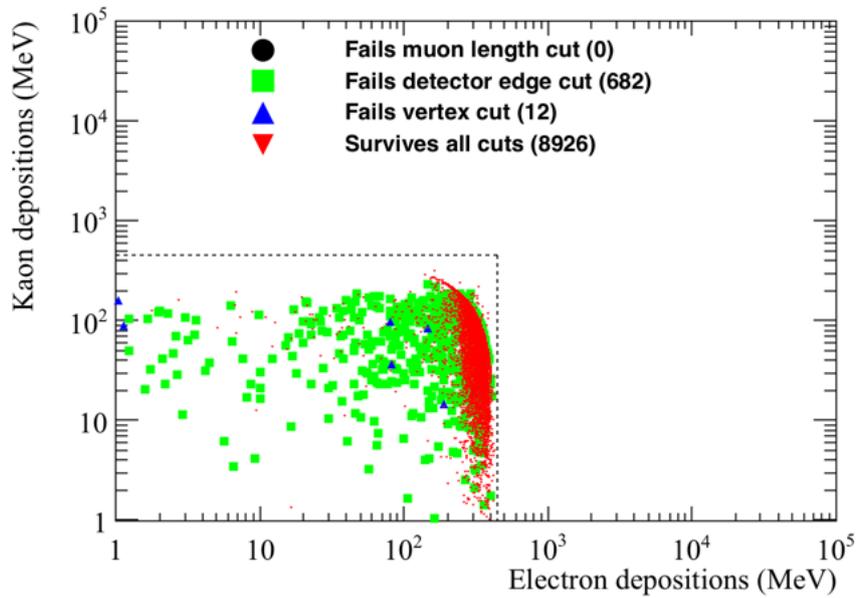
Each of these figures will be discussed in turn below. The events in the nucleon decay sample which pass all cuts have been plotted with smaller markers, this is to allow the less numerous events which fail the cuts to be visible.

Additionally, the figures outlined above have a box in the lower left corner of the plots produced by dashed grey lines. This constitutes the expected energy region where a signal event would lie on the graph, and has been drawn so as to contain as many of the simulated

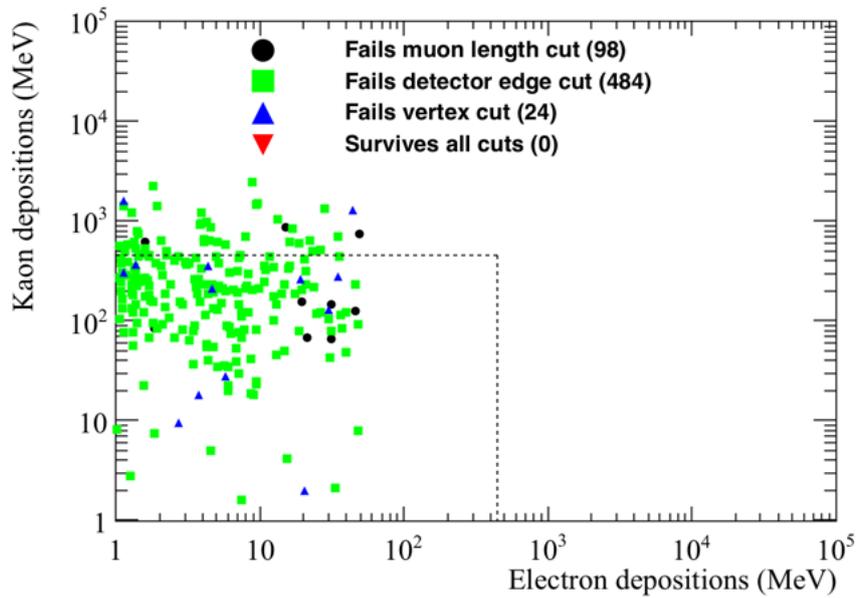
signal events as possible. These regions have been calculated using the values from Section 6.3.3 with an approximation for the effect Fermi motion added to them. The expected energy regions are propagated down to 0 MeV on both the x and y axis of all figures, and so it is possible for points not drawn here to pass the applied cuts. This is because it is possible for signal events to have no energy identified as “near” the common interaction vertex, such as the event shown in Figure 6.6. It is also possible for there to be no energy deposition which would be classed as “other” energy depositions.

From Figure 6.18, it can be seen that the electron energy distribution in the cosmic background sample, is very different from the one seen in the simulated neutron decay sample. The energies deposited by electrons in the nucleon decay sample are tightly concentrated between ~ 200 and ~ 400 MeV, whilst in the cosmic background sample, the energy deposited by the electron is almost always less than ~ 50 MeV. Many of the electrons in the cosmic background sample deposit less than 1 MeV of energy in the detector, and so are not shown in Figure 6.18b. This is why Figure 6.18b is very sparse despite 606 events being plotted. Realistically, these electrons are unlikely to be reconstructed due to their extremely low energies. From Figure 6.18a, it can be seen that some of the electrons produced in the nucleon decay events also deposit very little energy in the detector, though these events generally fail either the fiducial cut, or the vertex cut. An explanation as to why these events fail the cuts was presented in Section 6.3.2, when considering Figure 6.14.

From Figure 6.19a, it can be seen that as the energy deposited near the kaon and electron vertex increases, the sum of the energy deposited by the kaon and electron decreases. This is reasonable, because, when the spallation protons and neutrons have more energy, the kaon and electron will have less energy. The decrease in the energy deposited by the kaon and the electron, is roughly consistent with the increase in the amount of energy which is deposited near their shared vertex. This means that the sum of the three energies is generally around 450 MeV, as expected from the kinematic calculations shown in Section 6.3.3 where a two body decay was assumed. Though this is no longer valid when 3+ particles are produced, the majority of the rest mass energy of the neutron will be still be measured if the neutron energies are small. This is largely inconsistent with the simulated cosmic background events, where many events have no energy deposited near the shared vertex of the kaon and electron. However, the lack of energy deposited near the shared kaon and electron vertex cannot be used to discriminate against cosmogenic background events, as this is also observed in many simulated nucleon decay events. What can be used to separate cosmic background events from nucleon decay events though, is the sum of the three energy criteria, as it can be seen

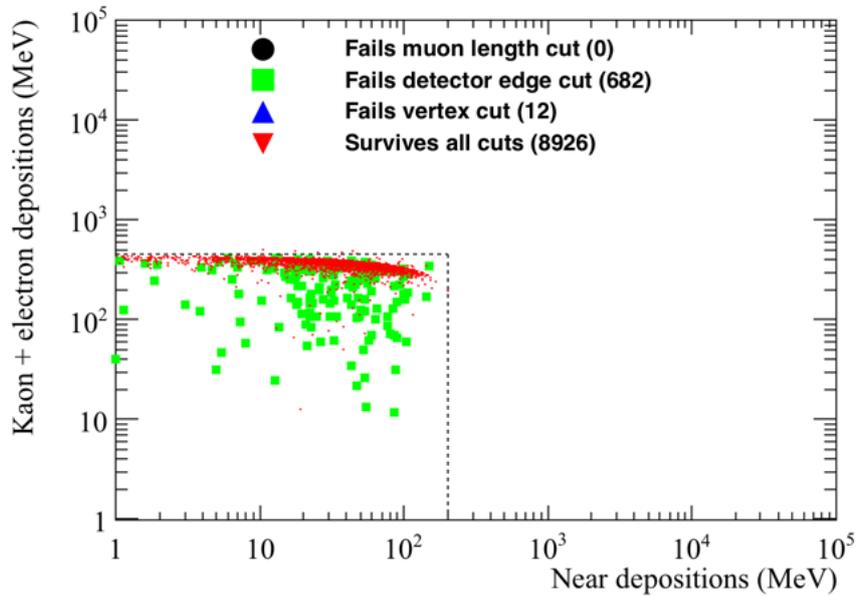


(a) Signal events.

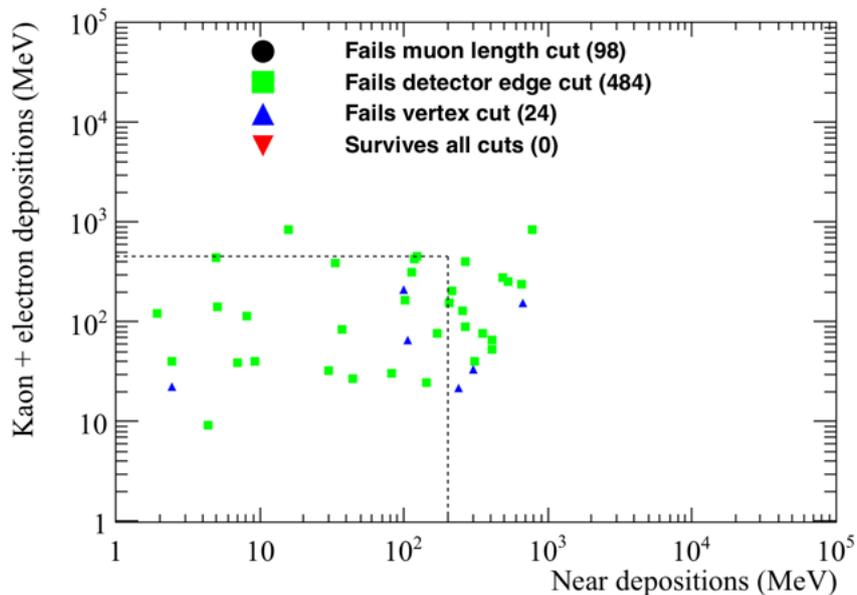


(b) Cosmic background events.

Fig. 6.18 The energy directly deposited by kaons versus the energy directly deposited by electrons, in the simulated nucleon decay (top) and cosmic background samples (bottom). The events failing the application of the muon length (black circle), fiducial (green box) and vertex (blue triangle) cuts, as well as the events passing all cuts (red triangle) are shown. The cuts are those defined at the start of Section 6.3.1.

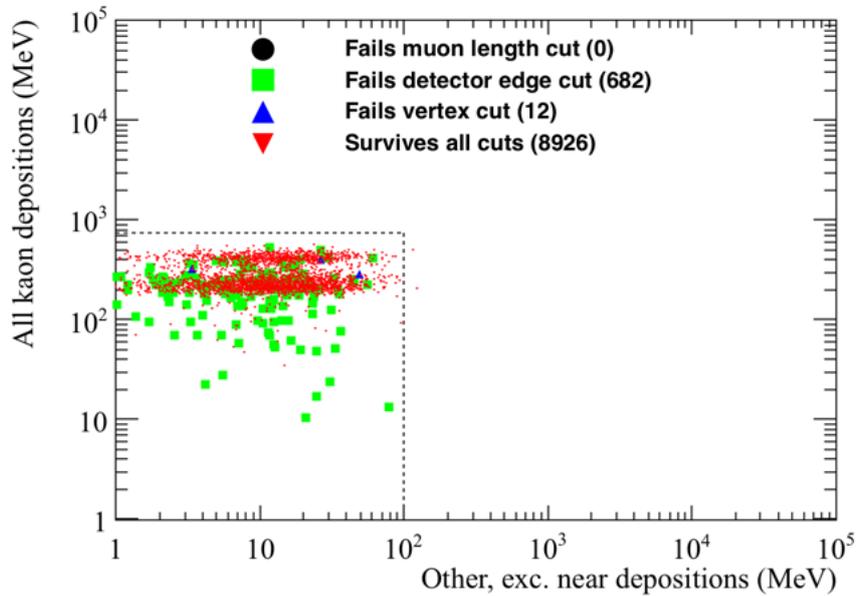


(a) Signal events.

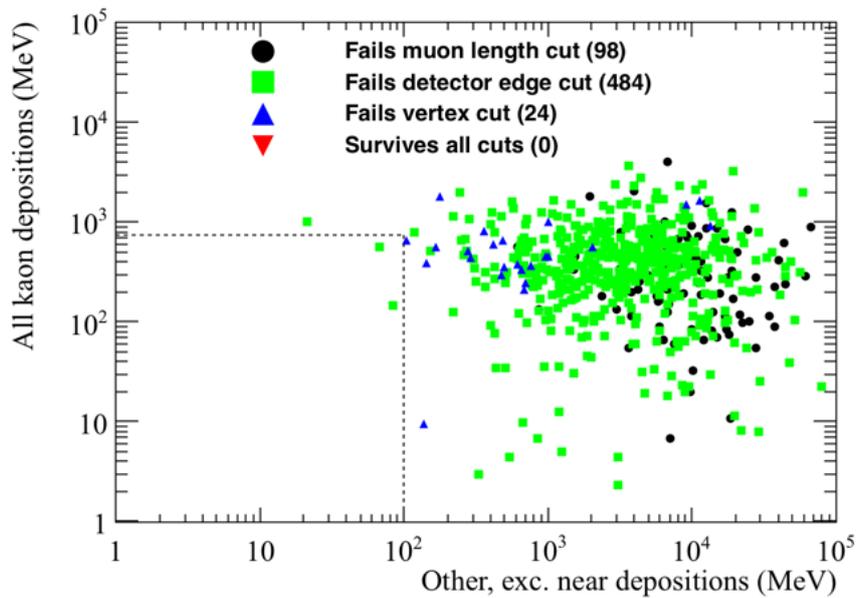


(b) Cosmic background events.

Fig. 6.19 The energy directly deposited by kaons, plus the energy directly deposited by electrons versus the energy deposited near the kaon and electron vertex, in the simulated nucleon decay (top) and cosmic background samples (bottom). The events failing the application of the muon length (black circle), fiducial (green box) and vertex (blue triangle) cuts, as well as the events passing all cuts (red triangle) are shown. The cuts are those defined at the start of Section 6.3.1.

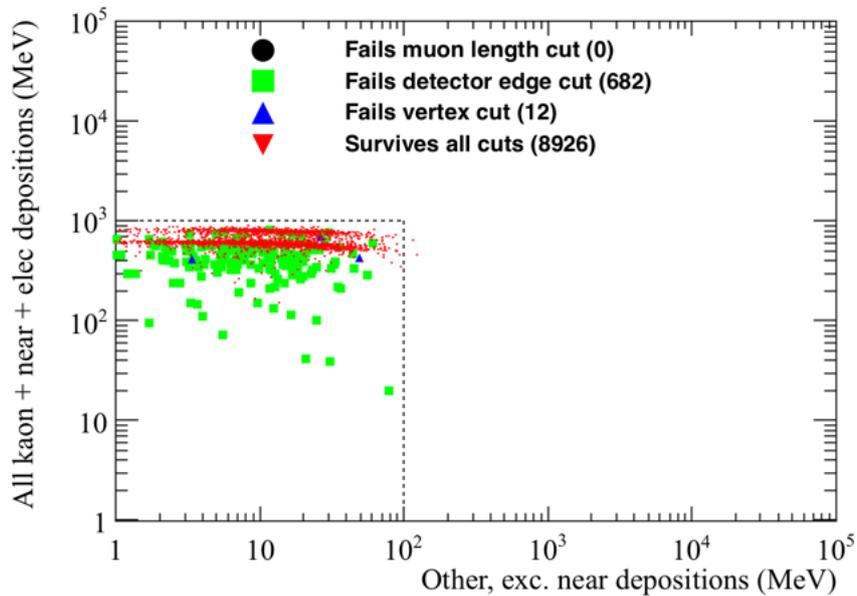


(a) Signal events.

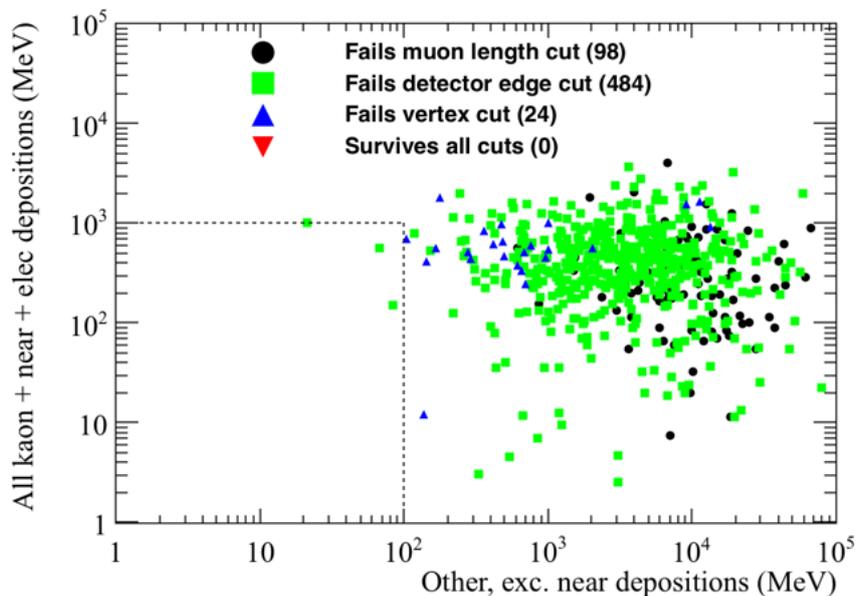


(b) Cosmic background events.

Fig. 6.20 The energy directly deposited by kaons, plus the energy deposited by the kaon decay products versus the energy depositions which do not fit any of the other criteria, in the simulated nucleon decay (top), and cosmic background samples (bottom). The events failing the application of the muon length (black circle), fiducial (green box) and vertex (blue triangle) cuts, as well as the events passing all cuts (red triangle) are shown. The cuts are those defined at the start of Section 6.3.1.



(a) Signal events.



(b) Cosmic background events.

Fig. 6.21 The energy directly deposited by kaons, plus the energy deposited by the kaon decay products, plus the energy directly deposited by electrons, plus the energy deposited near the kaon and electron vertex versus the energy depositions which do not fit any of the other criteria, in the simulated nucleon decay (top) and cosmic background samples (bottom). The events failing the application of the muon length (black circle), fiducial (green box) and vertex (blue triangle) cuts, as well as the events passing all cuts (red triangle) are shown. The cuts are those defined at the start of Section 6.3.1.

that this would rarely be around 450 MeV in the cosmic background sample.

Figure 6.20 shows that the sum of all energy deposits attributed to the kaon, against the sum of all energy depositions which are considered to be “other” energy depositions, is very different in the two samples. The total energy deposited in the detector which is attributed to the kaon is found by combining the energy directly deposited by the kaon, with the energy deposited by the kaon decay products. This sum is seen to have a banded structure in Figure 6.20a, this is due to the amount of energy from the kaon decay which is not reconstructed, and was discussed when considering Figure 6.13. The difference seen in the cosmogenic background and signal events, is primarily due to the large amount of “other” energy depositions which are seen in the cosmogenic background sample. This is expected, as in muon-induced events many particles may enter the detector, causing energy depositions which are not connected to the kaon track, the electron shower, or their common vertex, should one exist. This gives the most powerful mechanism for the discrimination between signal and background events, as in many background events the “other” energy deposited in the detector will be very large, and will cause the total energy deposited in the detector to be more than the rest mass of a neutron. Also, the “other” energy deposited by background events in the detector, will likely be concentrated in track or shower like structures. This is in contrast to signal events, where the “other” energy depositions are likely to be isolated depositions due to the spallation neutrons interacting in the detector. Classifying the structure of the “other” energy depositions is not performed here, though it could be included should any cosmic events appear to mimic a signal event. However, given the clear separation in the amount of “other” energy deposited in the detector, it is not currently required.

Figure 6.21 shows the total energy attributed to both the kaon and to the electron, plus any energy deposited near their shared interaction vertex, versus any “other” energy depositions. Here, the total energy attributed to the kaon is plotted, including the energy deposited by the kaon decay products, as was the case in Figure 6.20. As was seen in Figure 6.20, there is a clear separation between the simulated signal events, and the cosmic background events. This is again due to the large amount of “other” energy depositions observed in the background sample. It is interesting to compare Figures 6.20b and 6.21b, as they are very similar. This is to be expected from Figure 6.18b, where it was seen that the energy deposited by the electron was rarely more than 10 MeV, and was never more than 100 MeV. This was in stark contrast to Figure 6.18a, where the energy deposited in almost all of the electron showers was more than 100 MeV. Hence, whilst there is a large difference in the energy plotted on the y axis of Figures 6.20a and 6.21a, due to the addition of significant energy depositions from

the electron showers, there is relatively little difference the energy plotted on the y axis of Figures 6.20b and 6.21b.

A significant difference between Figures 6.20a and 6.21a, is the relatively narrow band of energies plotted on the y axis of Figure 6.21a, when compared to Figure 6.20a. This is because of the relationship between the energy deposited by the kaon, the energy deposited by the electron shower, and the energy deposited near their shared interaction vertex. This was seen in Figure 6.19a, where the sum of the three energies was roughly 450 MeV. This means that for events where the kaon deposited relatively little energy, shown by the population of events at low total kaon energies in Figure 6.20a, the energy deposited by the electron shower, plus the energy deposited near the shared interaction vertex, will be large.

Following the application of all energy cuts, as outlined by the grey dashed boxes in Figures 6.18, 6.19, 6.20, and 6.21, only 24 of the 8,926 simulated signal events, which passed all subsequent cuts, are removed. If only the energy cuts, and the requirement that there be at least one kaon track, and at least one electron shower in the event, are applied, it is found that only 2 background candidates would survive the application of cuts. None of these events present serious contenders for signal events though, as they both fail the vertex and fiducial cuts, and have electrons which have very little energy (< 10 MeV).

Many of the signal events which are removed by the energy cuts have large amounts of “other” energy depositions, though they would no longer be removed if the limit on “other” depositions were raised to 200 MeV, as seen in Figure 6.21a. Though increasing the cut on “other” energy depositions to 200 MeV would still result in all background events being removed, it would mean that the limit was much closer to the regions of “other” energy depositions which are populated in Figures 6.20b, and 6.21b. As a result if a candidate event were to have between 100 and 200 MeV of energy depositions considered to be “other” energy depositions, its energy distributions would have to be carefully checked in order to verify that it is definitively not due to a background event. It is envisioned that the process by which this would be done, would be to consider how the “other” energy deposits are distributed in the detector. As discussed earlier, in a signal event it is likely that they would be isolated depositions, whereas in a background event, there would likely be “track-like” objects disconnected from the kaon and electron vertex.

Therefore, the result of a preliminary study on background rejection for the $n \rightarrow K^+ + e^-$ decay channel, found no muon-induced background events which would survive the ap-

plication of cuts, in a sample of 2×10^9 muons representing 401.6 years of detector for a single FD module. This corresponds to a limit of the background rate of less than $0.44 \text{ events} \cdot \text{Mt}^{-1} \cdot \text{year}^{-1}$ at the 90% confidence level, using double sided errors [26] and a fiducial mass of 13.8 kt to give an exposure of 5.542 Mt·year.

Using the cuts outlined above, only 10.98% of signal events are removed, representing a high signal efficiency of 89.02%, where almost all of the signals which are removed are lost due to at least one of the particles produced in the decay escaping the detector. When only particles which are fully contained in the active volume of the detector are considered, the signal efficiency increases to 96.8%, where the 2 cm fiducial cut has removed most of the 3.2% of rejected signal events. These events have particles which travel close to the edge of the detector but do not quite leave it. For example, if the event shown in Figure 6.14 was a few cm further away from the edge of the detector, then this would be an example of an event which is fully contained in the detector but is rejected by the fiducial cut. If it is apparent that a candidate event is contained in the detector, even though it deposits significant amounts of energy near the detector edge, it may be possible to identify it as a neutron decay, due to the lack of signal mimicking background events which were observed in this study. Therefore, it may be possible to increase the signal efficiencies quoted above by relaxing the 2 cm fiducial cut, though considering events where particles leave the active volume of the detector will mean that PID is more difficult, and the reconstructed energy may be much less than the rest mass of the neutron.

6.3.5 Future improvements to nucleon decay studies and conclusions

Thus far, the study of the background to nucleon decay has only focused on the muon-induced background. However, as was mentioned in Section 1.2.2 atmospheric neutrinos are also able to produce signal mimicking events in underground detectors. There have been no studies to date concerning the atmospheric background to the $n \rightarrow K^+ + e^-$ decay channel, though this will need to be performed before a full background estimation can be made for this decay channel. However, this is a separate study which will have to be performed at a later date.

The study of muon-induced background to nucleon decay presented in this thesis has only been performed on Monte Carlo truth information, and so has not used reconstructed objects. The extension of the analysis to include work on tracks is an important next step, as then the full analysis which would be applied to real data can be tested. Preliminary studies have begun on hit reconstruction, and involve running a filter on the muons used in the earlier

analyses. This is because the number of events which are saved to disk would be prohibitive to running the full reconstruction process. As such, only events which meet the following criteria will be reconstructed [180];

- A minimum of 10 MeV deposited in the detector volume.
- A maximum of 3,000 MeV deposited in the detector volume.
- A maximum of 5 MeV deposited within 10 cm of the detector edge.

These criteria are designed to be soft enough that the full range of nucleon decay modes can be studied, including di-nucleon decay modes and neutron-antineutron oscillations. This is why the maximum deposited energy greatly exceeds the rest mass of a single nucleon. This also accounts for the fact that during reconstruction some energy depositions may not be reconstructed, so even though the total energy deposited in the detector is 3 GeV, the total reconstructed energy may be less.

Upon performing reconstruction, and hence energy and position smearing, it is likely that the energy cuts which were outlined in Section 6.3.4 will have to be broadened. This may result in more background events being contained within the expected energy distributions for signal events, however the effectiveness of the other cuts to remove background events should be largely unaffected. The reason for this is that upon performing reconstruction, the separation of the kaon and electron track will not decrease, and so the cut is expected to remain very efficient. The only cut which may be less effective is the requirement that there only be one kaon track, and only one electron shower in the event. This is because some of the low energy electron showers may not be reconstructed. However, even when this cut was relaxed no background events survived the application of all cuts.

The reduced reconstruction efficiency will also affect the nucleon decay signals. This, combined with the broader distributions of energies which will be reconstructed, will mean that the efficiency with which signal events can be identified will probably decrease. Studies have been performed to establish the event identification efficiencies when using reconstructed objects [181–184], though currently these have only been done with respect to the $p \rightarrow K^+ + \bar{\nu}_e$ decay channel. These studies have found that when using the current reconstruction algorithms particles may not be fully reconstructed, and also that there are errors in PID [182]. Therefore, the 100% efficient reconstruction which was assumed in this thesis could not be realised given the current state of reconstruction in DUNE. Using the current algorithms kaons are identified with an efficiency of 50.4% and a purity of

91.2% [182], though it is imagined that this will improve significantly before DUNE is constructed. The high purity of tracks which are identified as kaons hints that the problem of muons or pions being misidentified as kaons may be already be largely resolved; though a significant improvement is required in the efficiency of kaon track reconstruction.

However, it is envisioned that there will still be no muon-induced background events which survive the application of all cuts, though the impact of atmospheric neutrinos remains to be seen. This initial study, which only considered backgrounds from cosmogenic muons, shows that the observation of a single event in a 10 year run of the DUNE detector, representing an exposure of 400 kt·yr, would provide evidence of nucleon decay at a significance of over 3σ . Where a Poissonian distribution around the observation of 0 background events in the simulated background sample has been considered. The significance of this observation could be increased by simulating a larger sample of background events.

Chapter 7

Concluding remarks

This thesis has included work done in connection to the DUNE single phase LArTPC design, with a strong emphasis on the Phase II run of 35 ton detector. Simulations have also been performed in relation to the cosmic background to neutrino interactions when considering a surface detector, and the muon-induced background to nucleon decay searches in an underground detector.

Chapter 3 concerned the installation and operation of a system of cameras in the 35 ton detector. These cameras were found to be fully operational, and proved to be an extremely useful monitoring tool during the Phase II run. However, they were unable to conclusively observe any high voltage breakdowns in the LAr, though this capability had been shown in tabletop tests. Nevertheless, there are plans to install similar systems in future LArTPCs such as SBND [137].

Chapter 4 outlined the simulation efforts which were performed in preparation for the 35 ton Phase II run, culminating in a test of the method used to perform Particle IDentification (PID). The method of PID was seen to effectively identify muons and protons when they were simulated in the 35 ton detector in isolation. However, the method was seen to become less effective when the simulation was extended to include realistic samples of cosmic-ray events, where each “event” contained many particles. This was primarily because of the large number of reconstructed δ rays, which were associated with high energy muons passing through the detector. After the removal of these high energy muons, and the short tracks which were associated with them, a separation of proton and muon tracks was seen. This separation could be improved by using the true stopping point of the simulated particles from Monte Carlo truth. This showed that there was promise in performing the analysis on the data collected by the 35 ton data, though the selection of stopping particles would have to be

fully developed in order for this to be possible, as otherwise through going muons would saturate any sample of stopping particles.

A discussion of the features and issues surrounding the 35 ton dataset were presented in Chapter 5, concluding in the proposal of a new method of interaction time determination using the effects of longitudinal diffusion. Two metrics were proposed whereby this could be done, though considering the *RMS/Charge* of hits in reconstructed tracks was found to be the better identifier, where the *RMS* is the width of the reconstructed hit in ticks (500 ns), and the *Charge* is the integrated area of the reconstructed hit (ADC). It was found that the interaction time of through-going tracks could be determined with a systematic offset 171 μs , where the distribution has a FWHM of 210 μs , over a drift window of around 5,200 μs . When this is converted to a prediction of the x (drift) position of the particle, the systematic offset is found to be 18.5 cm, with a FWHM of 23.0 cm, over a drift length of 223 cm. This was then compared to a simulated low-noise detector, where the systematic offset of the interaction time (x position) determination was found to be 3 μs (0.4 cm) with a FWHM of 114 μs (12.6 cm). The effect of changing detector conditions in Monte Carlo was then observed with relation to; the electronics noise level of the detector, the applied electric field, the electron lifetime, and the magnitude of the longitudinal diffusion constant. As expected, lower noise levels, and larger longitudinal diffusion constants were seen to increase the accuracy of the method. Lower electron lifetimes were seen to cause the method to be more effective, as this corresponds to larger decreases in hit charge, giving the method a stronger handle on the x (drift) position of the particle. The method was seen to be largely unaffected by the electric field which was applied to the detector, this is reassuring as future detectors will run at 500 $\text{V}\cdot\text{cm}^{-1}$, and not at 250 $\text{V}\cdot\text{cm}^{-1}$ as the 35 ton detector did.

Finally, Chapter 6 concerns simulations of background events in LArTPCs. The backgrounds to neutrino interactions in a surface detector built on work presented in [159], where a system of cuts designed to reject background events had been developed. The author's contribution concerned the addition of an accurate detector geometry, and an accurate surface profile to simulations. The accurate detector geometry was seen to significantly affect the predicted background rate, causing the hadronic flux to be reduced by an order of magnitude due to the addition of extra shielding, whilst the muon-induced background was seen to increase due to the detector having a larger surface area. The background rate was seen to be largely unaffected by the addition of the accurate surface profile. The combined annual background rate from muons, protons and neutrons, after all cuts were applied, was seen to be reduced from ≈ 5 events when using a simplified detector geometry and surface profile, to

3.07 ± 0.25 events when using the accurate detector geometry and surface profile.

Chapter 6 concluded with a study concerning the muon-induced background rate to the $n \rightarrow K^+ + e^-$ decay channel. A series of cuts was developed to reject these events, concerning event features such as the proximity of hits to the detector walls, and the proximity of the kaon and electron tracks, as well as energy constraints assuming that the neutron decays at rest. After the cuts are applied to a sample of 2×10^9 muons, representing 401.6 years of detector live time for a single DUNE far detector module, no events are found which would mimic a nucleon decay event. This corresponds to a limit on the background rate of less than $0.44 \text{ events} \cdot \text{Mt}^{-1} \cdot \text{year}^{-1}$ at the 90% confidence level, using double sided errors [26] and a fiducial mass of 13.8 kt to give an exposure of 5.542 Mt·year. When the cuts are applied to a sample of 10,000 simulated signal events, a signal efficiency of 89.02% is found, this rises to 96.8% when only fully contained events are considered. This study suggests that DUNE is likely to be able to do a background free study of the $n \rightarrow K^+ + e^-$ decay channel. However, the analysis did not consider backgrounds due to atmospheric neutrinos, and did not use any reconstructed quantities, as it only used Monte Carlo truth energy depositions from simulated cosmic muons, and so needs to be extended in the future.

References

- [1] G. Aad *et al.* (ATLAS Collaboration), Phys. Lett. **B716**, 1 (2012).
- [2] S. Chatrchyan *et al.* (CMS Collaboration), Phys. Lett. **B716**, 30 (2012).
- [3] B. P. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), Phys. Rev. Lett. **116**, 061102 (2016).
- [4] Y. Fukuda *et al.* (Super-Kamiokande Collaboration), Phys. Rev. Lett. **81**, 1562 (1998).
- [5] Q. R. Ahmad *et al.* (SNO Collaboration), Phys. Rev. Lett. **89**, 011301 (2002).
- [6] W. Pauli, “*Open letter to the Gauverein meeting in Tübingen,*” (1930).
- [7] E. Fermi, Z. Phys. **88**, 161 (1934).
- [8] O. Kofoed-Hansen, Phys. Rev. **71**, 451 (1947).
- [9] G. C. Hanna and B. Pontecorvo, Phys. Rev. **75**, 983 (1949).
- [10] L. M. Langer and R. J. D. Moffat, Phys. Rev. **88**, 689 (1952).
- [11] C. L. Cowan *et al.*, Science **124**, 103 (1956).
- [12] G. Danby *et al.*, Phys. Rev. Lett. **9**, 36 (1962).
- [13] R. Davis *et al.*, Phys. Rev. Lett. **20**, 1205 (1968).
- [14] B. T. Cleveland *et al.*, Astrophys. J. **496**, 505 (1998).
- [15] B. Pontecorvo, Sov. Phys. JETP **6**, 429 (1957), [Zh. Eksp. Teor. Fiz.33,549(1957)].
- [16] K. S. Hirata *et al.*, Phys. Rev. Lett. **63**, 16 (1989).
- [17] M. L. Perl *et al.*, Phys. Rev. Lett. **35**, 1489 (1975).
- [18] K. Kodama *et al.*, Physics Letters B **504**, 218 (2001).
- [19] Q. R. Ahmad *et al.* (SNO Collaboration), Phys. Rev. Lett. **87**, 071301 (2001).
- [20] Z. Maki *et al.*, Progress of Theoretical Physics **28**, 870 (1962).
- [21] S. Andringa *et al.* (SNO+ Collaboration), Adv. High Energy Phys. **2016**, 6194250 (2016).

- [22] R. Arnold *et al.*, The European Physical Journal C **70**, 927 (2010).
- [23] D. Michael *et al.* (Minos Collaboration), Nucl. Instrum. Meth. A **596**, 190 (2008).
- [24] G. Tzanankos *et al.* (MINOS+ Collaboration), *MINOS+: a Proposal to FNAL to run MINOS with the medium energy NuMI beam*, Tech. Rep. FERMILAB-PROPOSAL-1016 (2011).
- [25] H. Nunokawa *et al.*, Prog. Part. Nucl. Phys. **60**, 338 (2008).
- [26] C. Patrignani *et al.* (Particle Data Group), Chin. Phys. **C40**, 100001 (2016).
- [27] N. Cabibbo, Phys. Rev. Lett. **10**, 531 (1963).
- [28] M. Kobayashi and T. Maskawa, Prog. Theor. Phys. **49**, 652 (1973).
- [29] L. Wolfenstein, Phys. Rev. D **17**, 2369 (1978).
- [30] S. P. Mikheev and A. Yu. Smirnov, Sov. J. Nucl. Phys. **42**, 913 (1985), [Yad. Fiz.42,1441(1985)].
- [31] G. Fogli *et al.*, Physics Letters B **583**, 149 (2004).
- [32] J. L. Hewett *et al.*, *Fundamental Physics at the Intensity Frontier*, Tech. Rep. ANL-HEP-TR-12-25, SLAC-R-991, FERMILAB-CONF-12-879-PPD (2012).
- [33] I. Esteban *et al.*, JHEP **01**, 087 (2017).
- [34] P. A. R. Ade *et al.* (Planck Collaboration), Astron. Astrophys. **594**, A13 (2016).
- [35] S. Mertens, Physics Procedia **61**, 267 (2015).
- [36] S. Schael *et al.* (SLD Electroweak Group, DELPHI, ALEPH, SLD, SLD Heavy Flavour Group, OPAL, LEP Electroweak Working Group, L3), Phys. Rept. **427**, 257 (2006).
- [37] C. Athanassopoulos *et al.* (LSND Collaboration), Phys. Rev. Lett. **81**, 1774 (1998).
- [38] A. Aguilar *et al.* (LSND Collaboration), Phys. Rev. D **64**, 112007 (2001).
- [39] A. A. Aguilar-Arevalo *et al.* (MiniBooNE Collaboration), Phys. Rev. Lett. **110**, 161801 (2013).
- [40] W. J. Marciano and G. Senjanović, Phys. Rev. D **25**, 3092 (1982).
- [41] H. Georgi and S. L. Glashow, Phys. Rev. Lett. **32**, 438 (1974).
- [42] H. Weyl, Zeitschrift für Physik **56**, 330 (1929).
- [43] G. Senjanovic, *15th International Symposium on Particles, Strings and Cosmology (PASCOS 2009) Hamburg, Germany, July 6-10, 2009*, AIP Conf. Proc. **1200**, 131 (2010).
- [44] R. Foot *et al.*, Zeitschrift für Physik C Particles and Fields **44**, 441 (1989).

- [45] I. Dorsner and P. Fileviez Perez, Nucl. Phys. **B723**, 53 (2005).
- [46] S. P. Martin, Adv. Ser. Direct. High Energy Phys **21**, 1 (2010).
- [47] D. J. Castano *et al.*, Phys. Rev. **D49**, 4882 (1994).
- [48] M. E. Gomez, in *Proceedings, European Network Meeting on Physics Beyond the Standard Model: Trieste, Italy, February 24-27, 1999* (1999) p. 006.
- [49] G. R. Farrar and P. Fayet, Physics Letters B **76**, 575 (1978).
- [50] J. C. Pati *et al.*, Physics Letters B **133**, 330 (1983).
- [51] R. N. Mohapatra and R. E. Marshak, Phys. Rev. Lett. **44**, 1316 (1980).
- [52] C. Berger *et al.* (Fréjus Collaboration), Physics Letters B **269**, 227 (1991).
- [53] F. Vissani, Phys. Rev. **D52**, 4245 (1995).
- [54] M. Severson, Phys. Rev. **D92**, 095026 (2015).
- [55] G. Battistoni *et al.*, Physics Letters B **133**, 454 (1983).
- [56] M. C. Goodman (Soudan-2 Collaboration), in *Proceedings of the 26th International Cosmic Ray Conference. August 17-25, 1999. Salt Lake City, Utah, USA* (1999) p. 364.
- [57] W. Gajewski *et al.*, *10th and Final Workshop on Grand Unification Chapel Hill, North Carolina, April 20-22, 1989*, Phys. Rev. **D42**, 2974 (1990).
- [58] K. Abe *et al.* (Super-Kamiokande Collaboration), Phys. Rev. D **95**, 012004 (2017).
- [59] K. Abe *et al.*, “*Letter of Intent: The Hyper-Kamiokande Experiment — Detector Design and Physics Potential —*,” (2011), arXiv:1109.3262.
- [60] A. Blake, *Studying Atmospheric Neutrino Backgrounds to $p \rightarrow K^+ \nu$ Decay*, Tech. Rep. LBNE-doc-8836 (2014).
- [61] S. Ritz *et al.* (HEPAP Subcommittee), “*Building for Discovery: Strategic Plan for U.S. Particle Physics in the Global Context*,” (2014), INSPIRE-1299183.
- [62] LBNE Collaboration, *Long-Baseline Neutrino Experiment (LBNE) Project, Volume 1: The LBNE Project*, Tech. Rep. LBNE-doc-5235 (2012).
- [63] LBNE Collaboration, *Long-Baseline Neutrino Experiment (LBNE) Project, Volume 2: The Beamline at the Near Site*, Tech. Rep. LBNE-doc-4317 (2012).
- [64] LBNE Collaboration, *Long-Baseline Neutrino Experiment (LBNE) Project, Volume 3: Detectors at the Near Site*, Tech. Rep. LBNE-doc-4724 (2012).
- [65] LBNE Collaboration, *Long-Baseline Neutrino Experiment (LBNE) Project, Volume 4: The Liquid Argon Detector at the Far Site*, Tech. Rep. LBNE-doc-4892 (2012).

- [66] LBNE Collaboration, *Long-Baseline Neutrino Experiment (LBNE) Project, Volume 5: Conventional Facilities at the Near Site*, Tech. Rep. LBNE-doc-4623 (2012).
- [67] LBNE Collaboration, *Long-Baseline Neutrino Experiment (LBNE) Project, Volume 6: Conventional Facilities at the Far Site*, Tech. Rep. LBNE-doc-5017 (2012).
- [68] A. Stahl *et al.*, *Expression of Interest for a very long baseline neutrino oscillation experiment (LBNO)*, Tech. Rep. CERN-SPSC-2012-021, SPSC-EOI-007 (2012).
- [69] R. Acciarri *et al.* (DUNE Collaboration), “*Long-Baseline Neutrino Facility (LBNF) and Deep Underground Neutrino Experiment (DUNE), Volume 4: The DUNE Detectors at LBNF*,” (2016), arXiv: 1601.02984.
- [70] J. Strait *et al.* (DUNE Collaboration), “*Long-Baseline Neutrino Facility (LBNF) and Deep Underground Neutrino Experiment (DUNE), Volume 3: Long-Baseline Neutrino Facility for DUNE*,” (2016), arXiv: 1601.05823.
- [71] S. D. Holmes *et al.*, in *Proc. 6th International Particle Accelerator Conference, Richmond, VA, USA* (2015) pp. 3982–3985.
- [72] R. Acciarri *et al.* (DUNE Collaboration), “*Long-Baseline Neutrino Facility (LBNF) and Deep Underground Neutrino Experiment (DUNE), Volume 2: The Physics Program for DUNE at LBNF*,” (2015), arXiv: 1512.06148.
- [73] R. Acciarri *et al.* (DUNE Collaboration), “*Long-Baseline Neutrino Facility (LBNF) and Deep Underground Neutrino Experiment (DUNE), Volume 1: The LBNF and DUNE Projects*,” (2016), arXiv: 1601.05471.
- [74] DUNE Collaboration, “*Proposals and Design Reports*,” (2017).
- [75] V. Chepel and H. Araujo, *JINST* **8**, R04001 (2013).
- [76] E. Worcester, *Presenting Long-Baseline Sensitivities*, Tech. Rep. DUNE-doc-2329 (2017).
- [77] M. Bass *et al.*, *Phys. Rev.* **D91**, 052015 (2015).
- [78] M. V. Diwan, *Heavy quarks and leptons. Proceedings, 7th International Conference, San Juan, Puerto Rico, June 1-5, 2004*, Frascati Phys. Ser. **35**, 89 (2004).
- [79] E. Worcester, *LBL MH Sensitivity: Standard Exposures*, Tech. Rep. DUNE-doc-2335 (2017).
- [80] E. Worcester, *LBL CPV Sensitivity: Exposure Plots*, Tech. Rep. DUNE-doc-2401 (2017).
- [81] E. Worcester, *LBL CPV Sensitivity: Standard Exposures*, Tech. Rep. DUNE-doc-2332 (2017).
- [82] E. Worcester, *LBL Sensitivity: Staging Plots*, Tech. Rep. DUNE-doc-2377 (2017).
- [83] M. C. Gonzalez-Garcia *et al.*, *JHEP* **11**, 052 (2014).

- [84] K. Abe *et al.* (Hyper-Kamiokande proto-), (2016).
- [85] J. Beringer *et al.* (Particle Data Group), Phys. Rev. D **86**, 010001 (2012).
- [86] H. Nishino *et al.* (Super-Kamiokande Collaboration), Phys. Rev. **D85**, 112001 (2012).
- [87] A. Bueno *et al.*, JHEP **04**, 041 (2007).
- [88] J. Klinger *et al.*, Phys. Lett. **B746**, 44 (2015).
- [89] M. Antonello *et al.*, Adv. High Energy Phys. **2013**, 260820 (2013).
- [90] L. Lin and M. Goodman, *Comprehensive list of nucleon decay modes*, Tech. Rep. DUNE-doc-679 (2016).
- [91] C. Adams *et al.* (LBNE Collaboration), “*The Long-Baseline Neutrino Experiment: Exploring Fundamental Symmetries of the Universe*,” (2013), arXiv: 1307.7335.
- [92] K. Scholberg, Annual Review of Nuclear and Particle Science **62**, 81 (2012).
- [93] R. Laha and J. F. Beacom, Phys. Rev. **D89**, 063007 (2014).
- [94] R. M. Bionta *et al.*, Phys. Rev. Lett. **58**, 1494 (1987).
- [95] K. Hirata *et al.*, Phys. Rev. Lett. **58**, 1490 (1987).
- [96] M. G. Aartsen *et al.* (IceCube Collaboration), Phys. Rev. Lett. **110**, 131302 (2013).
- [97] K. Choi *et al.* (Super-Kamiokande Collaboration), Phys. Rev. Lett. **114**, 141301 (2015).
- [98] J. F. Beacom, Ann. Rev. Nucl. Part. Sci. **60**, 439 (2010).
- [99] O. L. Caballero *et al.*, Astrophys. J. **745**, 170 (2012).
- [100] O. L. Caballero *et al.*, Phys. Rev. **D80**, 123004 (2009).
- [101] M. Thomson, *DUNE Collaboration, Strategy and Detector Requirements*, Tech. Rep. LBNE-doc-11100 (2015).
- [102] M. Adamowski *et al.*, JINST **9**, P07005 (2014).
- [103] D. Montanari *et al.* (LBNE Collaboration), in *Proceedings, International Cryogenic Engineering Conference and International Cryogenic Materials Conference 2014 (ICEC 25 – ICMC 2014): Enschede, The Netherlands, July 7-11, 2014*, Vol. 67 (2014) pp. 308–313.
- [104] A. Hahn *et al.* (LBNE Collaboration), in *Proceedings, 21st Symposium on Room-Temperature Semiconductor X-ray and Gamma-ray Detectors (RTSD 2014): Seattle, WA, USA, November 8-15, 2014* (2014) p. 7431158.
- [105] E. D. Church, “*LArSoft: A Software Package for Liquid Argon Time Projection Drift Chambers*,” (2013), arXiv: 1311.6774 .

- [106] The LArSoft collaboration, “*Software for LArTPCs,*” (2017).
- [107] R. Brun and F. Rademakers, *New computing techniques in physics research V. Proceedings, 5th International Workshop, AIHENP '96, Lausanne, Switzerland, September 2-6, 1996*, Nucl. Instrum. Meth. **A389**, 81 (1997).
- [108] S. Agostinelli *et al.*, Nucl. Instrum. Meth. A **506**, 250 (2003).
- [109] C. Andreopoulos *et al.*, “*The GENIE Neutrino Monte Carlo Generator: Physics and User Manual,*” (2015), arXiv: 1510.05494.
- [110] D. Casper, *Proceedings, 1st International Workshop on Neutrino-nucleus interactions in the few GeV region (NuInt 01): Tsukuba, Japan, December 13-16, 2001*, Nucl. Phys. Proc. Suppl. **112**, 161 (2002).
- [111] C. Hagman *et al.*, *Cosmic-ray Shower Library CRY*, Tech. Rep. UCRL-TM-229453 (2012).
- [112] C. Haggmann *et al.*, *Monte carlo simulation of proton-induced cosmic-ray cascades in the atmosphere*, Tech. Rep. UCRL-TM-229452 (2012).
- [113] D. Heck *et al.*, *CORSIKA: A Monte Carlo code to simulate extensive air showers*, Tech. Rep. FZKA-6019 (1998).
- [114] T. K. Gaisser, *Cosmic rays and particle physics* (Cambridge, UK: Univ. Pr. (1990) 279 p, 1990).
- [115] T. K. Warburton, “*Plan to add T0 calculations to calorimetry,*” 35 ton working group meeting (2015).
- [116] V. Kudryavtsev, Computer Physics Communications **180**, 339 (2009).
- [117] V. A. Kudryavtsev *et al.*, Nucl. Instrum. Meth. **A505**, 688 (2003).
- [118] P. Antonioli *et al.*, Astropart. Phys. **7**, 357 (1997).
- [119] V. A. Kudryavtsev *et al.*, Phys. Lett. **B471**, 251 (1999).
- [120] T. Alion, “*LBNE Geometries,*” (2014).
- [121] J. Asaadi *et al.*, *The Current Status of Simulation and Reconstruction for the DUNE Far Detector*, Tech. Rep. DUNE-doc-1689 (2016).
- [122] C. Harris and M. Stephens, in *In Proc. of Fourth Alvey Vision Conference* (1988) pp. 147–151.
- [123] P. V. C. Hough, in *Proceedings, 2nd International Conference on High-Energy Accelerators and Instrumentation, HEACC 1959: CERN, Geneva, Switzerland, September 14-19, 1959*, Vol. C590914 (1959) pp. 554–558.
- [124] J. S. Marshall and M. A. Thomson, Eur. Phys. J. **C75**, 439 (2015).
- [125] M. Wallbank, *EMShower Reconstruction*, Tech. Rep. DUNE-doc-1369 (2016).

- [126] R. Sulej and D. Stefan, “*Reconstruction: Projection Matching*,” Reconstruction working group meeting (2016).
- [127] J. B. Birks, Proc. Phys. Soc. **A64**, 874 (1951).
- [128] R. Acciarri *et al.* (ArgoNeuT Collaboration), JINST **8**, P08005 (2013).
- [129] J. Thomas and D. A. Imel, Phys. Rev. A **36**, 614 (1987).
- [130] A. Blatter *et al.*, Journal of Instrumentation **9**, P04006 (2014).
- [131] S. C. Delaquis *et al.*, JINST **8**, T12001 (2013).
- [132] K. Mavrokoridis *et al.*, JINST **9**, P02006 (2014).
- [133] E. Erdal *et al.*, JINST **10**, P11002 (2015).
- [134] N. McConkey *et al.*, JINST **12**, P03014 (2017).
- [135] G. Ghibaud and F. Balestra, Microelectronics Reliability **37**, 1353 (1997).
- [136] H. Zhao and X. Liu, Cryogenics **55**, 79 (2013).
- [137] M. Antonello *et al.* (LAr1-ND, ICARUS-WA104 and MicroBooNE Collaborations), “*A Proposal for a Three Detector Short-Baseline Neutrino Oscillation Program in the Fermilab Booster Neutrino Beam*,” (2015), arXiv:1503.01520 .
- [138] H. Bethe, Annalen der Physik **397**, 325 (1930).
- [139] F. Bloch, Annalen der Physik **408**, 285 (1933).
- [140] X. Qian, “*A Four-Plane proposal for Single-Phase LArTPC*,” Far Detector Sim/Reco working group meeting (2016).
- [141] J. Insler, “*Module to unstick ADC codes*,” 35 ton working group meeting (2016).
- [142] MicroBooNE Collaboration, *Noise Characterization and Filtering in the MicroBooNE TPC*, Tech. Rep. MICROBOONE-NOTE-1016-PUB (2016).
- [143] J. Cooley and J. Turkey, Math. Comp. **19**, 297 (1965).
- [144] N. Wiener, *Extrapolation, Interpolation, and Smoothing of Stationary Time Series* ([Cambridge]: Technology Press of the Massachusetts Institute of Technology., 1942).
- [145] M. Johnson *et al.*, *35T Observations and Measurements*, Tech. Rep. DUNE-doc-1704 (2016).
- [146] L. Bagby *et al.*, *35T Noise Studies Summary - APA wire readout noise*, Tech. Rep. DUNE-doc-1834 (2016).
- [147] T. Yang, “*Update on reconstruction of 35t data*,” 35 ton working group meeting (2016).
- [148] P. Torr and A. Zisserman, Computer Vision and Image Understanding **78**, 138 (2000).

- [149] M. Thiesse, “*Hit Reconstruction and Purity*,” 35 ton working group meeting (2016).
- [150] C. Anderson *et al.*, JINST **7**, P10019 (2012).
- [151] T. Yang, “*Update on 35 ton analysis*,” 35 ton working group meeting (2016).
- [152] S. Derenzo *et al.*, Nuclear Instruments and Methods **122**, 319 (1974).
- [153] S. E. Derenzo, “*LBL Physics Note No. 786: Electron diffusion and positive ion charge retention in liquid-filled high-resolution multi-strip ionization-mode chambers*.” (1974), unpublished.
- [154] E. Shibamura *et al.*, Phys. Rev. A **20**, 2547 (1979).
- [155] V. M. Atrazhev and I. V. Timoshkin, IEEE Trans. Dielectrics and Electrical Insulation **5**, 450 (1998).
- [156] Y. Li *et al.*, Nucl. Instrum. Meth. **A816**, 160 (2016).
- [157] D. Brailsford, “*Looking at electron diffusion in 35t data and MC*,” 35 ton working group meeting (2016).
- [158] C. Bromberg *et al.*, Journal of Instrumentation **10**, P07015 (2015).
- [159] M. Richardson, *Cosmic ray background estimations in large liquid argon detectors for proton decay via $p \rightarrow K^+ \bar{\nu}$ and ν_e appearance from ν_μ beam*, Ph.D. thesis, The University of Sheffield (2016).
- [160] N. S. Bowden, *Review Report: Mini-Review of Operation of the LBNE Far Detector at the Surface*, Tech. Rep. LBNE-doc-6493 (2012).
- [161] D. Wright, “*Shielding Physics List Description*,” (2012).
- [162] B. Baller, *LAr-FD Parameters Spreadsheet*, Tech. Rep. LBNE-doc-3383 (2015).
- [163] D. Cherdack, *The LBNE Fast MC*, Tech. Rep. LBNE-doc-7806 (2013).
- [164] D. Barker *et al.*, “*Muon-induced background for beam neutrinos at the surface*,” LBNE Reconfiguration Report (2012).
- [165] D. Mei *et al.*, *USD Cosmogenics Backgrounds Simulation Overview*, Tech. Rep. LBNE-doc-6621 (2012).
- [166] D. Coelho *et al.*, *Improving Electron / Gamma Separation*, Tech. Rep. LBNE-doc-8458 (2012).
- [167] A. Hitachi *et al.*, Phys. Rev. B **27**, 5279 (1983).
- [168] M. Aglietta *et al.* (LVD Collaboration), Phys. Rev. D **58**, 092005 (1998).
- [169] A. Dar, Phys. Rev. Lett. **51**, 227 (1983).
- [170] T. Blackwell *et al.*, *Cosmic-ray Spectra from CRY*, Tech. Rep. LBNE-doc-7517 (2013).

- [171] G. J. Feldman and R. D. Cousins, *Phys. Rev. D* **57**, 3873 (1998).
- [172] D. M. Mei *et al.*, *Astropart. Phys.* **34**, 33 (2010).
- [173] F. E. Gray *et al.*, *Nucl. Instrum. Meth.* **A638**, 63 (2011).
- [174] J. Heise, *Proceedings, Workshop on Dark Matter, Neutrino Physics and Astrophysics CETUP* 2013: 7th International Conference on Interconnection between Particle Physics and Cosmology (PPC 2013): Lead/Deadwood, South Dakota, USA, July, 8-13, 2013*, AIP Conf. Proc. **1604**, 331 (2014).
- [175] M. L. Cherry *et al.*, *Phys. Rev. D* **27**, 1444 (1983).
- [176] N. Abgrall *et al.* (MAJORANA Collaboration), “*Muon Flux Measurements at the Davis Campus of the Sanford Underground Research Facility with the Majorana Demonstrator Veto System,*” (2016), arXiv: 1602.07742.
- [177] E. Church, *September 2016 Progress Report from the Atmospheric Neutrinos, Nucleon Decay, and Cosmogenics Working Groups*, Tech. Rep. DUNE-doc-1644 (2016).
- [178] J. K. de Jong, *Atmospheric Muon Simulation for LBNE*, Tech. Rep. LBNE-doc-3144 (2012).
- [179] V. A. Kudryavtsev *et al.*, *Muon simulations for LBNE using MUSIC and MUSUN*, Tech. Rep. LBNE-doc-9673 (2014).
- [180] M. Robinson, “*Cosmogenic background to nucleon decay,*” Cosmogenics working group meeting (2017).
- [181] K. Wood, “*Reconstruction Efficiencies of Scintillation Light from PDK events,*” Cosmogenics working group meeting (2016).
- [182] A. Higuera, “*Proton Decay at DUNE $p \rightarrow K^+ + \bar{\nu}_e$,*” Cosmogenics working group meeting (2017).
- [183] T. Yang, “*Updates on reconstruction,*” Cosmogenics working group meeting (2017).
- [184] A. Higuera, “*Proton Decay at DUNE $p \rightarrow K^+ + \bar{\nu}_e$ and ^{39}Ar ,*” Cosmogenics working group meeting (2017).
- [185] J. Barrow *et al.*, *Nucleon Decay, Atmospheric Neutrino, and Cosmogenic Physics for FDTF Final Report*, Tech. Rep. DUNE-doc-2689 (2017).
- [186] D. Stefan and A. M. Ankowski, *Trends in nuclear physics. Proceedings, 43rd Conference on Nuclear Physics, Zakopane, Poland, September 1-7, 2008*, *Acta Phys. Polon.* **B40**, 671 (2009).

Appendix A

Supporting figures to Monte Carlo studies concerning determining interaction times using the effects of diffusion

The following figures support the figures shown in Section 5.5.3.

Figures A.1, A.2, A.3, and A.4 correspond to changing levels of the electronics noise.

Figures A.5, A.6, A.7, and A.8 correspond to changing electron lifetimes.

Figures A.9, A.10, A.11, and A.12 correspond to changing electric fields.

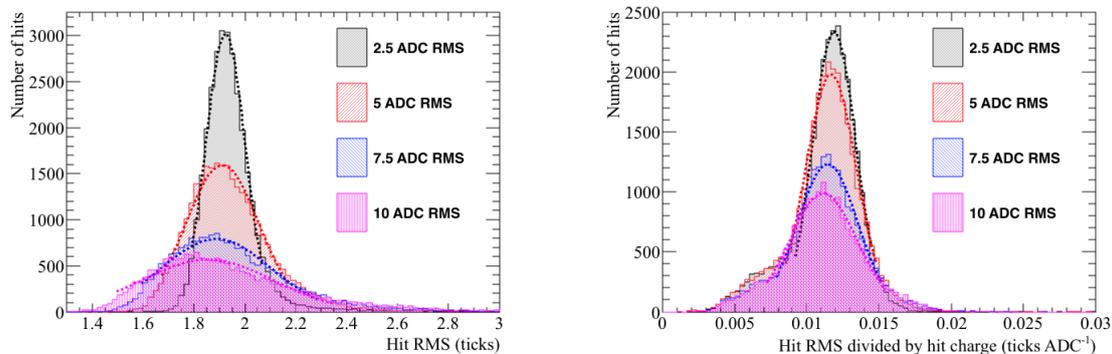
Figures A.13, A.14, A.15, and A.16 correspond to changing constants of longitudinal diffusion.

Figures A.1, A.5, A.9, and A.13 show how the distributions of the hit *RMS* and hit *RMS/Charge* change for hits between 20 and 30 cm from the APAs.

Figures A.2, A.6, A.10, and A.14 show how the most probable values of hit *RMS* changes with drift distance for tracks associated with counter differences of 4.

Figures A.3, A.7, A.11, and A.15 show how the most probable value of hit *RMS* next to the APAs changes with increasing counter difference.

Figures A.4, A.8, A.12, and A.16 show the normalised hit charge distributions for increasing noise levels, and the cut which is applied to remove the tails of the distribution.



(a) The distribution of hit RMS values for hits between $x = 20$ cm and $x = 30$ cm.

(b) The distribution of hit $RMS/Charge$ values for hits between $x = 20$ cm and $x = 30$ cm.

Fig. A.1 The distributions of the hit RMS and hit $RMS/Charge$ values for hits between $x = 20$ cm and $x = 30$ cm, for tracks with a counter difference of 4, for different values of the electronics noise.

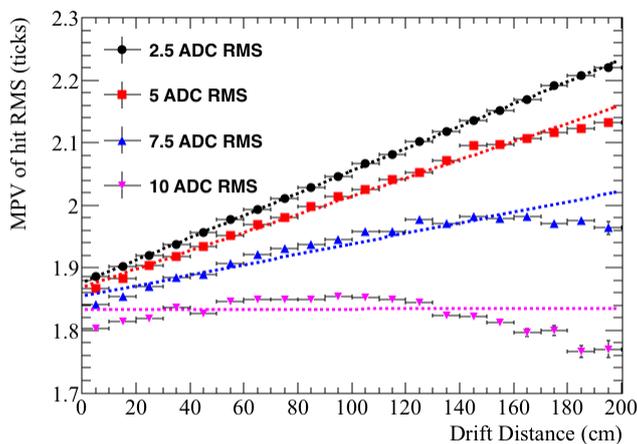


Fig. A.2 The most probable values of hit RMS as a function of drift distance, for tracks associated with a coincidence that had a counter difference of 4, for different values of the electronics noise.

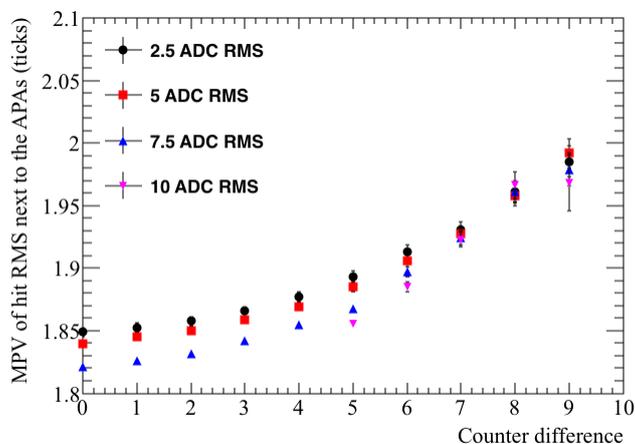


Fig. A.3 The most probable values of hit *RMS* within 10 cm of the APAs, as a function of the counter difference of the coincidence that the track was associated with, for different values of the electronics noise.

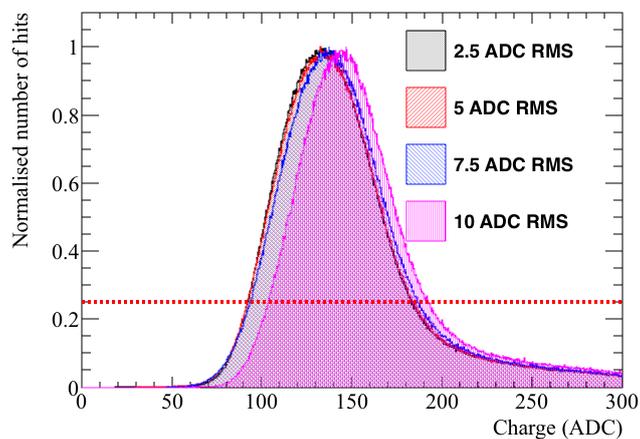
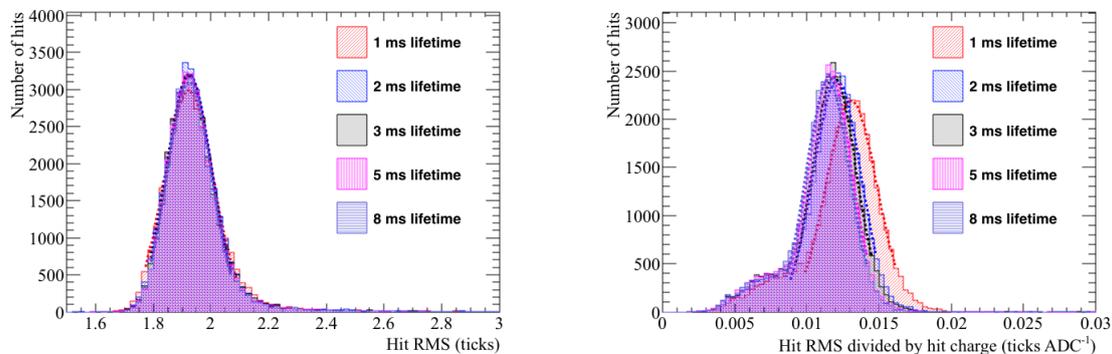


Fig. A.4 The normalised hit charge distribution for different values of the electronics noise. The hit charge is shown in units of ADC, and is normalised so that the most common hit charge has a value of 1. A cut on the normalised number of hits being greater than 0.25 is shown, the aim of this cut is to remove the tails of the hit charge distribution.



(a) The distribution of hit RMS values for hits between $x = 20$ cm and $x = 30$ cm.

(b) The distribution of hit $RMS/Charge$ values for hits between $x = 20$ cm and $x = 30$ cm.

Fig. A.5 The distributions of the hit RMS and hit $RMS/Charge$ values for hits between $x = 20$ cm and $x = 30$ cm, for tracks with a counter difference of 4, for different values of the electron lifetime.

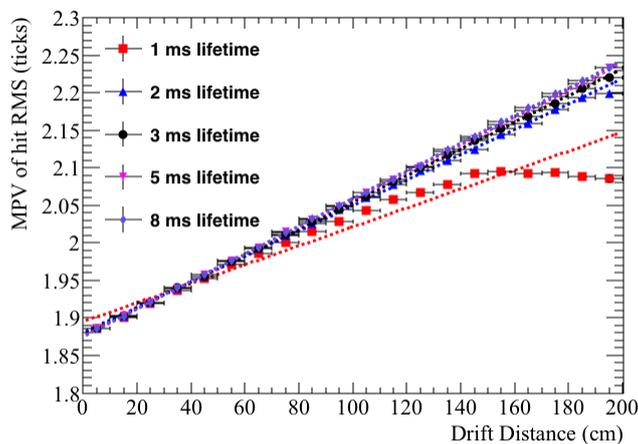


Fig. A.6 The most probable values of hit RMS as a function of drift distance, for tracks associated with a coincidence that had a counter difference of 4, for different values of the electron lifetime.

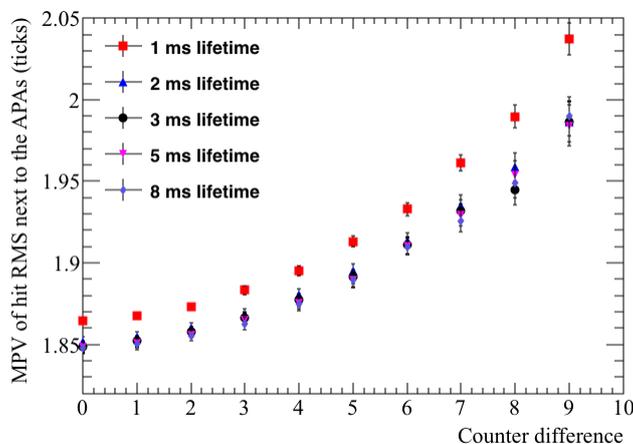


Fig. A.7 The most probable values of hit *RMS* within 10 cm of the APAs, as a function of the counter difference of the coincidence that the track was associated with, for different values of the electron lifetime.

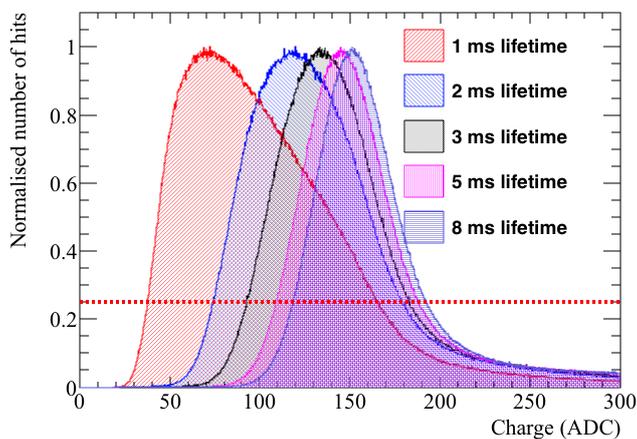
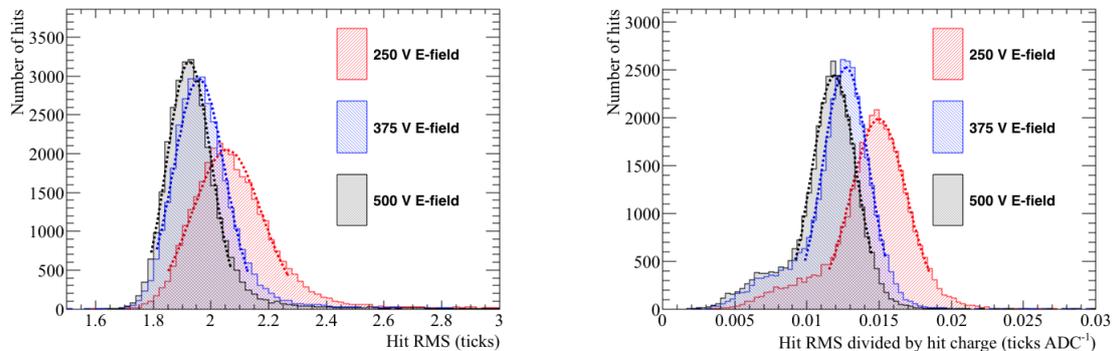


Fig. A.8 The normalised hit charge distribution for different values of the electron lifetime. The hit charge is shown in units of ADC, and is normalised so that the most common hit charge has a value of 1. A cut on the normalised number of hits being greater than 0.25 is shown, the aim of this cut is to remove the tails of the hit charge distribution.



(a) The most probable hit *RMS* values for hits between $x = 20$ cm and $x = 30$ cm.

(b) The most probably hit *RMS/Charge* values for hits between $x = 20$ cm and $x = 30$ cm.

Fig. A.9 The distributions of the hit *RMS* and hit *RMS/Charge* values for hits between $x = 20$ cm and $x = 30$ cm, for tracks with a counter difference of 4, for different values of the electric field.

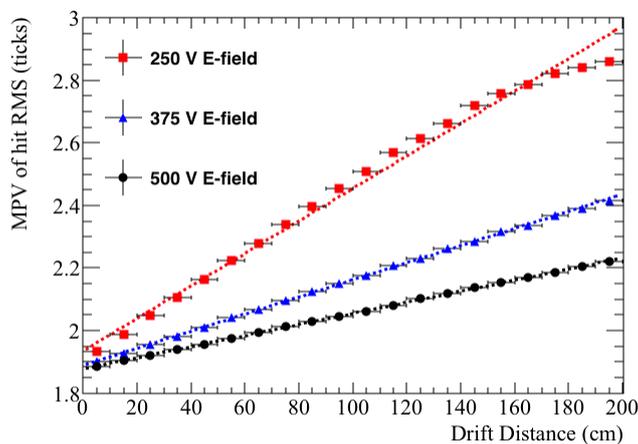


Fig. A.10 The most probable values of hit *RMS* as a function of drift distance, for tracks associated with a coincidence that had a counter difference of 4, for different values of the electric field.

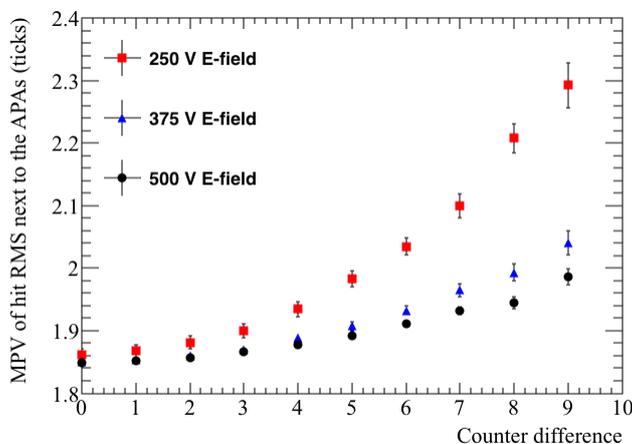


Fig. A.11 The most probable values of hit *RMS* within 10 cm of the APAs, as a function of the counter difference of the coincidence that the track was associated with, for different values of the electric field.

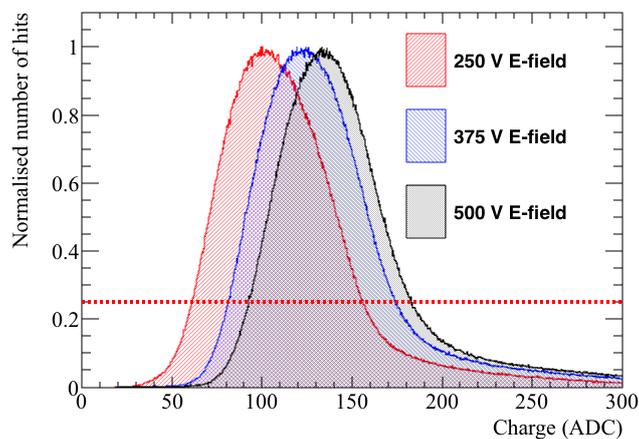
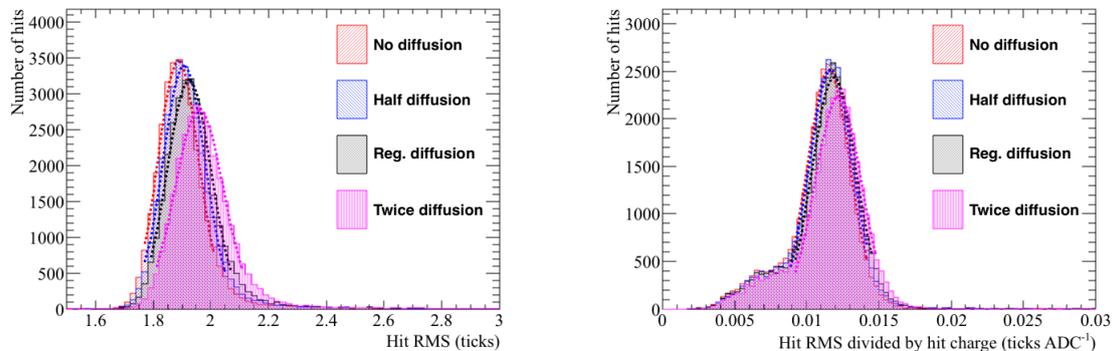


Fig. A.12 The normalised hit charge distribution for different values of the electric field. The hit charge is shown in units of ADC, and is normalised so that the most common hit charge has a value of 1. A cut on the normalised number of hits being greater than 0.25 is shown, the aim of this cut is to remove the tails of the hit charge distribution.



(a) The most probable hit RMS values for hits between $x = 20$ cm and $x = 30$ cm.

(b) The most probably hit $RMS/Charge$ values for hits between $x = 20$ cm and $x = 30$ cm.

Fig. A.13 The distributions of the hit RMS and hit $RMS/Charge$ values for hits between $x = 20$ cm and $x = 30$ cm, for tracks with a counter difference of 4, for different values of the constant of longitudinal diffusion.

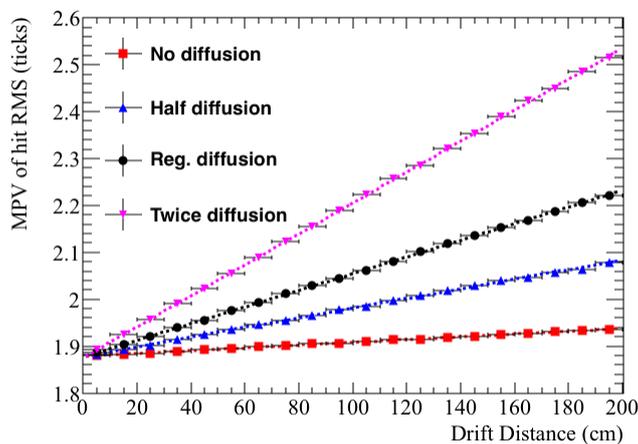


Fig. A.14 The most probable values of hit RMS as a function of drift distance, for tracks associated with a coincidence that had a counter difference of 4, for different values of the constant of longitudinal diffusion.

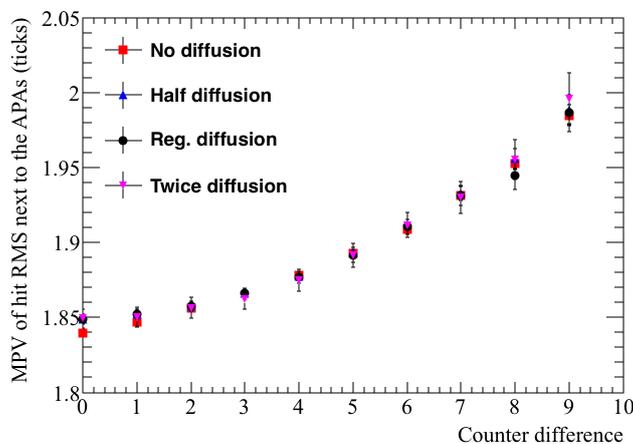


Fig. A.15 The most probable values of hit *RMS* within 10 cm of the APAs, as a function of the counter difference of the coincidence that the track was associated with, for different values of the constant of longitudinal diffusion.

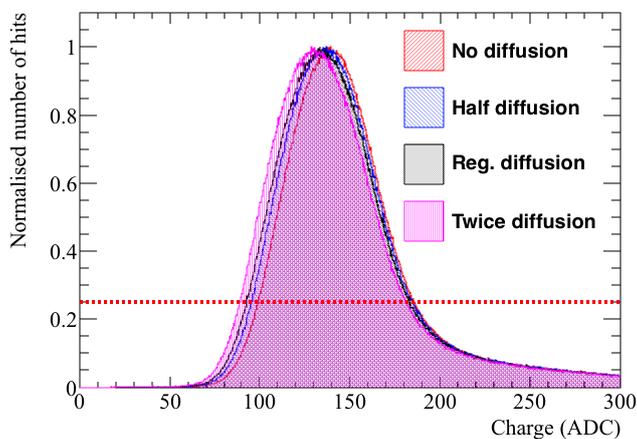


Fig. A.16 The normalised hit charge distribution for different values of the constant of longitudinal diffusion. The hit charge is shown in units of ADC, and is normalised so that the most common hit charge has a value of 1. A cut on the normalised number of hits being greater than 0.25 is shown, the aim of this cut is to remove the tails of the hit charge distribution.