

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



The
University
Of
Sheffield.

Thomas Karl Warburton

Department of Physics and Astronomy
University of Sheffield

This dissertation is submitted for the degree of
Doctor of Philosophy

March 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 highly level summaries of the theory necessary to support this thesis, and the various facets of the DUNE experiment. These summaries have been taken from a number of articles, and are referenced as such. When discussing the LArSoft software which DUNE uses, it is necessary to point out that this is a multi-experiment project, and so many people have contributed to its development in their own small ways, 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, and Matthew Thiesse, with only very minor input from the author. All work performed at Fermilab was done in conjunction with Michael Wallbank, and would have been impossible without the assistance of Dr. Linda Bagby.

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 and Dr. Tom Junk for their helpful insights in developing the methods shown here.

Chapter 5 details all aspects of the 35 ton data which was 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 and Dr. Dominic Brailsford 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 authors 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. All work was shown in his thesis, with according references, and so all tables have been taken from there, as these had used an improved analysis to the one which we had used in partnership. 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. 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

March 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 four 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, who 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 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 were installed, and cosmic-ray data was 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 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	10
List of tables	17
Nomenclature	19
1 Introduction and Theory	22
1.1 Neutrino physics	23
1.1.1 The history of neutrino oscillations	24
1.1.2 The theory of neutrino oscillations	25
1.1.3 Current experimental limits, and unanswered questions	32
1.2 Grand Unified Theories and nucleon decay	35
1.2.1 Overview of grand unified theories	35
1.2.2 Backgrounds to nucleon decay	37
2 The Deep Underground Neutrino Experiment	40
2.1 Brief experimental overview	40
2.2 The single phase detector design	41
2.3 The physics capabilities of DUNE	45
2.3.1 Neutrino physics	45
2.3.2 Nucleon decay	53
2.3.3 Additional physics opportunities	55
2.4 Path to building DUNE - The 35 ton prototype	57
2.5 The DUNE software	62
3 The 35 ton camera system	71
3.1 The selection and characterisation of cameras	72
3.2 The design of the camera system	73
3.3 Performance in the 35 ton	77

Table of contents	8
4 Simulations of the 35 ton prototype	81
4.1 Determination of interaction times	81
4.2 Calibrating calorimetric constants	86
4.3 Discerning reconstruction efficiencies	87
4.4 Performing particle identification	96
5 The 35 ton data sample	117
5.1 Organisation of the data structure	117
5.2 Reformatting the data to the offline structure	121
5.3 Observations on data quality and noise mitigation	122
5.4 Performance of reconstruction algorithms	127
5.5 Measuring interaction times using electron diffusion	136
5.5.1 Determining interaction times in 35 ton data	139
5.5.2 Determining interaction times in a low-noise detector using Monte Carlo, and differences with data	146
5.5.3 Impact of changing detector properties using Monte Carlo samples .	156
5.5.4 The limitations of and future improvements to the method of interaction time determination using diffusion	167
6 Simulations of the DUNE Far Detector	169
6.1 Simulations of the LBNE surface detector	169
6.1.1 Classifying signal and background events	172
6.1.2 Description of cuts used	173
6.1.3 Generating particles for background studies	175
6.1.4 Results from background simulations	177
6.1.5 Summary of simulations for the LBNE surface detector	181
6.2 The use of MUSUN in LArSoft	182
6.3 Nucleon decay channels in DUNE	185
6.3.1 Cosmogenic background to the $n \rightarrow K^+ + e^-$ decay channel	188
6.3.2 Signal events in the $n \rightarrow K^+ + e^-$ decay channel	195
6.3.3 Kinematics of nucleon decay	207
6.3.4 Energy constraints on the cosmogenic background to the $n \rightarrow K^+ + e^-$ decay channel	208
6.3.5 Future improvements to nucleon decay studies and conclusions	217
7 Concluding remarks	219

Table of contents	9
References	222
Appendix A Supporting figures to Monte Carlo studies concerning determining interaction times using the effects of diffusion	231

List of figures

1.1	A schematic representation of the two mass hierarchies, showing the fractional components of flavour states in each mass state	33
1.2	Schematic of a muon event mimicking a nucleon decay	38
2.1	A representation of the DUNE experimental setup	42
2.2	A cross section of the DUNE single phase detector design, showing 8 TPCs	44
2.3	An illustration of the wire wrapping in the DUNE single phase design	46
2.4	The ν_e appearance probability at 1300 km as a function of neutrino energy, for a range of values of δ_{CP}	48
2.5	The significance with which DUNE will be able to determine the neutrino mass hierarchy, for all values of δ_{CP}	49
2.6	The significance with which DUNE will be able to determine the value of δ_{CP} , for all values of δ_{CP}	50
2.7	The resolution with which DUNE will be able to determine the value of δ_{CP} , for increasing beam exposures	50
2.8	The significance to which the octant of θ_{23} can be determined, for different values of θ_{23}	51
2.9	The resolution with which DUNE will be able to determine the value of $\sin^2 \theta_{23}$ for increasing beam exposures	52
2.10	The resolution with which DUNE will be able to determine the value of $\sin^2 \theta_{13}$ for increasing beam exposures	52
2.11	The resolution with which DUNE will be able to determine the value of Δm_{31}^2 for increasing beam exposures	53
2.12	A comparison of current, and future, nucleon decay lifetime limits, compared with the ranges predicted by Grand Unified Theories.	55
2.13	A kaon event which was observed in the ICARUS T600 detector, in the CNGS data	56

2.14 An overview of the DUNE prototyping schedule, including complementary experiments	58
2.15 A schematic representation of the 35 ton prototype detector during the Phase II run	60
2.16 A representation of the counter locations in the 35 ton	61
2.17 The LArSoft coordinate system as it is represented in the 35 ton detector	64
2.18 A selection of reconstructed hits from simulated energy depositions	66
2.19 Performing disambiguation with different wire pitches.	68
3.1 The components which made up a camera module used in the 35 ton camera system	74
3.2 Two camera modules which are mounted in the 35 ton prototype detector	75
3.3 The system diagram for the 35 ton camera system	76
3.4 The calibration images for the 8 cameras used in the 35 ton camera system	78
3.5 The signal degradation over time for two cameras in the 35 ton camera system	79
4.1 Matching tracks and flashes in the 35 ton using positions in the yz plane	83
4.2 The central x position of a reconstructed track versus the number of detected photoelectrons	84
4.3 The predicted x positions of flashes using the relationship between photo-electron and drift distance	84
4.4 The number of events as a function of the difference between Monte Carlo and photon detector times	85
4.5 The calibration of the calorimetric constants in the 35 ton	88
4.6 The reconstruction efficiencies for simulated events as a function of the track length in the detector from Monte Carlo truth.	91
4.7 The reconstruction efficiencies for simulated events as a function of the deposited energy from Monte Carlo truth.	91
4.8 The reconstruction efficiencies for simulated events as a function of the θ track angle from Monte Carlo truth track.	92
4.9 The reconstruction efficiencies for simulated events as a function of the ϕ track angle from Monte Carlo truth track.	92
4.10 The reconstruction efficiencies for simulated events as a function of the θ and ϕ track angles from Monte Carlo truth track.	93
4.11 The mean energy loss per unit track length of different particle masses in different materials	97

4.12	Stopping power for different particle masses as a function of residual range in liquid argon	99
4.13	The distribution of PIDA values, calculated using Monte Carlo truth, for different particle masses	100
4.14	The reconstruction efficiencies for protons in a sample generated using CRY.	101
4.15	The reconstruction efficiencies for the simulated isolated muon and proton samples in the 35 ton detector.	103
4.16	The calculated PIDA values for the simulated isolated proton and muon samples in the 35 ton detector.	105
4.17	The $\frac{dE}{dx}$ versus residual range plot for the simulated isolated proton and muon samples in the 35 ton detector.	106
4.18	The calculated PIDA values for the simulated proton enriched sample in the 35 ton detector	110
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	112
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 . .	113
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	114
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	116
5.1	The 35 ton data sample	118
5.2	The 35 ton data structure	120
5.3	Dropped TPC data in the 35 ton	123
5.4	Recovering stuck ADC codes in the 35 ton	124
5.5	Removing coherent noise in the 35 ton	125
5.6	Applying Wiener filters to the 35 ton data	126
5.7	The effect of noise removal algorithms in the 35 ton data	128
5.8	dQ/dx in the 35 ton as a function of drift time	130
5.9	The numbering scheme for the east - west counters in the 35 ton	131
5.10	The dot product of the track and vector joining the centres of the coincidence counters in the yz plane	132
5.11	The alignment of reconstructed tracks with the vectors joining the centres of the coincidence counters	132

5.12 Reconstruction efficiencies of through going tracks in the 35 ton data	135
5.13 Schematic showing the process of diffusion	137
5.14 A simulated event display showing multiple tracks and flashes in the 35 ton detector	138
5.15 The effect of adding a noise baseline to a hit	140
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	141
5.17 The drift distance dependence of diffusion in the 35 ton dataset for coincidences with a counter difference of 4	142
5.18 The angular dependence of diffusion in the 35 ton dataset for hits within 10 cm of the APAs	143
5.19 The distribution of normalised hit charge in the 35 ton dataset	144
5.20 The difference between the predicted and reconstructed hit times in the 35 ton dataset	145
5.21 The accuracy of the hit <i>RMS</i> method in the 35 ton dataset	147
5.22 The accuracy of the hit <i>RMS/Charge</i> method in the 35 ton dataset	148
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	150
5.24 The drift distance dependence of diffusion in the 35 ton dataset and Monte Carlo for coincidences with a counter difference of 4	151
5.25 The angular dependence of diffusion in the 35 ton dataset and Monte Carlo for hits within 10 cm of the APAs	152
5.26 The distribution of normalised hit charge in the 35 ton dataset and a Monte Carlo sample	153
5.27 Comparing the accuracy of the hit <i>RMS</i> method in the 35 ton dataset and a Monte Carlo simulation	154
5.28 Comparing the accuracy of the hit <i>RMS</i> method in the 35 ton dataset and a Monte Carlo simulation	155
5.29 Comparing the accuracy of the hit <i>RMS</i> method, as the electronic noise changes	158
5.30 Comparing the accuracy of the hit <i>RMS</i> method, as the electronic noise level changes	159
5.31 Comparing the accuracy of the hit <i>RMS</i> method, as the electron lifetime changes	160
5.32 Comparing the accuracy of the hit <i>RMS</i> method, as the electron lifetime changes	161
5.33 Comparing the accuracy of the hit <i>RMS</i> method, as the electric field changes	162

5.34 Comparing the accuracy of the hit <i>RMS</i> method, as the electric field changes	163
5.35 Comparing the accuracy of the hit <i>RMS</i> method, as the constant of longitudinal diffusion changes	164
5.36 Comparing the accuracy of the hit <i>RMS</i> method, as the constant of longitudinal diffusion changes	165
6.1 The <i>complex</i> detector geometry used in the LBNE surface detector simulations	171
6.2 A diagram of the PoCA, and the angle w.r.t the beam calculations in the LBNE surface detector simulations	174
6.3 The correlation between the surface profile, and the distribution of azimuthal angles at the DUNE far detector site	184
6.4 The distributions of muon parameters for a sample of 10^6 muons generated by MUSUN in LArSoft	186
6.5 The initial positions of muons generated by MUSUN around a DUNE 10 kt module	187
6.6 A simulated $n \rightarrow K^+ + e^-$ decay which occurred in a gap between TPCs .	191
6.7 The energy distribution of background events surviving the application of sequential cuts.	193
6.8 The energy distribution of background events per MeV of deposited energy surviving the application of sequential cuts	194
6.9 A simulated $n \rightarrow K^+ + e^-$ decay which occurred near the edge of the detector volume	196
6.10 A simulated $n \rightarrow K^+ + e^-$ decay which is fully contained in a single TPC volume	197
6.11 The energy distribution of signal events surviving the application of sequential cuts in the $n \rightarrow K^+ + e^-$ channel	198
6.12 The energy distribution of signal events per MeV of deposited energy surviving the application of sequential cuts	198
6.13 The number of events as a function of the energy deposited by the kaon decay products.	200
6.14 A simulated $n \rightarrow K^+ + e^-$ decay where the kaon did not deposit any energy in the active volume	202
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.	203
6.16 The separation of the kaon and the electron produced in the simulated $n \rightarrow K^+ + e^-$ decays	205

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	206
6.18 The energy directly deposited by kaons versus the energy directly deposited by electrons, in the simulated nucleon decay and cosmic background samples	211
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	212
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	213
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	214
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	232
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	233
A.3 The angular dependence of hits within 10 cm of the APAs, for different values of the electronics noise	233
A.4 The normalised hit charge distribution for different values of the electronics noise	234
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	234
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	235
A.7 The angular dependence of hits within 10 cm of the APAs, for different values of the electron lifetime	235
A.8 The normalised hit charge distribution for different values of the electron lifetime	236

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	236
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	237
A.11 The angular dependence of hits within 10 cm of the APAs, for different values of the electric field	237
A.12 The normalised hit charge distribution for different values of the electric field	238
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	238
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	239
A.15 The angular dependence of hits within 10 cm of the APAs, for different values of the constant of longitudinal diffusion	239
A.16 The normalised hit charge distribution for different values of the constant of longitudinal diffusion	240

List of tables

1.1	Three-flavour oscillation parameters from a fit to global data after the NOW-2016 and ICHEP-2016 conferences	33
2.1	The parameters of the wire planes in the DUNE FD single phase design	43
2.2	The beam exposure in units in units of $\text{kt}\cdot\text{MW}\cdot\text{years}$, as a function of time, assuming a staged DUNE construction	47
2.3	Nucleon decay limits in DUNE and Super-Kamiokande, in some favoured decay channels	56
4.1	Stopping power parameterisation for various particle types in liquid argon	98
4.2	The properties of initial particles simulated in the muon and proton samples. The angles θ_{xz} , and θ_{yz} , are defined as the angle that a vector makes in the xz , and yz planes, respectively.	102
4.3	A summary of the PIDA values calculated for the simulated isolated proton sample, as sequential cuts are applied	107
4.4	A summary of the PIDA values calculated for the simulated isolated muon sample, as sequential cuts are applied	108
4.5	A summary of the PIDA values calculated for the primary particles in the simulated isolated proton sample, as sequential cuts are applied	108
5.1	The angles which tracks, with given counter differences, have relative to the APA frames	134
6.1	The minimum energies of simulated particles, when determining the cosmogenic background for the LBNE surface detector	176
6.2	The normalised annual background rate for the <i>simple</i> detector geometry, using the flat surface profile	177

6.3	The normalised annual background rate, for events where a primary muon enters the active volume of the detector, for the <i>complex</i> geometry and accurate surface profile	178
6.4	The normalised annual background rate, for events where a primary muon misses the active volume of the detector, for the <i>complex</i> geometry and accurate surface profile	179
6.5	The normalised annual background rate, for proton induced events, for the <i>complex</i> geometry	180
6.6	The normalised annual background rate, for neutron induced events, for the <i>complex</i> geometry	181
6.7	The normalised annual background rate for the <i>simple</i> detector geometry and flat surface profile, and for the <i>complex</i> geometry and accurate surface profile	182
6.8	Muon flux parameters as calculated with MUSIC/MUSUN.	183
6.9	The number of events which could mimic a $n \rightarrow K^+ + e^-$ decay, when cuts are applied in isolation	194
6.10	The most common decay modes of charged kaons, and their probabilities .	199

Nomenclature

Roman Symbols

tick Unit of time equal to 500 ns

Acronyms / Abbreviations

art analysis reconstruction framework

ADC Analogue to Digital Converter

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 Minimally 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

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 SUperSymmetR \mathbf{Y}

TPC Time Projection Chamber

¹ Chapter 1

² Introduction and Theory

³ The “Standard Model of Particle Physics” (SM) is a set of theories which has been widely
⁴ tested and has been found to accurately predict the interactions of fundamental particles.
⁵ These tests have come in many forms throughout the 20th and 21st centuries, and include
⁶ the detection of all of the quarks and leptons which it predicts, as well as measurements of
⁷ the properties of these particles. The recent discovery of the Higgs boson [1, 2] “completed”
⁸ the SM, as this was the last particle which it predicted to be observed. However, despite its
⁹ many successes, the SM does not represent the “final” theory of fundamental particle physics,
¹⁰ should one exist. This is because there are many questions made by recent experimental
¹¹ observations which the SM is unable to address, some of these will be briefly discussed below.

¹²

¹³ Firstly, though the SM accurately predicts the interactions made by the electromagnetic,
¹⁴ weak nuclear, and strong nuclear forces, it makes no mention of gravity. This is a major flaw
¹⁵ of the SM as gravity is one of the driving forces in the formation of astronomical objects
¹⁶ such as planets, stars, and galaxies. With the recent detection of gravitational waves [3]
¹⁷ this issue has again be brought into focus. Secondly, the rotational velocities of galaxies is
¹⁸ measured to be far greater than the predicted value, hinting at the presence of a significant
¹⁹ amount of matter which we are unable to detect. The SM makes no prediction as to what this
²⁰ “dark-matter” is comprised of. Thirdly, measurements of distant supernovae appear to show
²¹ that the expansion of the universe is accelerating, not decelerating as would be expected,
²² this implies the presence of some form of unknown energy source. Again, the SM makes no
²³ prediction as to what this unknown energy source, or “dark-energy” is. A further point of
²⁴ consternation with the SM is that it is not “elegant,” as it has as many as 19 free parameters,
²⁵ which appear to be unrelated to each other. There are also unresolved questions regarding
²⁶ the particles which are predicted by the SM, such as, why charge is quantised, why there
²⁷ 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 should 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 conclude 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

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

11

12 1.1.1 The history of neutrino oscillations

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

27

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

tector to look for the electron neutrinos produced by the sun, and measured a flux which 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].

Conclusive proof for neutrino oscillations came in 2001, when the SNO experiment measured both the Neutral Current (NC) and Charged Current (CC) interactions of solar neutrinos. The SNO experiment 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, which is sensitive to all flavours of neutrinos, 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. 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 are 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

¹ the form shown in Equation 1.1.

$$\begin{pmatrix} v_e \\ v_\mu \\ v_\tau \end{pmatrix} = U_{PMNS} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \quad (1.1)$$

³ The matrix labelled U_{PMNS} in Equation 1.1, is then expressed by Equations 1.2, 1.3 and 1.4.
⁴ In these equations it has been assumed that U_{PMNS} is unitary, and that there are three angles
⁵ (θ_{12} , θ_{13} and θ_{23}) plus a CP-violating phase δ , which explain the mixing between the mass
⁶ and flavour eigenstates. The notation $s_{\alpha\beta}$ and $c_{\alpha\beta}$ has been used to denote $\sin \theta_{\alpha\beta}$ and
⁷ $\cos \theta_{\alpha\beta}$ respectively.

⁸

$$\begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{pmatrix} \quad (1.2)$$

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

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

¹³ Equation 1.2 shows how each element in the U_{PMNS} relates the flavour states to the mass
¹⁴ states, whilst Equation 1.3 shows the full mixing formalism for Dirac neutrinos. Finally,
¹⁵ Equation 1.4 separates the full formalism into three 3×3 matrices which each contain one of
¹⁶ the three mixing angles.

¹⁷

¹⁸ Should neutrinos be Majorana particles, then the U_{PMNS} matrices should be multiplied
¹⁹ by $\text{diag}(e^{i\alpha_1/2}, e^{i\alpha_2/2}, 1)$. The question of whether neutrinos are Majorana or Dirac particles
²⁰ does raise important questions for neutrino physics. This is because should neutrinos be
²¹ Majorana particles, their masses could be generated via a Majorana mass term. There are
²² next generation experiments such as SNO+ [21] and SuperNEMO [22], which will search for
²³ neutrinoless double beta decay as a means to test whether neutrinos are Majorana particles.
²⁴ However, the two Majorana phases do not affect neutrino oscillations, and so will not be
²⁵ covered further in this discussion of neutrino oscillations.

²⁶

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.

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

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 about $\sqrt{1+x^2}$ 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.

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

1.1 Neutrino physics

28

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

$$\begin{aligned} 6 \quad A(v_\alpha \rightarrow v_\beta) &= \langle v_\beta | v_\alpha(t) \rangle \\ 7 \quad &= \sum_k \sum_j \langle v_j | U_{\beta j} U_{\alpha k}^* e^{-im_k^2 L/2E} | v_k \rangle \\ 8 \quad &= \sum_k U_{\alpha k}^* U_{\beta k} e^{-im_k^2 L/2E} \end{aligned} \quad (1.12)$$

- 10 Equation 1.12, can then be used to get the probability for the original neutrino v_α to oscillate
 11 to a different flavour v_β :

$$\begin{aligned} 12 \quad P(v_\alpha \rightarrow v_\beta) &= |\langle v_\beta | v_\alpha(t) \rangle|^2 \\ 13 \quad &= \left| \sum_k U_{\alpha k}^* U_{\beta k} e^{-im_k^2 L/2E} \right|^2 \\ 14 \quad &= \sum_k U_{\alpha k}^* U_{\beta k} e^{-im_k^2 L/2E} \sum_j U_{\alpha j}^* U_{\beta j} e^{-im_j^2 L/2E} \end{aligned} \quad (1.13)$$

- 17 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
 18 the remainder of the discussion of neutrino mixing.

- 19
 20 In the interests of simplicity, an explicit calculation of the neutrino oscillation probability
 21 will be given assuming that there are only two neutrino flavour and mass states. The reason
 22 for this is that then there is only one mixing angle (θ), and no complex phase. We are also free
 23 to choose the simplest mixing matrix, such that the mixing matrix becomes Equation 1.15.

$$24 \quad \begin{pmatrix} v_\alpha \\ v_\beta \end{pmatrix} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \quad (1.15)$$

1.1 Neutrino physics

29

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.

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

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.

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

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). 1
- The distance the neutrino travels (L). 2
- The energy of the neutrino (E). 3
- The magnitude of the matter effects (a). 4

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.

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 $\overline{\nu}_\mu \rightarrow \overline{\nu}_e$. The asymmetry between neutrino and antineutrino oscillations is given by Equation 1.20.

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

However, the observed asymmetry between $\nu_\mu \rightarrow \nu_e$ and $\overline{\nu}_\mu \rightarrow \overline{\nu}_e$ 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.

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

10

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

18

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

26

27 **1.1.3 Current experimental limits, and unanswered questions**

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

32

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

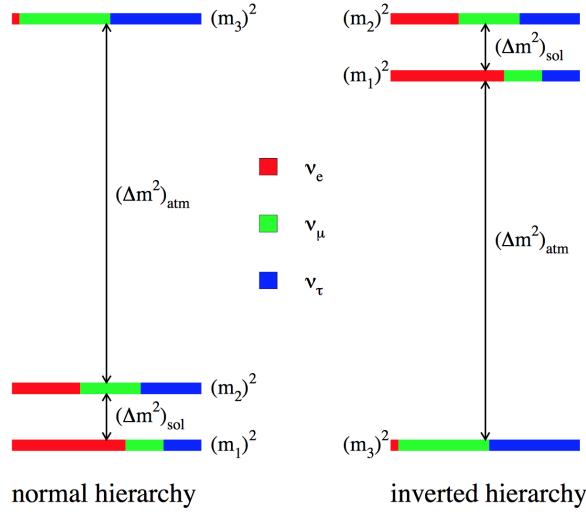


Fig. 1.1 A schematic representation of the two mass hierarchies, showing the fractional components of flavour states in each mass state. Left: the normal hierarchy, where the masses are ordered $m_1 < m_2 < m_3$. Right: the inverted mass hierarchy, where the masses are ordered $m_3 < m_1 < m_2$. The 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}	$145 \rightarrow 391$	$0 \rightarrow 360$
$\Delta m_{21}^2 \times 10^5 (\text{eV}^2)$	$7.50^{+0.19}_{-0.17}$	$7.03 \rightarrow 8.09$	$7.50^{+0.19}_{-0.17}$	$7.03 \rightarrow 8.09$	$7.03 \rightarrow 8.09$
$\Delta m_{3\ell}^2 \times 10^3 (\text{eV}^2)$	$+2.524^{+0.039}_{-0.040}$	$+2.407 \rightarrow +2.643$	$-2.514^{+0.038}_{-0.041}$	$-2.635 \rightarrow -2.399$	$\begin{bmatrix} +2.407 \rightarrow +2.643 \\ -2.629 \rightarrow -2.405 \end{bmatrix}$

¹ unanswered. Some of these have already been discussed earlier, and in no particular order
² are;

- ³ 1. **What is the quadrant of θ_{23} ?** As can be seen from Table 1.1, θ_{23} is very close to 45° ,
⁴ representing maximal mixing. However, whether it is more or less than 45° remains to
⁵ be seen.
- ⁶ 2. **What is the value of δ_{CP} ?** The value of δ_{CP} is largely unconstrained to 3σ , and as
⁷ discussed in Section 1.1.2, the observation of CP violation in neutrinos may help to
⁸ explain the matter-antimatter asymmetry in the universe.
- ⁹ 3. **What is the neutrino mass hierarchy?** As discussed in Section 1.1.2, though it is
¹⁰ known that the m_1 mass state is lighter than the m_2 mass state, it is not known whether
¹¹ these are lighter or heavier than the v_3 mass state.
- ¹² 4. **What is the absolute neutrino mass scale?** The absolute mass of the neutrino is
¹³ unknown, though cosmological constraints, from Planck CMB data, place an upper
¹⁴ limit of $\sum m_\nu < 0.23$ eV at the 95% confidence limit [34]. Neutrino experiments such
¹⁵ as KATRIN hope to reach sensitivities of this magnitude by looking for a cut-off in the
¹⁶ energy spectrum of β decays [35].
- ¹⁷ 5. **Are neutrinos Dirac or Majorana particles?** This concerns how neutrinos acquire
¹⁸ mass, as if they are Dirac particles then they would acquire mass through interactions
¹⁹ with the Higgs field, as the other particles in the SM do. However, if they were
²⁰ Majorana particles, then they would acquire at least some of their mass through self-
²¹ coupling. This would violate lepton number, and mean that neutrinos are their own
²² antiparticle. As discussed in Section 1.1.2, experiments are searching for this by
²³ looking for neutrinoless double β decay.
- ²⁴ 6. **Are there sterile neutrinos?** Experimental data from LEP tightly constrains the
²⁵ number of active neutrinos to those predicted by the SM [36]. However, in recent years
²⁶ there have been anomalous results which may hint at so-called sterile neutrinos [37–39].
²⁷ These sterile neutrinos would not interact, but would still be involved in oscillations.
²⁸ Sterile neutrinos are included in the neutrino mixing formalism by extending the PMNS
²⁹ matrix from a 3×3 matrix, to a $(3+N) \times (3+N)$ matrix, where N is the number of
³⁰ sterile neutrinos. Many short baseline experiments are searching for signatures which
³¹ 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

3

1.2 Grand Unified Theories and nucleon decay

4

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.

5

6

7

8

9

10

11

1.2.1 Overview of grand unified theories

12

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].

13

14

15

16

17

18

19

20

21

22

23

24

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

25

26

27

28

29

30

31

1.2 Grand Unified Theories and nucleon decay**36**

¹ proton lifetimes of the order $\tau_p \leq 10^{35}$ years [44, 45].

²

³ SUperSymmetrY (SUSY) is an extension of the standard model which aims to rectify
⁴ many of the inconsistencies seen in the SU(5) models. It does this by predicting that there
⁵ are “superpartners” to the bosons and fermions predicted by the SM, which differ from their
⁶ SM counterparts by a spin factor of $\frac{1}{2}\hbar$. An “s” is added to the names of the SM fermions
⁷ (sup, selectron), and “ino” is added to the names of the SM bosons (wino, gluino), to denote
⁸ the difference between SM particles and their superpartners [46]. The Minimally Super-
⁹ symmetric Standard Model (MSSM) is the simplest SUSY theory, and it predicts a single
¹⁰ superpartner to each of the known SM fermions and bosons [47]. Supersymmetric SU(5)
¹¹ models have been popular for many years, and even predicted the value of the Weinberg
¹² angle, and that the top quark would have a mass of around 200 GeV [43]. Though the initial
¹³ proton lifetimes of $\tau_p \simeq 10^{30-31}$ years have now been ruled out, it is possible to get proton
¹⁴ lifetimes of $\tau_p \simeq 10^{34-36}$ years, which are well within the current experimental limits [48].

¹⁵

¹⁶ In the MSSM it is possible to have proton decay lifetimes of less than 1 s, through
¹⁷ baryon and lepton number violating interactions, unless a symmetry called *R*-parity [49] is
¹⁸ introduced (Equation 1.21).

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

²⁰ where *B* is baryon number, *L* is lepton number, and *s* is spin. SM particles have *R* parity +1,
²¹ whilst supersymmetric particles have *R*-parity -1. The baryon - lepton number (*B* - *L*) is
²² generally conserved in GUTs, as a result of *R*-parity. However, in order to explain neutrino
²³ mass and mixing, *R*-parity must be broken [43]. This has the unfortunate consequence of
²⁴ making the lightest neutralino unstable, meaning that it cannot be the Lightest Supersym-
²⁵ metric Particle (LSP), as it would decay too quickly to explain dark matter. This then means
²⁶ that the best candidate for the LSP is the unstable gravitino [43]. However, after allowing
²⁷ for *R*-parity to be broken, additional channels of nucleon decay become possible, such as
²⁸ $n \rightarrow K^+ + l^-$ and $p \rightarrow K^+ + l^- + \pi^+$. It can be shown that the decay rate of the $n \rightarrow K^+ + l^-$
²⁹ channel may be an order of magnitude larger than that of the proton [43, 50]. As will be
³⁰ shown in Table 2.3, these channels currently have very low limits on their lifetimes and so
³¹ warrant further study, as nucleon decay signatures may potentially be observed with relatively
³² low lifetimes.

³³

³⁴ Finally, it is also possible to construct GUTs which use higher order gauge groups,
³⁵ such as the SO(10) group. When considering SO(10) theories, supersymmetric models are
³⁶ 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 range from $10^{33\text{--}36}$ years, with an upper limit of 10^{38} years in the $p \rightarrow K^+ \bar{\nu}$ decay channel [51].

Over the years there have been many experiments which have searched for the signatures of nucleon decay, these have included NUSEX [52], FREJUS [53], Soudan I/II [54], and IMB [55]. The current experimental limit on the proton lifetime is held by the SK experiment, which sets a lower partial lifetime in the $p \rightarrow e^+ \pi^0$ decay mode of 1.6×10^{34} years [56]. A successor to SK, called Hyper-Kamiokande (HK), is planned for the mid 2020s [57] and will increase the proton decay lifetime limits measured by SK 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 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

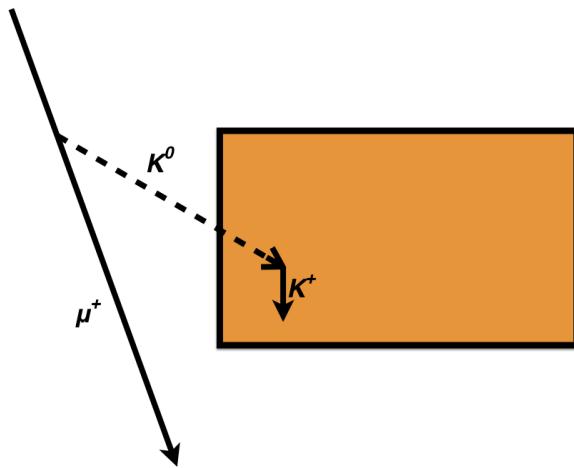


Fig. 1.2 Schematic of a muon event mimicking a nucleon decay. A cosmic muon produces a K_L^0 outside of the detector volume, which then interacts producing an isolated kaon. There would be no other charged tracks produced, and so the event would be indistinguishable from a real proton decay event.

1 tracks coming from outside the detector (the K_L^0 is neutral and so does not produce a track
 2 in LArTPCs). This event structure would look very similar to a nucleon decay event, and
 3 a schematic of such an interaction is shown in Figure 1.2. It is important to note that as
 4 LArTPCs are not magnetised it is very difficult to discern the charge of a kaon, and so a K^\pm
 5 are largely indistinguishable.

6
 7 Atmospheric neutrinos are also able to produce signals which mimic nucleon decays
 8 if they interact within the active volume. The simulations described in Chapter 6 concern
 9 backgrounds induced by cosmic ray muons. The backgrounds induced by atmospheric
 10 neutrinos will not be considered in this thesis.

11
 12 Though background inducing events like the one shown in Figure 1.2 are rare, the nucleon
 13 decay rate is very low, and so it is necessary to identify background events and reject them
 14 when searching for a signal. For this reason a fiducial cut is often applied, as many of the K_L^0 s
 15 will interact close to the detector edge. Should an interaction pass this cut, it is possible to
 16 apply strict energy criteria to the reconstructed energies. This is because the energies of the
 17 particles produced in a nucleon decay event are well defined. This will be discussed further
 18 in Chapter 6 in the context of the $n \rightarrow K^+ e^-$ decay channel. In the case of the $p \rightarrow K^+ \bar{\nu}$
 19 decay channel, one would expect the kaon to have a momentum of about 340 MeV, and a
 20 total energy of about 600 MeV. The kaon would also be expected to decay at rest, and so
 21 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.

1

2

3

¹ Chapter 2

² The Deep Underground Neutrino ³ Experiment

⁴

⁵ 2.1 Brief experimental overview

⁶ The Deep Underground Neutrino Experiment (DUNE) is a next-generation neutrino experi-
⁷ ment, which came about through the recommendations of the P5 report in 2014 [58]. It was
⁸ largely formed by the merger of two next-generation neutrino experimental proposals, the
⁹ Long Baseline Neutrino Experiment (LBNE) [59–64] a US-based experiment, and the Long
¹⁰ Baseline Neutrino Observatory (LBNO) [65] a European-based experiment.

¹¹

¹² The DUNE Far Detector (FD) will consist of four modules, each of which is a Liquid
¹³ Argon (LAr) Time Projection Chamber (TPC) (LArTPC), with an approximate fiducial mass
¹⁴ of 10 kton. The four FD modules will be 4850 ft below ground, at the Sanford Underground
¹⁵ Research Facility (SURF). The first module will be a single phase (liquid) LArTPC and will
¹⁶ be operational before the end of 2024. The subsequent modules will be a combination of
¹⁷ single phase and two phase (liquid/gas) detectors, each becoming operational in 2025, 2027
¹⁸ and 2028. No details will be provided in this thesis concerning the two phase detector design,
¹⁹ for details see [66]. A staged approach is used to ensure that physics measurements can begin
²⁰ before all of the modules are constructed.

²¹

²² The FD location at SURF is 1300 km away from Fermi National Laboratory (FNAL,
²³ Fermilab), from which DUNE will receive a neutrino beam with a beam power of 1.2 MW.

2.2 The single phase detector design**41**

The beam will become operational in 2026, meaning that two of the four FD modules will
1
be constructed before the beam is finished. Therefore, the first two FD modules will have
2
been fully characterised, and will have taken atmospheric and solar neutrino data before the
3
beam is operational. The beam power will increase to 2.4 MW during the lifetime of the
4
experiment. The exact specifications of the beam design are currently being developed [67],
5
though it will use the Main Injector beamline at FNAL after the PIP-II upgrade [68]. The
6
neutrino beam will be able to be configured to run in either neutrino or antineutrino mode by
7
“sign-selecting” the mesons which are produced after protons hit a graphite target.
8

DUNE will also have a high resolution and high precision Near Detector (ND) located
10
at Fermilab, approximately 600 m downstream from the target. The exact nature of the
11
ND is still being decided, though the reference design is a Fine-Grained Tracker (FGT).
12
This FGT consists of a Straw-Tube Tracking module, and an Electromagnetic Calorimeter,
13
inside a 0.4 T dipole magnet. The DUNE ND will also have a set of muon identifiers in
14
the dipole magnet. A detector of this structure will allow the DUNE ND to make precision
15
measurements of the neutrino flux and interaction cross section, as well as the signal and
16
background rates. There is also the possibility of building an additional LArTPC ND, though
17
this would suffer from neutrino pile-up, due to the slow drift time of LArTPCs, and the high
18
luminosity of the beam.
19

A schematic of the DUNE experimental setup is shown in Figure 2.1. Given this exper-
21
imental setup, DUNE will have a wide range of physics opportunities, these are discussed
22
in detail in Section 2.3. There is extensive documentation concerning all aspects of DUNE,
23
this can be found in the Conceptual Design Report (CDR) documents [69, 70, 67, 66], and
24
numerous annexes [71]. What follows is a high level summary of the information contained
25
in these CDR volumes, and annexes.
26

2.2 The single phase detector design

28

As the work presented in this thesis has been performed on a prototype of the single phase
29
design (Chapters 3, 4 and 5), and a simulation of the FD single phase design (Chapter 6),
30
the DUNE FD single phase design will be briefly described below. As no work has been
31
performed in relation to the two phase design, no information will be presented concerning
32
this design. Detailed descriptions of both the single and two phase FD designs can be found
33

2.2 The single phase detector design

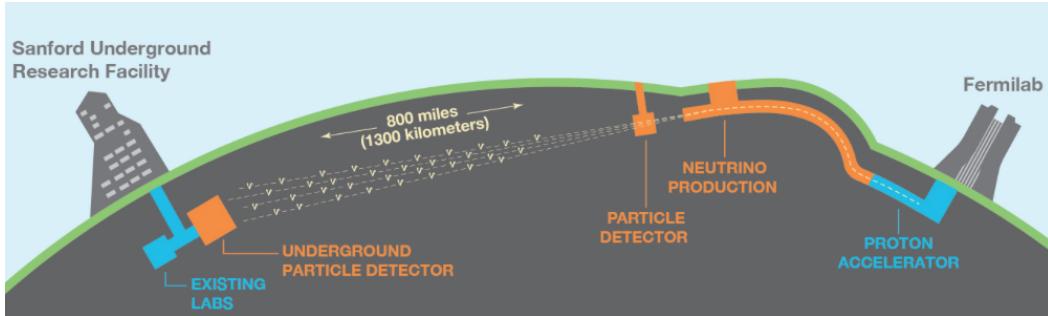


Fig. 2.1 A representation of the DUNE experimental setup. Fermilab, the host lab is shown right. The neutrino beam will be produced at Fermilab, and a high resolution near detector will also be located there. SURF is shown left, it is 1300 km away from Fermilab, and the far detector will be located there. 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 [69].

¹ in [66].

²

³ In the single phase design, the charge generation, electron drift, and charge collection, all
⁴ occur in the LAr. As the charge collection occurs in the LAr, this means that all of the TPC
⁵ structure, including the electronics, are submerged. The active volume of 14.1 kt, measuring
⁶ 12 m high, 14.5 m wide (in the drift direction), and 58 m long (in the beam direction), is
⁷ too big to be contained in a single TPC, and so many individual TPCs are required. This
⁸ is because it is not practical to collect electrons after they drift over many metres, or to
⁹ construct an Anode Plane Assembly (APA) frame and Cathode Plane Assembly (CPA) frame
¹⁰ across 10 metres or more. The mine shafts would also not be able to accommodate APA
¹¹ and CPA frames which were over 10 metres long or wide. This limit on size also affects
¹² the material used to make the cryostat, and so it will be constructed using a stainless steel
¹³ membrane design. This design is commonly used for liquefied natural gas storage and
¹⁴ transport tanker ships, and so is an established technology. However, it had not been used in
¹⁵ LArTPC experiments before the 35 ton prototype detector, described in Section 2.4.

¹⁶

¹⁷ The active volume of the DUNE FD is divided into 200 TPCs, each produced by applying
¹⁸ voltages to APAs and CPAs, and contained within a field cage. Each TPC measures 6 m
¹⁹ high, 3.6 m wide (in the drift direction), and 2.3 long (along the beam direction). The TPCs
²⁰ are arranged such that there are 4 TPCs across the width of the detector, 2 TPCs stacked
²¹ vertically on top of each other, and 25 TPCs along the length of the detector, to give the full
²² active volume size of $12 \times 14.5 \times 58 \text{ m}^3$. This experimental setup means that the detector is
²³ 58 m long in the beam direction, offering a large region where neutrinos interaction can occur.

2.2 The single phase detector design

43

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

Function	Orientation (°)	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

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 wires 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 are not read out. The next two wire planes are induction wire planes (U and V), and have a bias 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

2.2 The single phase detector design

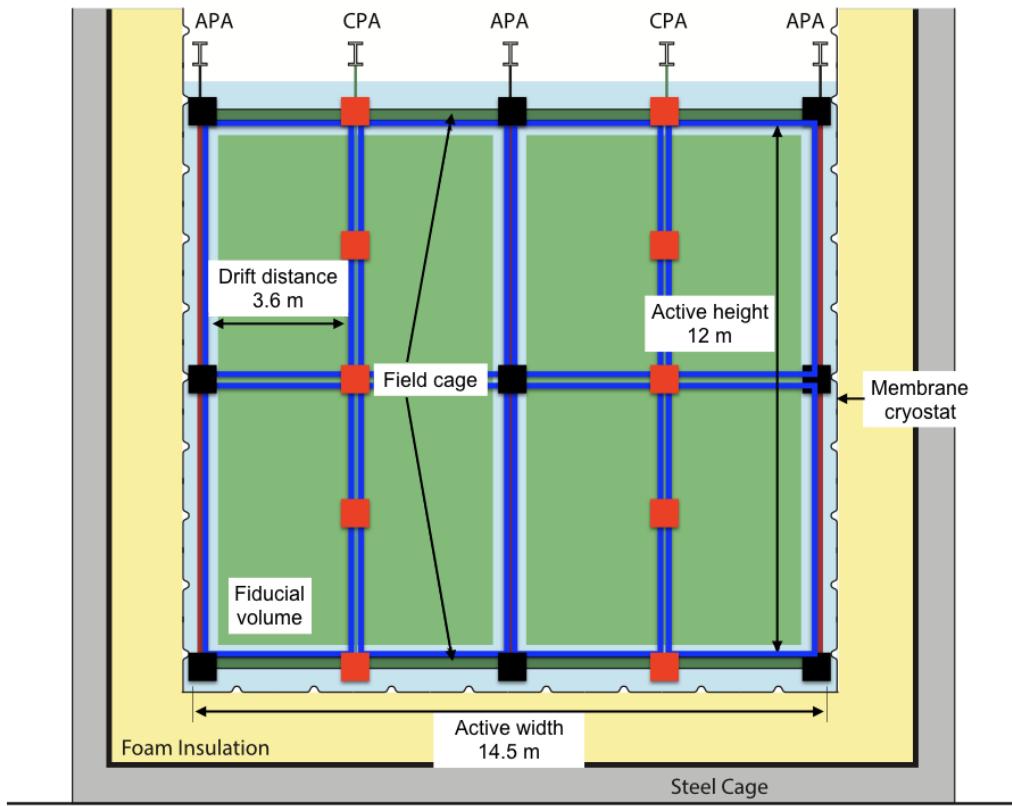


Fig. 2.2 A cross section of the DUNE signal phase detector design, showing 8 TPCs. The cross section shown is looking along the beamline, such that the detector extends a total of 58 m into the page. The TPCs (blue squares) are bound between a pair of APAs and CPAs, and are surrounded by a field cage (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 above 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 24 more cross sections going into the page to produce a total of 200 TPCs. The figure has been modified from [66].

2.3 The physics capabilities of DUNE

45

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.

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 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 [70]:

- 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.

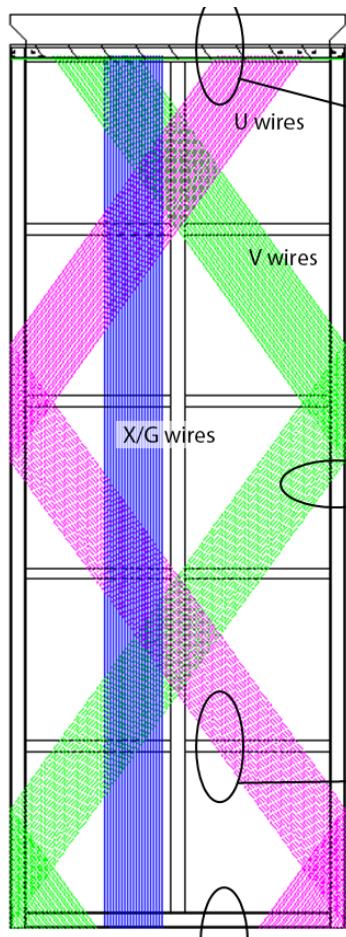


Fig. 2.3 An illustration of the wire wrapping in the DUNE single phase design. A small number of the wires in each plane are shown. The instrumented wire planes, in the order that the electrons drift past them, are shown as follows; the 1st induction plane (U) is magenta, the 2nd induction plane (V) is green, and the collection plane (X) is blue. The grid plane (G), is in front of the 1st induction plane and is parallel to the collection plane. The figure is taken from [66].

2.3 The physics capabilities of DUNE

47

Table 2.2 The beam exposure in units in units of $\text{kt}\cdot\text{MW}\cdot\text{years}$, as a function of time, assuming a staged DUNE construction. This table uses the fiducial mass of the detector (10 kt), the luminosity of the beam (1.2-2.4 MW), and the time that the detector is exposed to the beam (years). The table is taken from [72].

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

- Measuring a wide range of neutrino cross-sections, using the ND. 1
- Measuring nuclear effects, particularly neutrino final-state interactions, using the ND. 2

The beam exposure as a function of time which DUNE will receive is shown in Table 2.2, 3
and is important when considering the figures presented below. 4

As presented in Section 1.1, both the matter effect, and δ_{CP} introduce an asymmetry 5
between neutrino and anti-neutrino oscillations. As the matter effect is caused by the difference 6
in the presence/absence of electrons/positrons in the Earth, it increases with distance. The 7
result of this is that for baselines longer than approximately 1000 km the two effects can be 8
resolved [73]. For this reason DUNE will have a baseline of 1300 km so that it is able to 9
unambiguously determine the neutrino mass hierarchy, and determine δ_{CP} [74]. 10

Figure 2.4 shows the oscillation probabilities of $\nu_\mu \rightarrow \nu_e$ for both neutrino and anti- 13
neutrino modes at a baseline of 1300 km. The reason for requiring a broadband neutrino 14
beam at this baseline is clearly apparent, as it can be seen that whilst the energy of the first 15
neutrino oscillation maximum is relatively unaffected by the value of δ_{CP} , the energies of the 16
higher oscillation maxima are strongly affected. It is therefore vital that DUNE is able to 17
accurately measure the rate of ν_e appearance at the lowest energies of the neutrino beam. It 18
can also be seen that there are large differences in the expected oscillation probabilities for 19
neutrinos and anti-neutrinos. Therefore, in order to measure the effect of δ_{CP} , the neutrino 20
beam will have to be able to operate in both neutrino and anti-neutrino mode. 21

Figures 2.5 and 2.6 show the DUNE sensitivities to the mass hierarchy, and the value of 23
 δ_{CP} , with beam exposures corresponding to 7 and 10 years worth of data taking [72]. These 24
beam exposures are calculated given the phased approach to DUNE which was outlined 25

2.3 The physics capabilities of DUNE

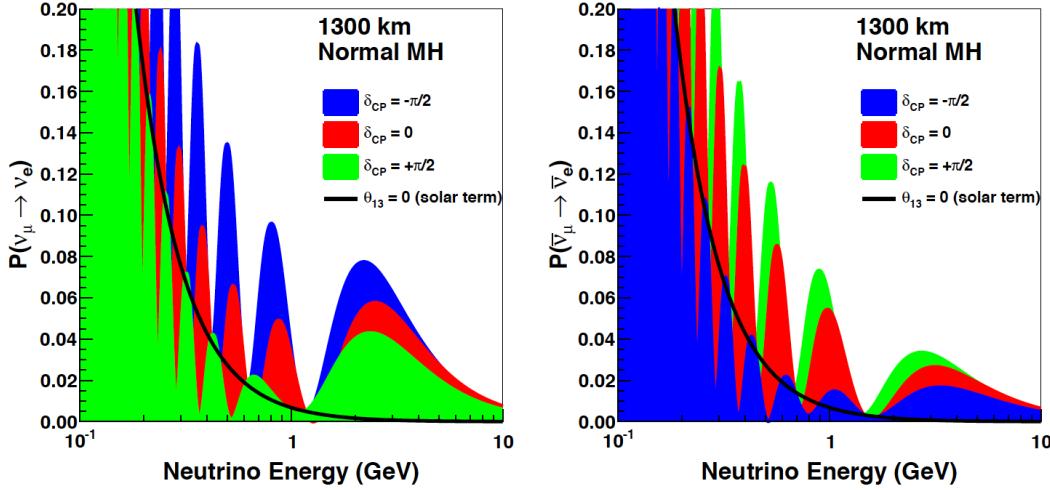


Fig. 2.4 The ν_e appearance probability at 1300 km as a function of neutrino energy, for a range of values of δ_{CP} . Left: the oscillation probability for neutrinos. Right:, the oscillation probability for anti-neutrinos. Both figures assume normal mass hierarchy (ν_1 is the lightest state). The probabilities 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 [70].

1 in Table 2.2, meaning that a beam exposure of 7 years corresponds to a beam exposure
 2 300 kt·MW·years, whilst a beam exposure of 10 years corresponds to a beam exposure of
 3 556 kt·MW·years. The best fit values from NuFit 2016 [33] are used when making these fig-
 4 ures, and equal running in neutrino and antineutrino mode is assumed. Figures 2.8, 2.9, 2.10
 5 and 2.11 are taken from the CDR [70], and show the effect that the two beam designs [67]
 6 have on the sensitivities, assuming beam exposures of 300 kt·MW·years.
 7

8 Figure 2.5 shows the significance with which DUNE will be able to determine the neu-
 9 trino mass hierarchy, for all values of δ_{CP} . It can be seen that the mass hierarchy can be
 10 determined with a significance of $\sqrt{\Delta\chi^2} = 5$, for all values of δ_{CP} after 7 years of data
 11 taking. It can also be seen that the mass hierarchy can be more conclusively determined if it
 12 is inverted.
 13

14 Figure 2.6 shows the significance with which DUNE will be able to determine the value of
 15 δ_{CP} . It can be seen that even with 10 years worth of data, there are regions where the value of
 16 δ_{CP} cannot be determined accurately because of the complex phase being 0 when δ_{CP} is equal
 17 to $-\pi$, 0, or π . Therefore, for values of δ_{CP} around these values, the significance to which
 18 δ_{CP} can be determined approaches 0. As such, even at very large beam exposures of over
 19 800 kt·MW·years, corresponding to around 13 years of data taking according to Table 2.2,

2.3 The physics capabilities of DUNE

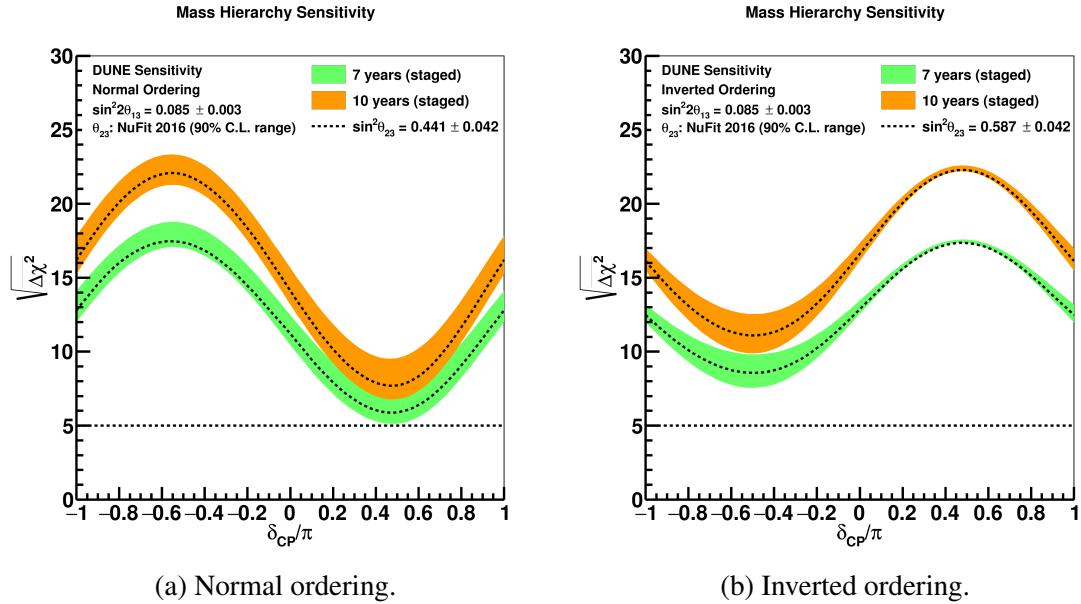


Fig. 2.5 The significance with which DUNE will be able to determine the neutrino mass hierarchy, for all values of δ_{CP} . Left: the sensitivity assuming normal ordering. Right: the sensitivity assuming inverted ordering. The shaded region shows the range of sensitivities for the 90% confidence level range for θ_{23} values, the dashed line shows the sensitivity for the NuFit central value of θ_{23} . The figure is taken from [75].

only 75% of the δ_{CP} values can be determined to a significance of 3σ [76]. 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) [76]. 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 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.

2.3 The physics capabilities of DUNE

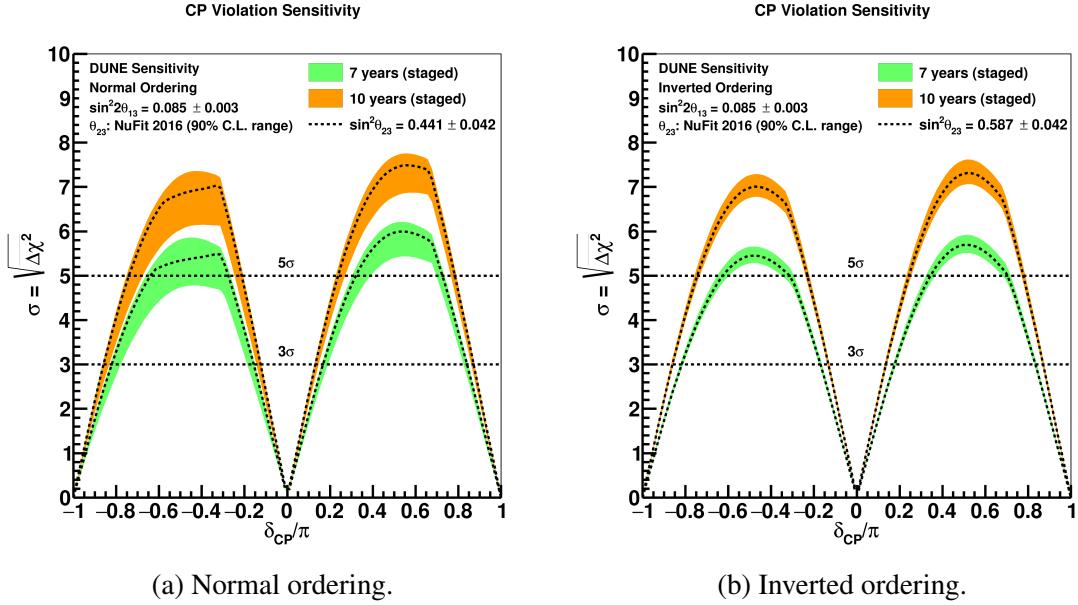


Fig. 2.6 The significance with which DUNE will be able to determine the value of δ_{CP} , for all values of δ_{CP} . Left: the sensitivity assuming the normal ordering. Right: the sensitivity assuming inverted ordering. The shaded region shows the range of sensitivities for the 90% confidence level range for θ_{23} values, the dashed line shows the sensitivity for the NuFit central value of θ_{23} . The figure is taken from [77].

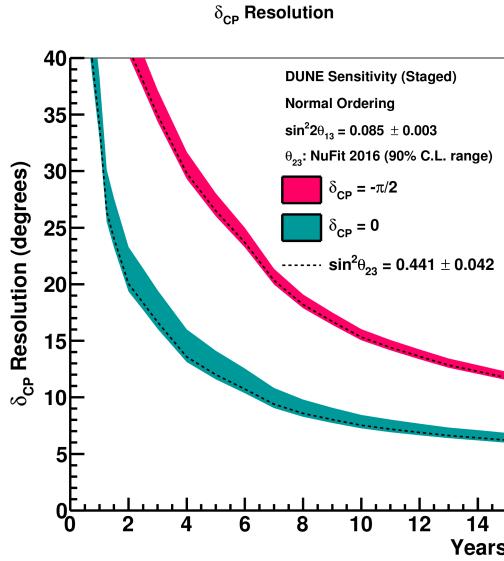


Fig. 2.7 The resolution with which DUNE will be able to determine the value of δ_{CP} , for increasing beam exposures (in years). Two bands are shown, one for a value of δ_{CP} which could cause maximal CP-violation ($\delta_{CP} = 90^\circ$), and another where there would be no CP-violation ($\delta_{CP} = 0^\circ$). 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 [78].

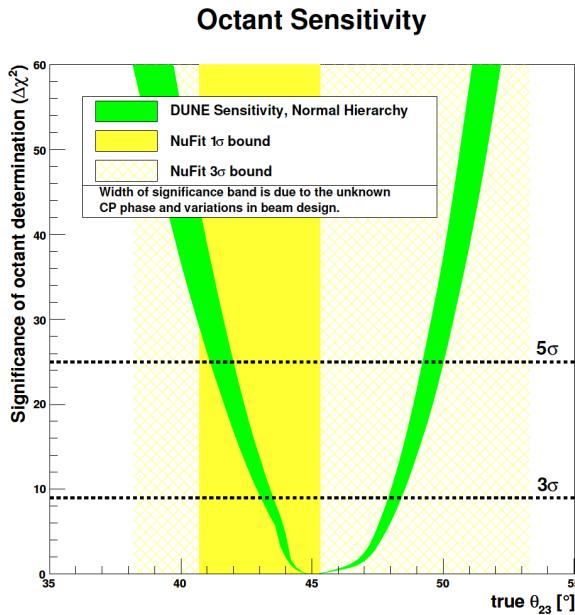


Fig. 2.8 The significance to which the octant of θ_{23} can be determined, for different values of θ_{23} . A beam exposure of 800 kt·MW·years is assumed, offering a 3σ determination of the value of δ_{CP} for 75% of the values of δ_{CP} . The green band shows the effect of different values of δ_{CP} and beam configurations [67]. The yellow regions show the 1σ and 3σ bands. The figure is taken from [70].

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 $v_\mu \rightarrow v_\mu$ and $v_\mu \rightarrow v_e$, 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 v_e and \overline{v}_e appearance, as opposed to \overline{v}_e disappearance, as 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 [70].

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17

2.3 The physics capabilities of DUNE

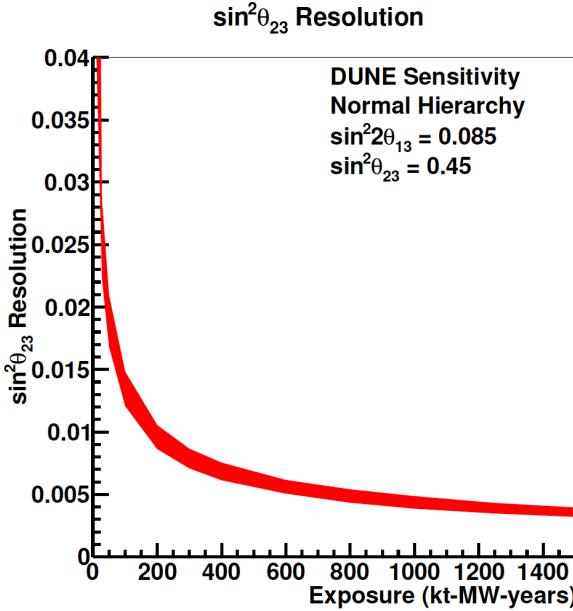


Fig. 2.9 The resolution with which DUNE will be able to determine the value of $\sin^2 \theta_{23}$ for increasing beam exposures. A normal hierarchy is assumed, and the band shows the sensitivity to the potential beam designs [67]. The values for $\sin^2 2\theta_{13}$ and $\sin^2 \theta_{23}$ are taken from [79]. The figure is taken from [70].

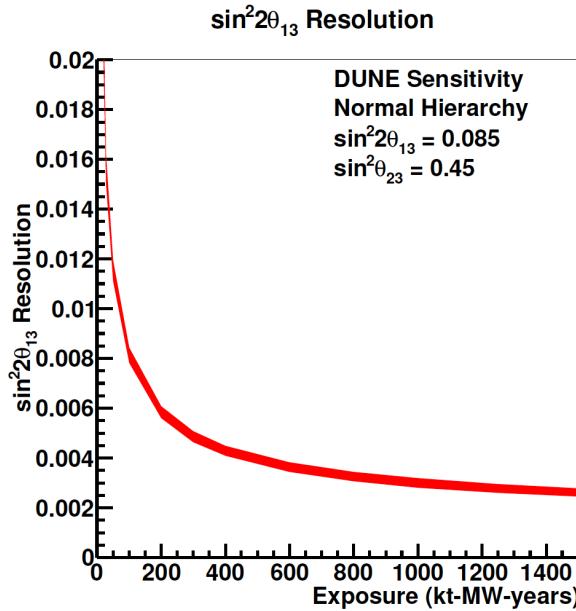


Fig. 2.10 The resolution with which DUNE will be able to determine the value of $\sin^2 \theta_{13}$ for increasing beam exposures. A normal hierarchy is assumed, and the band shows the sensitivity to the potential beam designs [67]. The values for $\sin^2 2\theta_{13}$ and $\sin^2 \theta_{23}$ are taken from [79]. The figure is taken from [70].

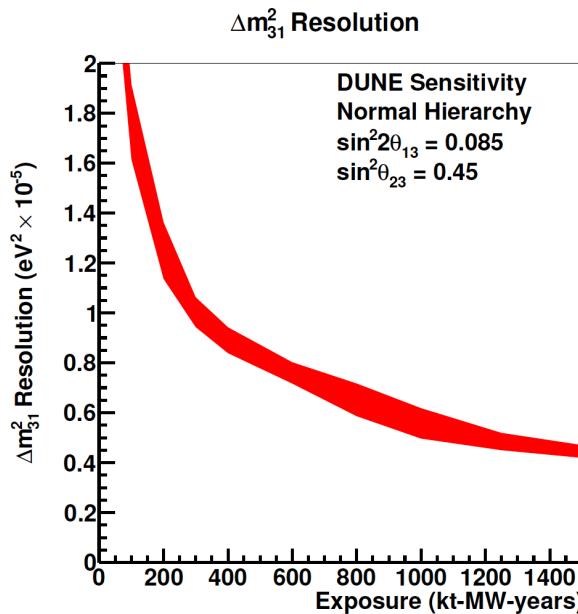


Fig. 2.11 The resolution with which DUNE will be able to determine the value of Δm_{31}^2 for increasing beam exposures. A normal hierarchy is assumed, and the band shows the sensitivities to the potential beam designs [67]. The values for $\sin^2 2\theta_{13}$ and $\sin^2 \theta_{23}$ are taken from [79]. The figure is taken from [70].

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 [70].

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. 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.

2.3 The physics capabilities of DUNE

1 The search for nucleon decay has been primarily focused on two proton decay modes,
 2 $p \rightarrow e^+ \pi^0$ and $p \rightarrow K^+ \bar{\nu}$. In the first of these decay modes, the total mass of the proton
 3 should be converted into the electromagnetic showers produced by the two particles, and
 4 their net momentum should be zero. This is a signal which can be clearly identified in
 5 water Cherenkov detectors, such as Super-Kamiokande (SK), and so is the main decay mode
 6 which these detectors look for. This decay mode should also produce a clear signal in a
 7 LArTPC such as DUNE, though the limit that DUNE will place on this decay mode will
 8 not be competitive with that of Hyper-Kamiokande (HK), the successor to SK, due to the
 9 much smaller mass of DUNE. However, the second decay mode is particularly interesting in
 10 DUNE, as kaons should be able to be accurately identified in a LArTPC. This is not the case
 11 in a water Cherenkov detector, as the kaons are not energetic enough to produce Cherenkov
 12 light. This is true for all decay modes which have a kaon in the final state, including neutron
 13 decay modes. It is also important to note that DUNE will search for all types of nucleon
 14 decays, including bound neutron decays, di-nucleon decay modes and neutron-antineutron
 15 oscillations. An analysis concerning the $n \rightarrow K^+ + e^-$ decay channel will be presented in
 16 Section 6.3.

17
 18 It is hoped that DUNE will be able to reach sensitivities to nucleon decay lifetimes of
 19 between $10^{33} - 10^{35}$ years. A comparison of current, and potential lifetime limits, from
 20 a range of experiments is shown in Figure 2.12. It can be seen that DUNE will provide
 21 very stringent limits to nucleon decay lifetimes, which will compete with HK in all decay
 22 modes containing kaons in the final state. Cherenkov detectors are able to perform relatively
 23 background free studies in the $p \rightarrow e^+ \pi^0$ channel, and so due to the much larger mass of HK,
 24 it is able to achieve a limit which is superior to DUNE. In the other channels however, the
 25 excellent spatial resolution of LArTPCs allows DUNE to compete with, and in many cases
 26 improve on, any limit which could be set by HK.

27
 28 When estimating the expected sensitivity of DUNE to a range of nucleon decay modes,
 29 previous studies have been used [82, 83]. Some of the channels where one would expect
 30 a LArTPC, such as DUNE, to have an advantage in signal efficiency when compared to a
 31 water Cherenkov detector, such as SK/HK, are shown in Table 2.3. The accurate tracking of
 32 the kaon, and its subsequent decay products, is the reason for the increased signal efficiency
 33 that is expected in LArTPCs. The ability for LArTPCs to perform tracking to this accuracy
 34 was seen by the ICARUS collaboration, using the T600 detector [84]. Figure 2.13, shows
 35 an example event where a kaon enters the detector and is seen to decay to a muon, which

2.3 The physics capabilities of DUNE

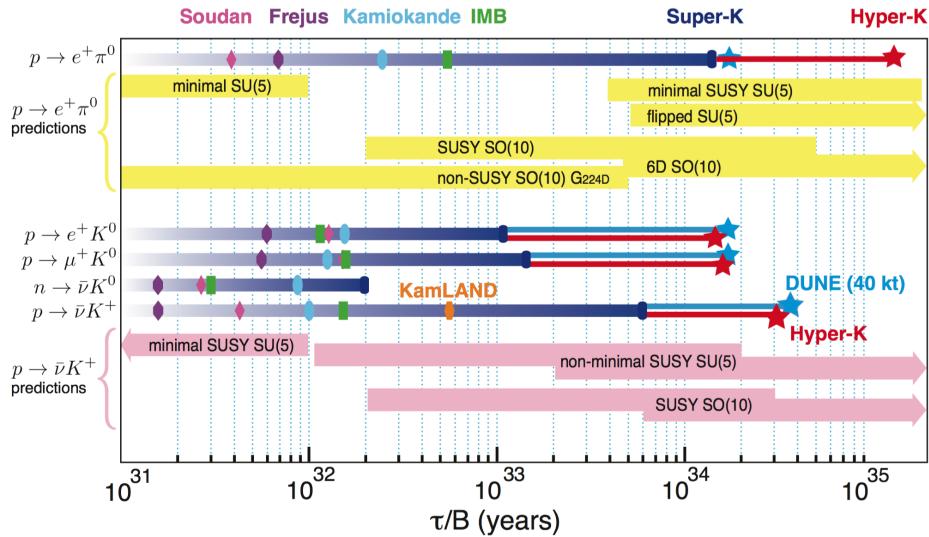


Fig. 2.12 A comparison of current, and future, nucleon decay lifetime limits, compared with the ranges predicted by Grand Unified Theories. Coloured bars are shown for published limits by a number of experiments [80, 81]. Stars are shown for projected limits by future experiments, these limits are calculated using Poisson statistics, and include predicted background rates. The lifetimes predicted by different models are shown for the $p \rightarrow e^+\pi^0$, and $p \rightarrow K^+\bar{\nu}_e$, decay modes. The figure is taken from [70].

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 per Mt yr [83, 86, 87]. With a background rate this low, the observation of a single, well reconstructed event, could provide evidence of nucleon decay [70].

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 [88, 89]. Many 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

2.3 The physics capabilities of DUNE

Table 2.3 Nucleon decay limits in DUNE and Super-Kamiokande, in some favoured decay channels. The current lifetime limits currently measured (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 [85], 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
$p \rightarrow \pi^+ \bar{v}_e$	0.016	41.9	19.8	54	N/A
$p \rightarrow K^+ \bar{v}_e$	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
$n \rightarrow \pi^0 \bar{v}_e$	0.112	45.1	26.0	30	7.1
$n \rightarrow K^+ e^-$	0.032	96	55.4	100	1.8

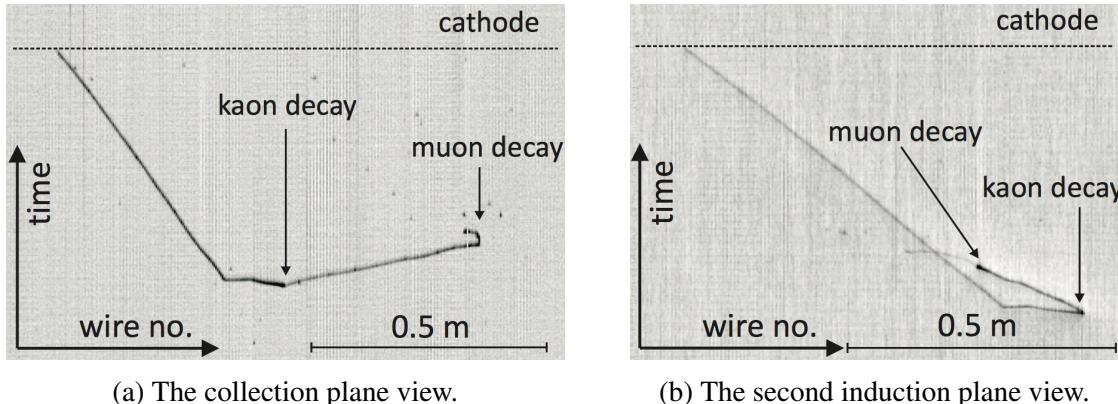


Fig. 2.13 A kaon event which was observed in the ICARUS T600 detector, in the CNGS data. Left: the signal on the collection plane. Right: the signal on the second induction plane. The kaon enters the detector and decays via $K \rightarrow \mu\nu$, the muon then decays via $\mu \rightarrow e\nu$. The figure is taken from [84].

were potentially accompanied by a few gamma rays. 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 [90, 91].

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 [92, 93], but have not observed any signals. It is hoped that with the increased angular resolution of LArTPCs, the number of background events could be substantially reduced [70]. Other aspects where DUNE could make contributions are in searches for the diffuse supernova neutrino background [94], searches for neutrinos from accretion disks [95], and searches for black hole-neutron stars mergers [96].

Hence, there is a very compelling case for building a LArTPC with a fiducial 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

2.4 Path to building DUNE - The 35 ton prototype

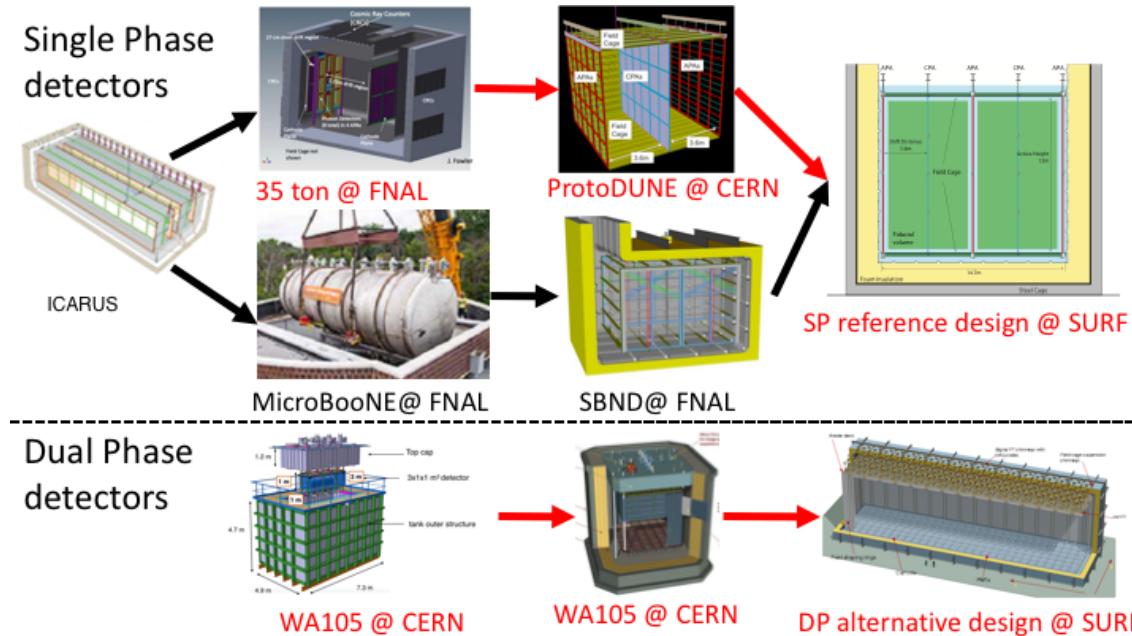


Fig. 2.14 An overview of the DUNE prototyping schedule, including complementary experiments. Top shows the experiments which utilise single phase LArTPC designs, whilst bottom shows experiments which utilise a two phase design. The experiments which are part of the DUNE prototyping program are shown in red. The figures are taken from [66, 97].

1 perform the physics which DUNE is expected to achieve. It was also designed to show that it
 2 was possible to achieve this high level of purity without the need to fully evacuate the detector
 3 prior to filling, but that this could instead be achieved by performing a “piston purge.” This
 4 “piston purge” had previously been demonstrated by the Liquid Argon Purity Demonstrator
 5 (LAPD) [98], which the 35 ton was built next to, at Fermilab’s PC-4 facility. The process by
 6 which the “piston purge” was performed is as follows. Firstly, room temperature argon is
 7 pumped into the detector, displacing the air in the detector. Once the purity is no longer seen
 8 to improve, a gas/liquid spray is used to slowly cool the detector. The injection of the spray
 9 introduces turbulence into the gas in the detector, which causes the entire cryostat to cool,
 10 avoiding large temperature differences. Upon the completion of the cooldown, filling of the
 11 LAr commences, and LAr purification is performed using the installed recirculation loop.
 12 Using this purification method, electron lifetimes of 3 ms were observed after a few days of
 13 purification. These purity levels were maintained for many days at a time, thus demonstrating
 14 that membrane cryostat technology is capable of producing, and maintaining, high purity
 15 LAr [99, 100].

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 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 for 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.

2.4 Path to building DUNE - The 35 ton prototype

60

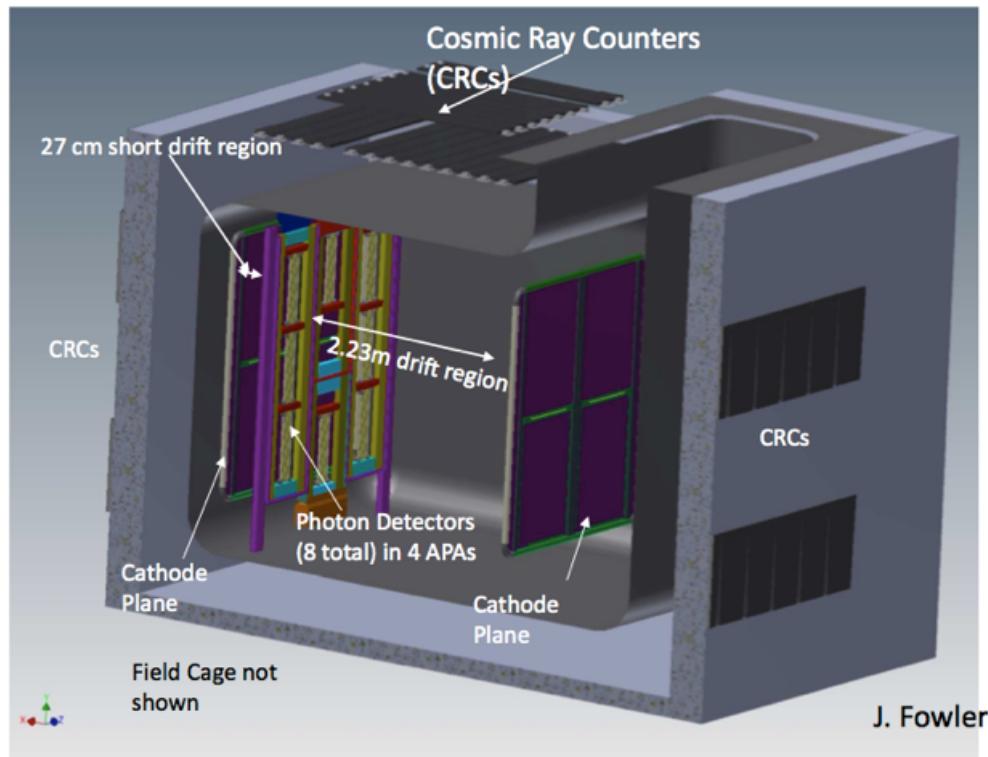


Fig. 2.15 A schematic representation of the 35 ton prototype detector during the Phase II run. The cathode planes (dark purple), photon detectors (yellow), APA frames (bounded light purple regions), cosmic rays counters (outer black rectangles), and the two drift volumes (2.23 m and 0.27 m) are shown. 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 [66].

2.4 Path to building DUNE - The 35 ton prototype

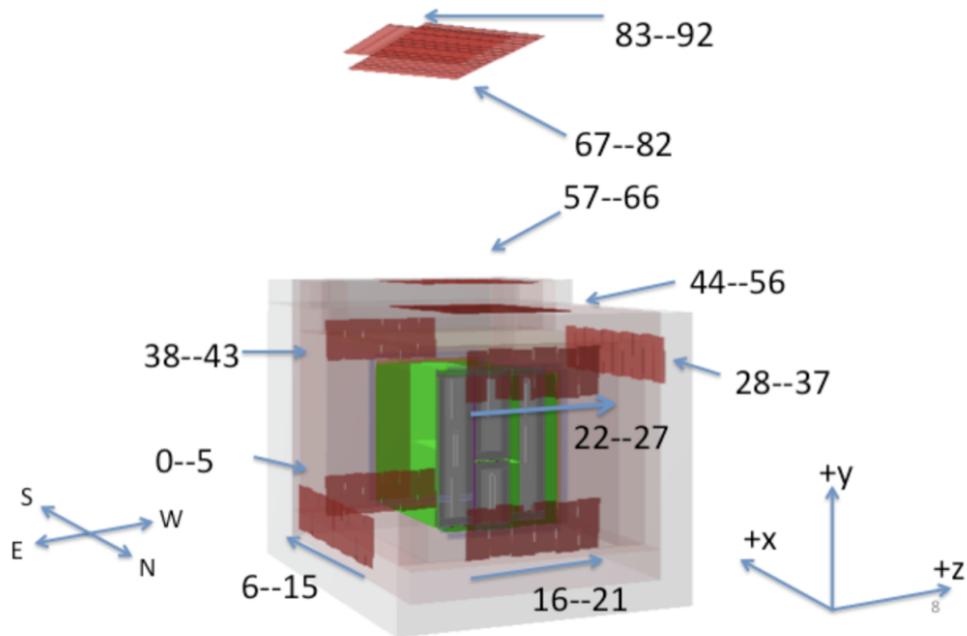


Fig. 2.16 A representation of the counter locations in the 35 ton, with the geographic, and LAr-Soft, coordinate systems shown. The other detector components can be seen inside the cryostat, such that the cosmic ray counters on the north wall are behind the short drift volume. The east - west counters 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.

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.

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

2.5 The DUNE software

62

1 electron lifetime which could be achieved was not limited by the presence of an instrumented
2 detector. The operation of the installed TPC also served to test many of the novel features of
3 the FD reference design. Some of the results from the Phase II run are shown in Chapter 5.

4

5 Following on from the 35 ton Phase II run, the single phase ProtoDUNE detector is under
6 construction at CERN, and is due to take data in the second half of 2018. The ProtoDUNE
7 detector is a small version of the full DUNE single phase reference design, as the APAs, and
8 the drift length, are the same size as those in the reference design. The detector will be in
9 a charged particle test beam, and so will be able to fully test the assumptions made in the
10 DUNE CDRs about the detector performance characteristics. The detector will contain two
11 sets of three APAs, 7.2 m apart, with one set of CPAs between them. This will give a total of
12 6 TPCs, each with a drift length of 3.6 m.

13

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

22

23 2.5 The DUNE software

24 The software package used by DUNE is called LArSoft [101, 102]. LArSoft is a simulation,
25 reconstruction and analysis package for LArTPCs that is being used by many experiments
26 in the US neutrino program. It has been developed to be detector agnostic, meaning that
27 much of the code is shared between experiments. To this end, it is envisioned that it will be
28 used as a platform for constant development in existing experiments, and those still in the
29 planning phases, such as DUNE. LArSoft is built around the Fermilab-supported *analysis*
30 *reconstruction framework (art)*. External packages such as ROOT [103] and GEANT4 [104]
31 are incorporated into LArSoft, meaning that the user does not have to coordinate specific
32 versions of the packages.

33

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) [105], which is used to study neutrino interactions and nucleon decays. Another package, Nuance [106], is a neutrino interaction generator specifically for LAr. Finally, Cosmic RaY shower library (CRY) [107, 108], and COsmic Ray SImulations for KAscade (CORSIKA) [109], 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. Recently the MUon Simulations UNderground (MUSUN) [110, 111] generator, which takes the output of MUon SImulation Code (MUSIC) [110, 112, 113], 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 most 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.
- ϕ - The angle between the z axis and the vector.

2.5 The DUNE software

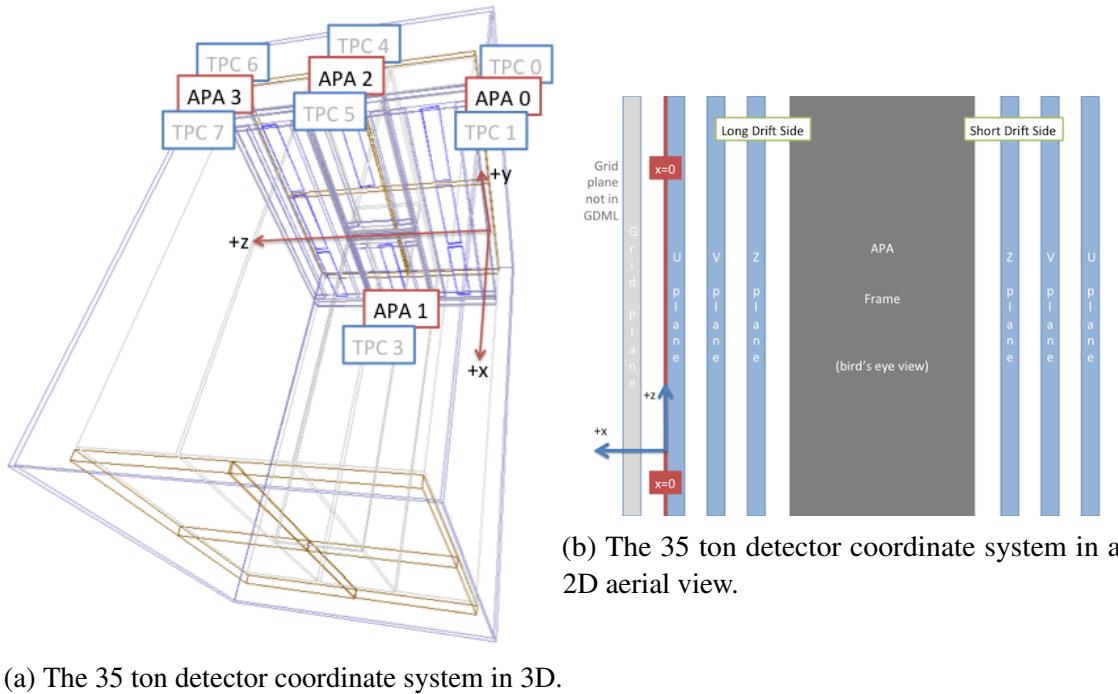


Fig. 2.17 The LArSoft coordinate system as it is represented in the 35 ton detector in 3D (left) and 2D (right). The location of the origin is shown relative to the TPC detector components. Left: the four APAs (purple outlines), and eight TPCs are shown, where the even numbered TPCs are on the “short” 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 [114].

¹ Figure 2.17 shows two schematic representations of how the location of the origin appears in
² the 35 ton prototype.

³
⁴ The simulation of particles is usually split into five processes, to reflect the different areas
⁵ in which development often progresses. These stages are as follows;

- ⁶ • Particle generator.
- ⁷ • Particle transport using GEANT4.
- ⁸ • Full detector simulation, including detector responses.
- ⁹ • Full event reconstruction.
- ¹⁰ • 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 [102, 115].

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 process of hit reconstruction. Figures 2.18a, and 2.18b, show multiple time-separated energy

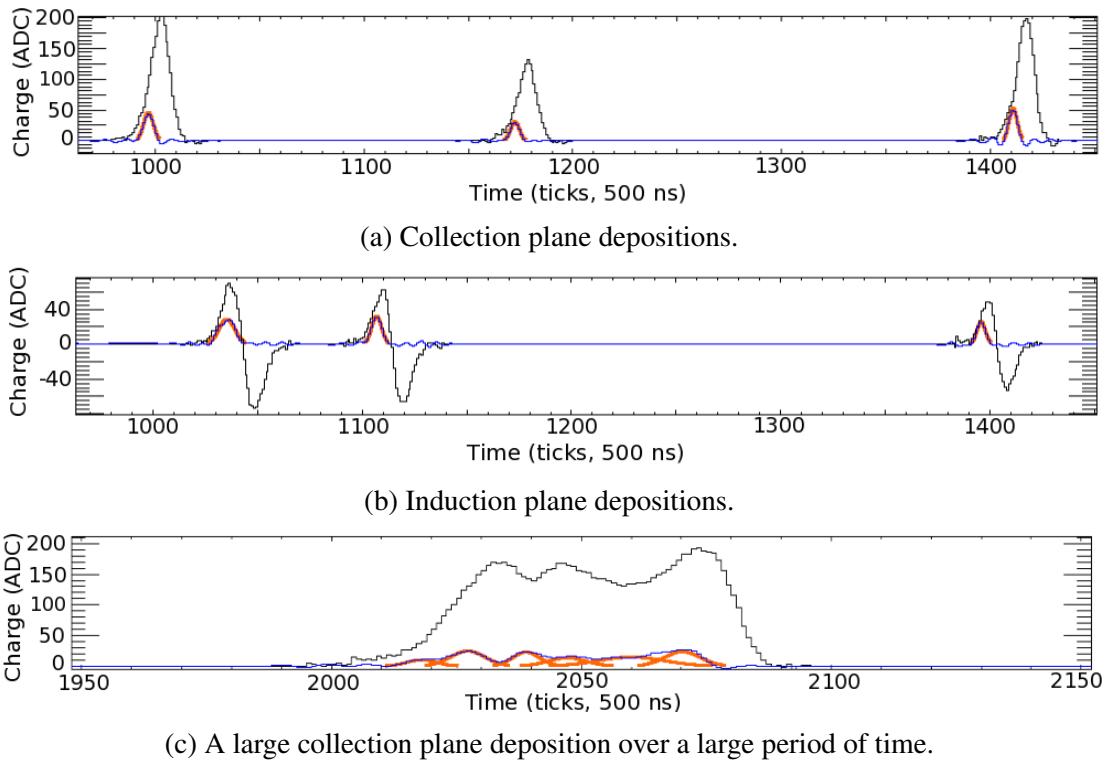


Fig. 2.18 The raw and deconvoluted signals with reconstructed hits for simulated energy depositions. The depositions are from particles generated by CRY and are not from a single event, as such they have been selected for demonstration purposes only. The plots are shown with increasing charge (ADC) on 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.

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 between TPCs in the DUNE FD, in order to reduce the amount of internal dead space, and

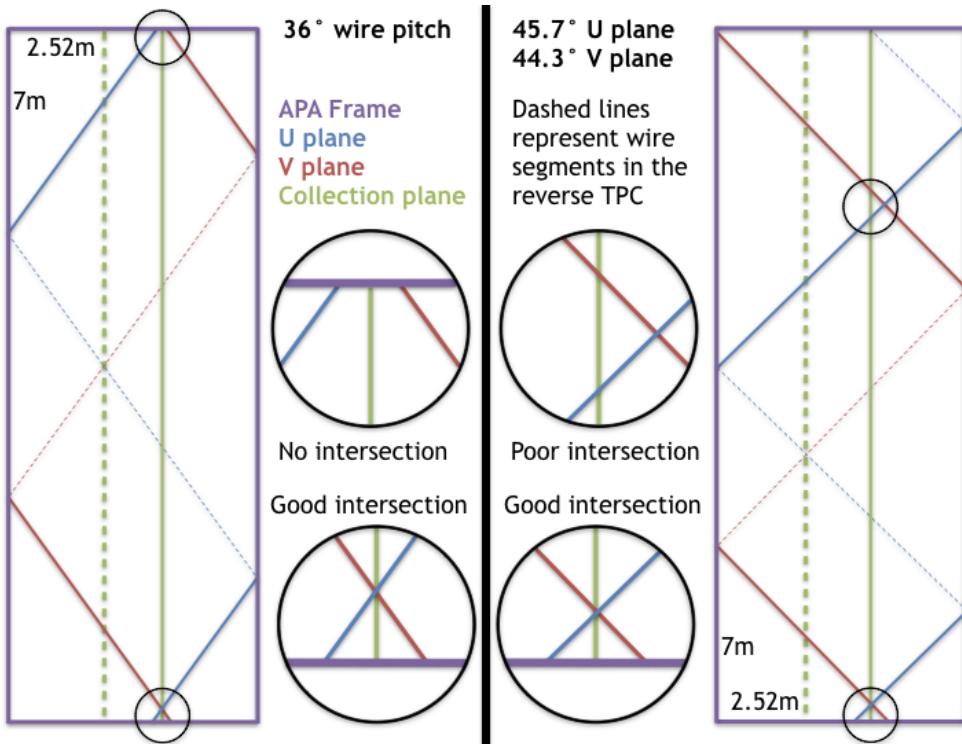


Fig. 2.19 The effect that different wire pitches have on the ability to perform disambiguation in APAs with the far detector geometry. Left: induction plane wires with a wire pitch of 36°, this is the wire pitch that is used in the FD reference design. Right: induction plane wires with a wire pitches 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.

¹ reading out APAs from the sides would introduce large regions of dead space. As a result of
² this, APAs at the top of the cryostat are read out from the top of the APA, whilst APAs at the
³ bottom of the cryostat are read out from the bottom of the APA.

⁴

⁵ Once the hits have been disambiguated they are combined to make clusters in each of the
⁶ three planes. The clustering process is usually performed in wire-tick space on each plane
⁷ separately, where all the hits from a single track or shower, should make a single cluster
⁸ in each plane. It is possible to seed the start of clusters by using imaging techniques such
⁹ as a Harris transform [116], or to identify straight lines by using Hough transforms [117],
¹⁰ though these are rarely used. As hits from a physical entity are unlikely to remain on a single
¹¹ 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 PM-Track [84] and Pandora [118], and the most common showering algorithm is EMShower [119]. 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 builds its outputs by looking at the projection of multiple clusters in all projections, and builds 3D objects which best represent the 2D inputs [102, 120]. 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 [102].

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 [121] and a modified Box model [122]. The modified Box model uses a correction to the Box model [123] at low values of $\frac{dE}{dx}$. Normally the modified Box model is used, as it holds for both large and small ionisations, whereas Birks model experiences difficulties at large ionisations, and the traditional Box model struggles at low $\frac{dE}{dx}$. 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.

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

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

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

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

$$1 \quad \frac{dE}{dx} = \frac{e^\alpha - Recomb_A}{\beta} \quad (2.2a)$$

$$2 \quad \alpha = \frac{10^3 \times \beta}{C_{GeV \rightarrow e^-}} \times (dQ/dx)_{Cor} \quad (2.2b)$$

$$3 \quad (dQ/dx)_{Cor} = \frac{dQ}{dx} \times \frac{C_{lifetime}}{C_{ADC \rightarrow e^-}} \quad (2.2c)$$

$$4 \quad \beta = \frac{Recomb_B}{\rho \times E_{field}} \quad (2.2d)$$

5
6 When performing calorimetry, it is also important that the interaction time is known, so
7 that the x positions of hits can be corrected, as they will initially be reconstructed assuming
8 an interaction time of 0 s. This assumption is made because the beam trigger is placed at a
9 time of $T = 0$ when considering beam induced events. An unknown interaction time causes
10 the hit and track positions to be calculated incorrectly, and is corrected by the user once the
11 interaction time of the track has been determined.

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 [66] on the cathode plane. As the breakdown field has recently been measured to be much lower than this, at around 40 kV cm^{-1} [124], some form of monitoring is required.

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 [124, 125, 98, 126, 127]. 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 [128]. 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. CMOS cameras are also found to have lower leakage currents, improved mobility, and lower thermal noise, at temperatures around 100K [129, 130].

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, and resolutions in both time and space, 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. 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

3.2 The design of the camera system

73

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

14

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 cycled when they were exposed to an elevated, though still cryogenic, temperature. 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

1
2
3
4
5
6
7
8
9
10
11
12
1314
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34

3.2 The design of the camera system

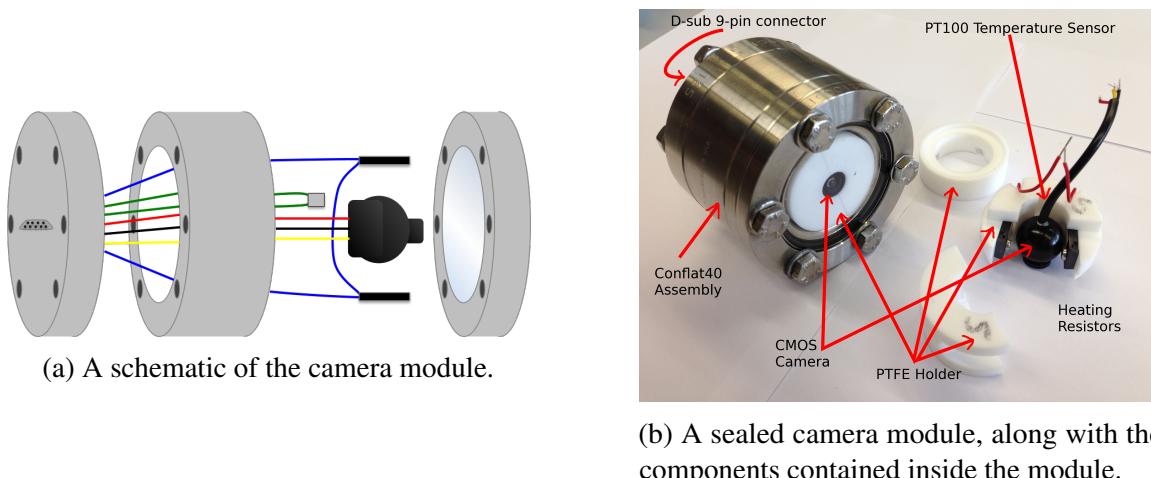


Fig. 3.1 The components which made up a camera module used in the 35 ton camera system. Left: a schematic representation of the components in the camera modules, with wires from the PT100 temperature sensor (green), the heating resistors (blue) and camera (red, blue, yellow), connected to a 9-pin D-sub 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.

¹ was also found that with the heater operational, the internal temperature of the module could
² rise by as much as 80 K after 17 minutes. During operation, the heaters and power supply
³ to the cameras were controlled using a custom-built, remotely controllable device, built by
⁴ Bob Bridgeland, at the University of Warwick. The temperature from the PT100 sensor was
⁵ processed by another custom-made device, the output of which was read into a computer
⁶ using a NI USB-6009 device.

⁷
⁸ A schematic of a camera module is shown in Figure 3.1a. Figure 3.1b shows a picture of
⁹ a sealed camera module, with the individual components inside a module next to it.

¹⁰
¹¹ It was necessary to mount the camera modules inside the 35 ton. This was achieved
¹² using a custom-designed mounting bracket, an example of which is shown in Figure 3.2. The
¹³ mounts were fixed to existing cryogenic pipework, and could be freely rotated, so that the
¹⁴ cameras could be pointed towards specific regions of interest. This meant that the camera
¹⁵ mounts had two degrees of freedom, as well as being able to freely placed on the cryogenic
¹⁶ pipes.

3.3 Performance in the 35 ton75



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.

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 ran 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.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15

3.3 Performance in the 35 ton

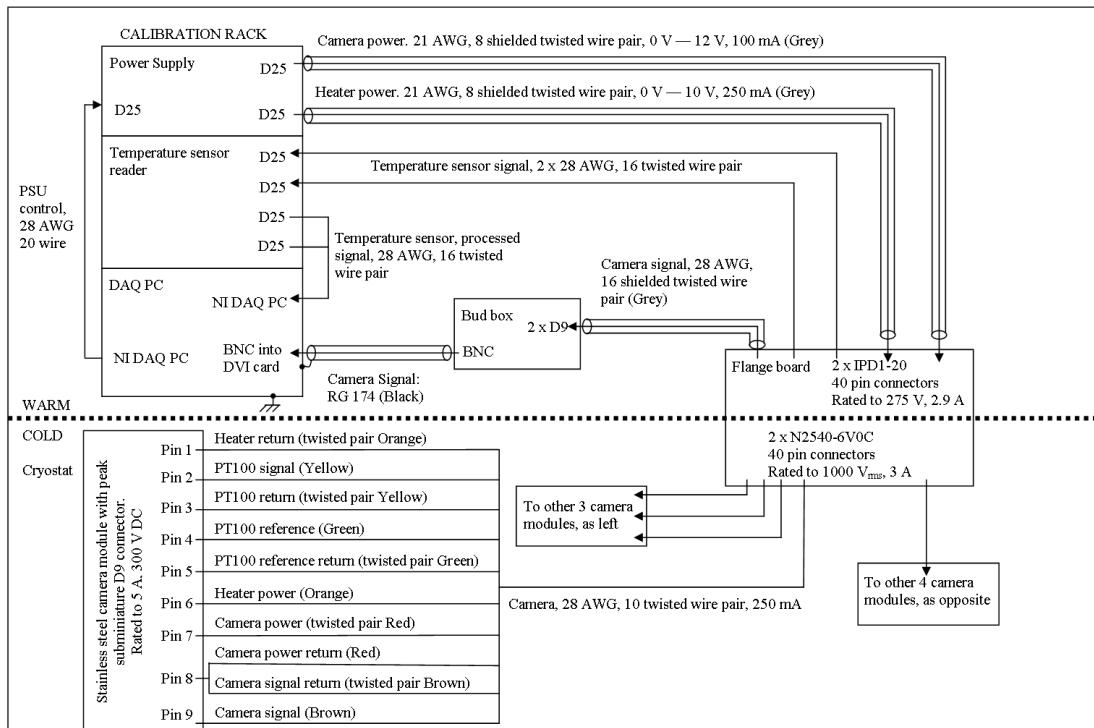


Fig. 3.3 The system diagram for the 35 ton camera system. The contents of the rack which was used in the camera system (top left), the wires which were terminated at the flange board (right), and the signals that each of the wires connected to the cameras carried (bottom left), are shown. In addition 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 feedthrough.
- 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 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

3.3 Performance in the 35 ton

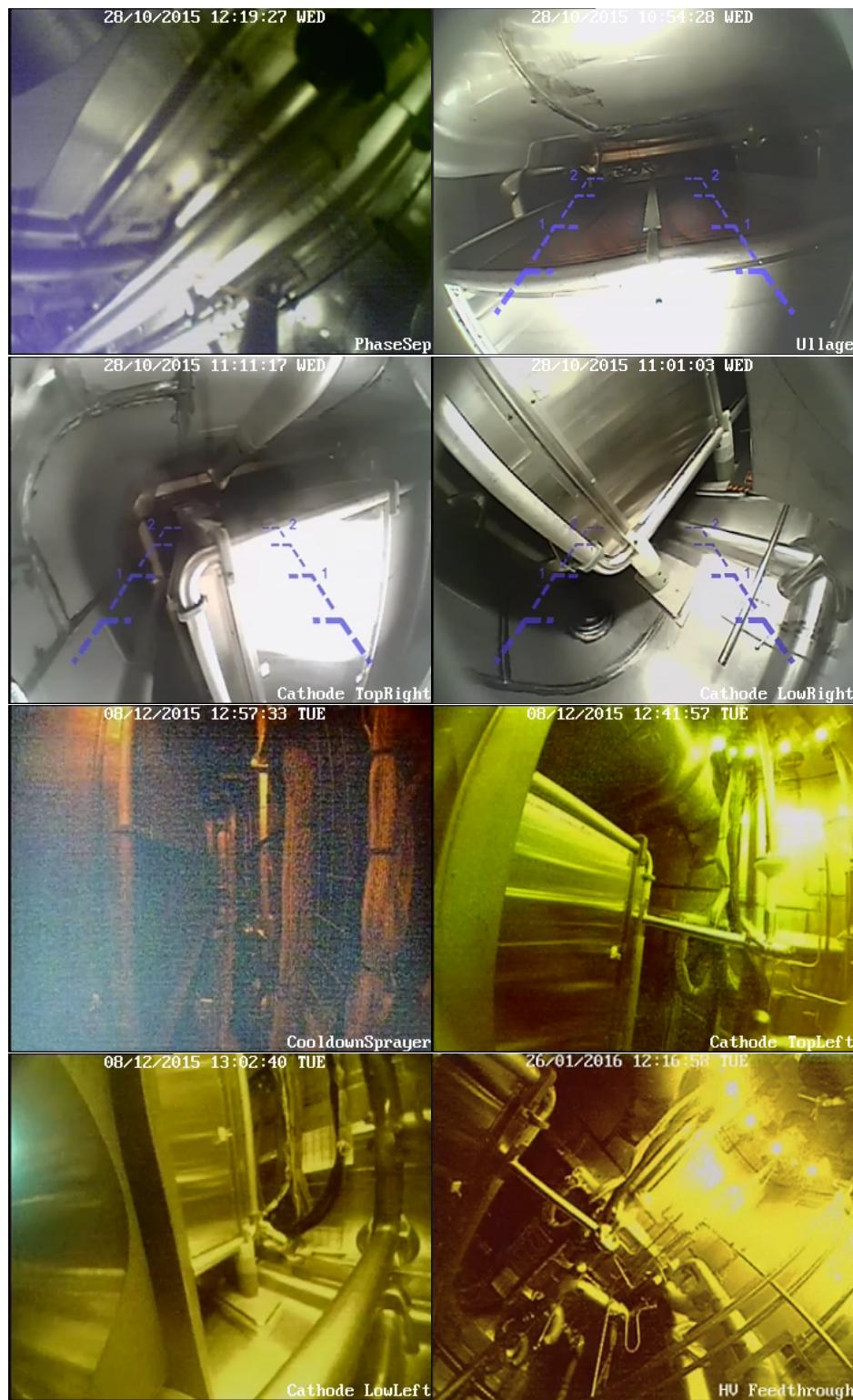
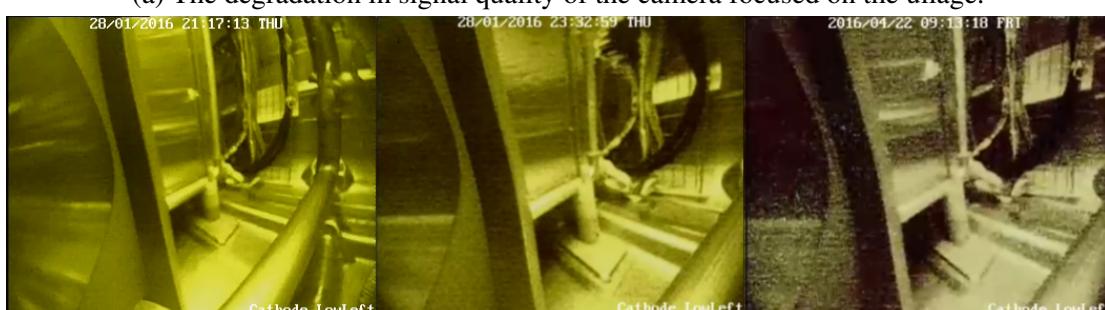


Fig. 3.4 The calibration images for the 8 cameras used in the 35 ton camera system. The images from left to right, top to bottom, show the following: phase separator, ullage, cathode top right, cathode bottom right, cool down sprayers, cathode top left, cathode bottom left, and high voltage feed through. 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.

3.3 Performance in the 35 ton

80

1 used to trigger on HV breakdowns was severely limited for some cameras. As shown, this
2 degradation was highly camera specific, with the resolution of the ullage camera deteriorating
3 severely, whilst the camera focusing on the lower left corner of the cathode is relatively
4 unchanged over the course of the run. A striking difference between the left-most images,
5 and the central and right-most images, in Figure 3.5, is the loss of colour in the images
6 which are produced. This is attributed to a partial failure of the on-board encoding circuits,
7 which resulted in the colour signal streams not being functional when exposed to cryogenic
8 temperatures. This loss of colour was seen as soon as the cameras were exposed to cryogenic
9 temperatures, and was also observed in the initial tests performed in Sheffield.

10

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

17

18 During normal operation of the 35 ton system, the high voltage operated stably at 60 kV,
19 and so no breakdowns were observed. However, after the cathode was raised to 135 kV in
20 low-purity argon, three of four breakdowns were detected by the camera system, and data
21 was written to disk as expected. Unfortunately, the cameras were not able to pinpoint the
22 exact locations of the breakdowns, and so their effectiveness at discerning the locations of
23 HV breakdowns is still largely untested. Despite this, however, the cameras proved to be a
24 very valuable monitoring tool in the 35 ton cryostat, and there has been interest in including
25 analogous systems in future LArTPCs, such as SBND [131].

26

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 the 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

4.1 Determination of interaction times

82

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

6

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

24

25 Another metric by which flashes could be assigned to reconstructed tracks, is by utilising
26 the relationship between the number of measured photoelectrons in the simulation, and the
27 distance from the APAs at which they were produced. When considering two flashes of
28 scintillation light that are produced at different distances from the APAs, it would be expected
29 that more photoelectrons would be collected when the photons were produced closer to the
30 APAs. This relationship is shown in Figure 4.2, where it can be seen that as drift distance
31 increases, the number of measured photoelectrons decays exponentially. This relationship
32 means that the distance between the APAs and the source of the flash, can be predicted from
33 the number of measured photoelectrons. The predicted separation of the flash source and the
34 APAs, can then be compared to the expected x position of a reconstructed track, given the
35 difference in flash time and hit times (Figure 4.3). The difference in these two quantities can
36 then be used as a second metric, as it gives a measure of how well matched the reconstructed

4.1 Determination of interaction times

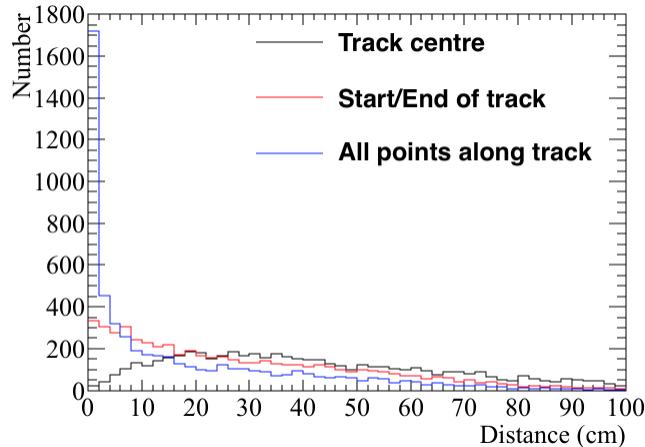


Fig. 4.1 The number of events as a function of the calculated distance between a reconstructed track, and a reconstructed flash, for various metrics. Black: the distance between the track centre and the flash centre. Red: the perpendicular distance between the flash centre, and the line joining the start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

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 metric has very accurately predicted the interaction time of the flash.

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 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 by the

4.1 Determination of interaction times

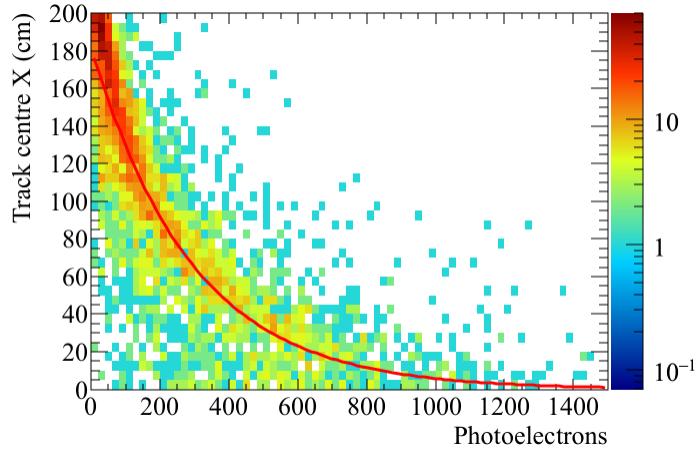


Fig. 4.2 The central x position of a reconstructed track versus the number of detected photoelectrons. The red line corresponds to a parameterisation of the distribution, which is used to predict the distance between the APAs and the flash source, given the number of measured photoelectrons in the flash.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

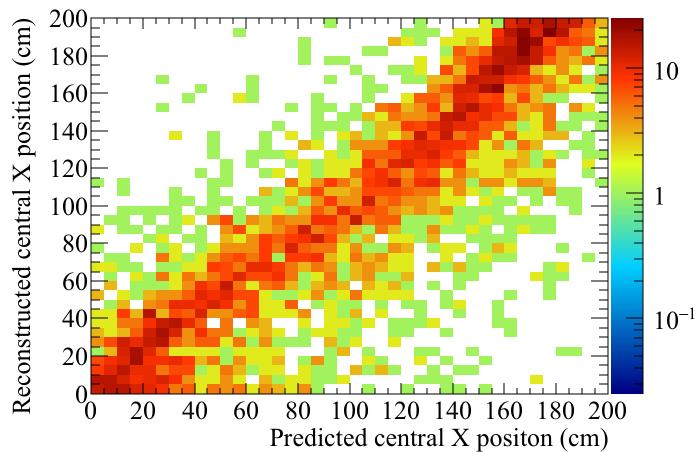
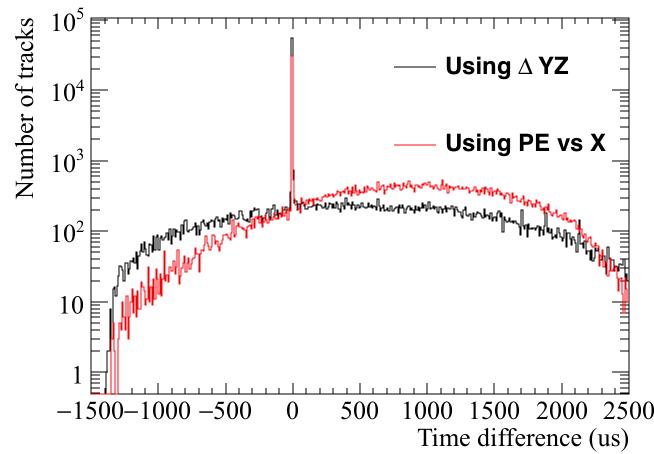
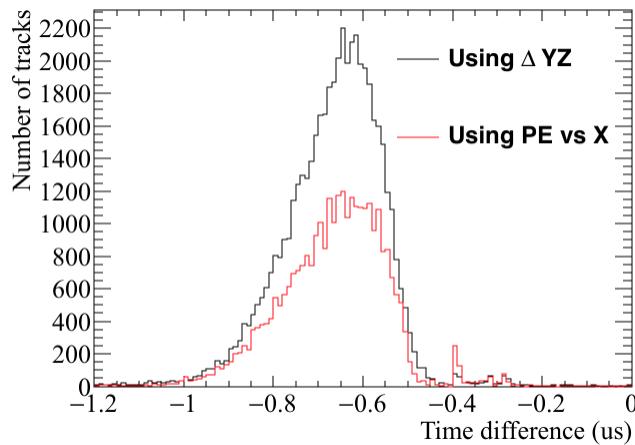


Fig. 4.3 A comparison of the x position predicted using the relationship in Fig 4.2, and the x position predicted using the difference in flash and hit times.

4.1 Determination of interaction times



(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 $0.6 \mu s$ due to an electronics offset in the simulation.

4.2 Calibrating calorimetric constants

86

1 photon detectors is often very accurate. However, there is also a significant number of tracks
 2 for which the photon interaction time does not match the Monte Carlo truth interaction time.
 3 This is due to, for example, the broad width of the distribution shown in Figure 4.2, meaning
 4 that the number of photoelectrons cannot necessarily accurately predict the distance between
 5 the APAs and the flash source. Improving the accuracy of the photon detector interaction time
 6 is required, as applying an incorrect interaction time correction will result in the x position
 7 of hits being incorrect. The peak seen at a time difference of 0 ms has been expanded in
 8 Figure 4.4b, where it can be seen that there is a systematic offset of 0.6 μ s, this is due to an
 9 electronics offset applied in the simulation to the photon detector system.

10

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

25

26 **4.2 Calibrating calorimetric constants**

27 Having the correct calorimetric responses is vital when trying to calculate $\frac{dE}{dx}$, as the mea-
 28 sured change in charge has to be correctly converted to the change in energy. The parameter
 29 which has to be tuned in order to ensure that this is done correctly, is the number of electrons
 30 that each ADC corresponds to. This was presented in Equations 2.1 and 2.2, as $C_{ADC \rightarrow e^-}$.
 31 Each plane will have a different response function, and so each plane has to be treated
 32 separately. These parameters have to be tuned in such a way as to make a known particle
 33 energy deposition have the correct $\frac{dE}{dx}$, the easiest deposition to tune against is the minimally
 34 ionising particle (MIP) peak, which in LAr should have a value of $1.8 \text{ MeV}\cdot\text{cm}^{-1}$. To do

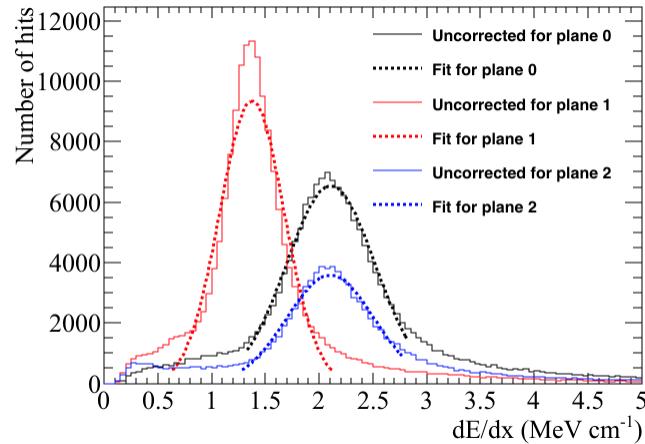
this, the sample of 10,000 positive muons made to calibrate the photon detector track/flash assignment will be used, as many of these particles will be MIPs.

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 Most Probable Value (MPV) of $\frac{dE}{dx}$ for that plane. If the MPVs are not equal to $1.8 \text{ MeV}\cdot\text{cm}^{-1}$, 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.

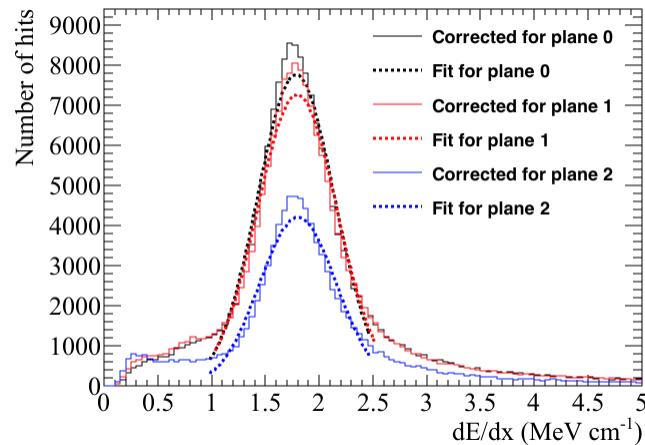
4.3 Discerning reconstruction efficiencies

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, and at an altitude above sea level of 0 m.

The criteria upon which to determine whether a particle is well reconstructed has to be carefully chosen, as every definition will have limitations. 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



(a) Before calibration is performed.



(b) After calibration is performed.

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

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 with regards to many aspects 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 ran on the particles. One reconstruction path uses no Monte Carlo information, and

4.3 Discerning reconstruction efficiencies

90

1 so the disambiguation is performed as outlined in Section 2.5, and the interaction time is
2 determined using the simulated photon detectors, as described in Section 4.1. The second
3 reconstruction path uses cheated disambiguation, whereby Monte Carlo truth information is
4 used to select the wire segments which hits occurred on, and Monte Carlo truth information
5 is also used when calculating the interaction time of the track.

6

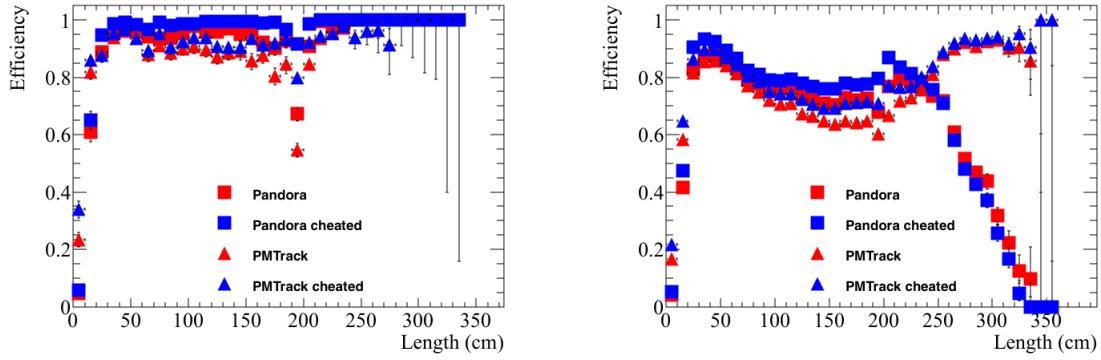
7 The calculation of reconstruction efficiencies also serves as an effective method upon
8 which reconstruction algorithms can be further developed, as it identifies aspects which do
9 not work as expected. For example, when the reconstruction efficiencies for the CRY sample
10 were initially calculated, they were significantly lower than for the positive muon sample (at
11 10%), but only when disambiguation was not cheated. It transpired that this was because the
12 disambiguation was only selecting the largest collection of hits on each plane, for each TPC.
13 This is not a problem when only 1 particle is simulated, as one would expect that there only
14 be one particle on a given plane, in each TPC. However, when considering a CRY sample
15 which lasts of the order of 10 ms, there will almost certainly be multiple particles on each
16 plane, in each TPC. Removing the hits from all but one of these multiple particles, will
17 cause them to have no reconstructed track, and thus cause the efficiency to drop significantly.
18 Upon making the disambiguation algorithm no longer have this restriction, the reconstruction
19 efficiencies of the positive muon and CRY samples were observed to become much more
20 similar, as presented here.

21

22 The reconstruction efficiencies given the current state of the most commonly used recon-
23 struction algorithms, Pandora [118] and PMTrack [84], are shown in Figures 4.6, 4.7, 4.8, 4.9
24 and 4.10. Efficiencies are shown for both the positive muon and CRY samples, where it can
25 be seen that the efficiency tends to be lower for the CRY sample. It is thought that this is due
26 to the more complex event structure in the CRY sample, as multiple primary particles will be
27 present in the detector at any given time. The relatively slow drift velocity of LAr may mean
28 that these tracks cross in wire-tick space. Tracks crossing in wire-tick space could cause
29 reconstruction errors, as the overlaps may be mistaken for interactions, which would cause
30 the tracks to be split, resulting in the interaction time calculated from the photon detectors to
31 be incorrect. This error, in the calculation of interaction time using the photon detectors, was
32 seen in Figure 4.4. The reconstruction efficiencies for the CRY sample are more realistic, as
33 events will rarely be isolated in the 35 ton detector, due to the large flux of cosmic particles
34 on the Earth’s surface.

35

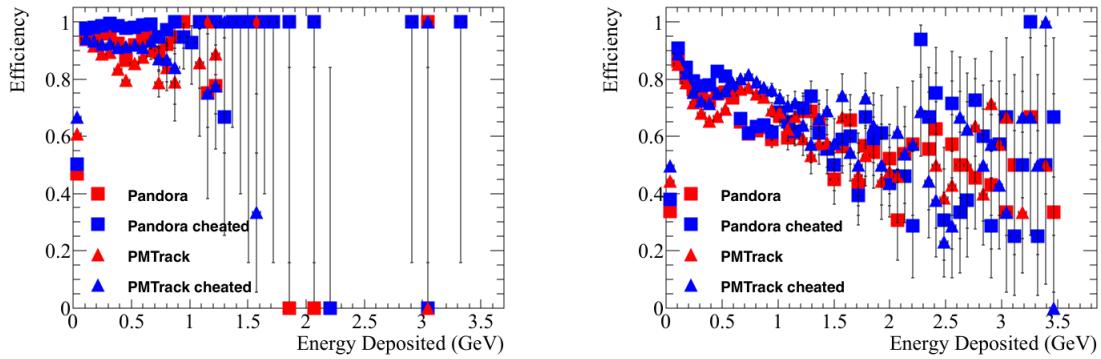
4.3 Discerning reconstruction efficiencies



(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 [118] (squares) and PMTrack [84] (triangles).



(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 [118] (squares) and PMTrack [84] (triangles).

4.3 Discerning reconstruction efficiencies

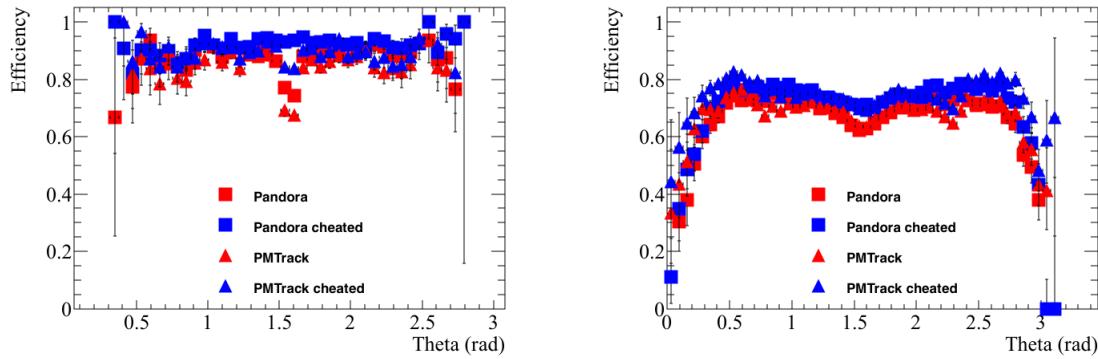


Fig. 4.8 The reconstruction efficiencies for simulated events as a function of the θ track angle from Monte Carlo truth track. The efficiencies are shown for 'non-cheated' reconstruction (red), and 'cheated' reconstruction (blue), for both Pandora [118] (squares) and PMTrack [84] (triangles).

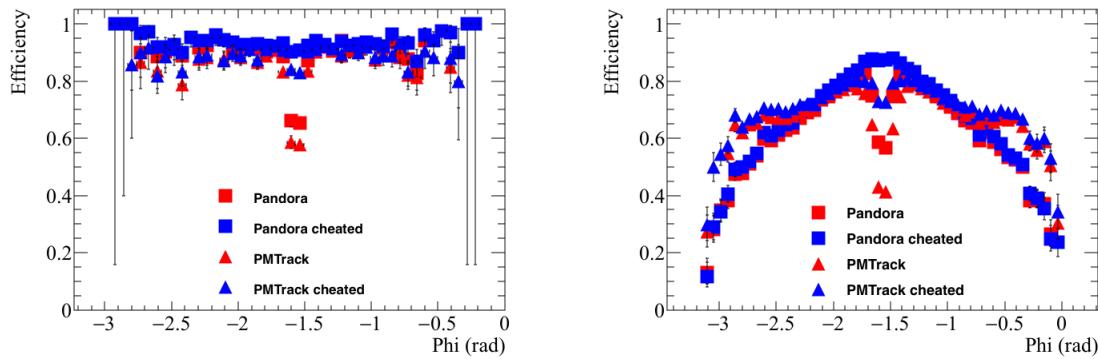
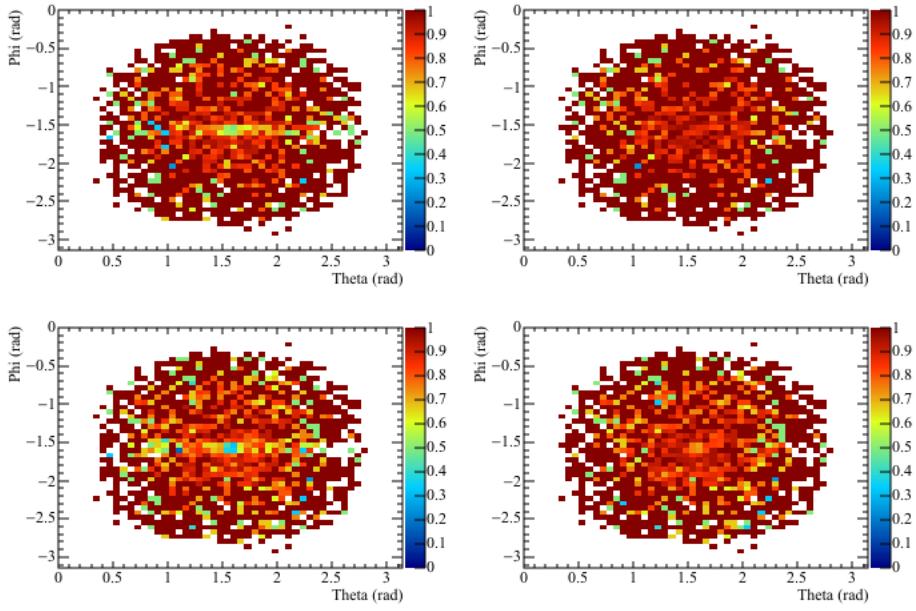
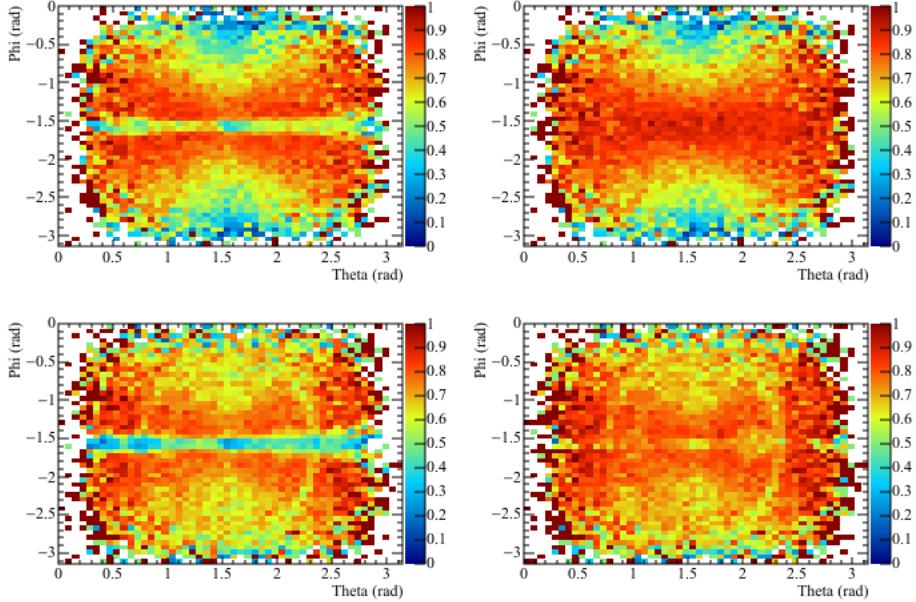


Fig. 4.9 The reconstruction efficiencies for simulated events as a function of the ϕ track angle from Monte Carlo truth track. The efficiencies are shown for 'non-cheated' reconstruction (red), and 'cheated' reconstruction (blue), for both Pandora [118] (squares) and PMTrack [84] (triangles).



(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 [118] (top) and PMTrack [84] (bottom).

4.3 Discerning reconstruction efficiencies

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

$$7 \quad x_{Hit} = T_{Hit} \times v_{Drift} \quad (4.1)$$

8 In Equation 4.1, x_{Hit} is the calculated x position of the hit, T_{Hit} is the measured time of the
 9 hit, and v_{Drift} is the electron drift velocity. An electron, in an electric field of $500 \text{ V}\cdot\text{cm}^{-1}$,
 10 in LAr, drifts at a speed of $0.160563 \text{ cm}\cdot\mu\text{s}^{-1}$. However, when the same is done for hits
 11 in the CRY sample, using particles with large interaction times, the x positions will have
 12 offsets proportional to the interaction time of the particle, unless the hit time is corrected by
 13 Equation 4.2.

$$14 \quad T_{Hit} = T_{Measured} - T_{Interaction} \quad (4.2)$$

15 In Equation 4.2, T_{Hit} is the corrected hit time, $T_{Measured}$ is the measured time of the hit, and
 16 $T_{Interaction}$ is the calculated interaction of the particle which caused the hit.

17
 18 The scale of these offsets in x positions can be seen by considering a particle which is
 19 generated at a time of $T = 12.5 \text{ ms}$, as then the offset in the reconstructed x position calculated
 20 by Equation 4.1, would be more than 20 m! Obviously the hits could not have occurred
 21 at these positions, as the drift distances are roughly 30 cm in the “short” drift volume, and
 22 225 cm in the “long” drift volume. However, as track segments are reconstructed separately
 23 in the “short” and “long” drift volumes, before later being merged, it is possible for there
 24 to be a discontinuity in x of more than 40 m, if these x offsets are not corrected for. As the
 25 interaction time of the track is calculated using the output of the tracking algorithms, it is not
 26 possible to prevent this by using the interaction time at present. It is however, possible to
 27 subtract this discontinuity in x from the total reconstructed track length when the stitched
 28 track is stored in the event. This will give the correct track length, though the user will still
 29 have to correct the positions of individual hits, using the calculated interaction times. This
 30 is what is done by PMTrack, hence it not exhibiting this rapid decrease in reconstruction
 31 efficiency for long tracks. The interaction time can be found from, among other things, the
 32 Monte Carlo truth generation time, or the photon detectors, as discussed in Section 4.1.

33

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 partially attributed to particles with track lengths of less than 1 cm in the active volume of the detector. These particles, which represent 30% of the particles with active volume track lengths of less than 10 cm, are too short to be reconstructed using the current reconstruction process. These particles 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

4.4 Performing particle identification

¹ large electromagnetic showers upon entering the LAr.

²

³ It is also interesting to note the pronounced decreases in reconstruction efficiencies for
⁴ particular angles, shown in Figures 4.8 and 4.9. The decrease in efficiency at $\phi = \frac{\pi}{2}$, seen in
⁵ Figure 4.9, can be attributed to the drop in efficiency for particles with track lengths between
⁶ 190 cm and 200 cm, seen in Figure 4.6. This is because the vertical height of the detector
⁷ is approximately 195 cm, and near vertical tracks will hit few collection wires, meaning
⁸ that determining the triple points needed by the disambiguation is very difficult. This is
⁹ verified by the large increase in efficiency achieved by cheating the disambiguation, as seen
¹⁰ in Figure 4.8a, where the reduction in reconstruction efficiency is seen to become much
¹¹ less pronounced. Similarly, the decrease in efficiency at $\theta = \frac{\pi}{2}$ can be attributed to particles
¹² which are perpendicular to the collection plane wires, as this track orientation also results in
¹³ few collection wires being hit.

¹⁴

¹⁵ The information from Figures 4.8 and 4.9 is combined in Figure 4.10, where the sharp
¹⁶ drop in efficiency at $\phi = \frac{\pi}{2}$ for the “non-cheated” CRY sample is particularly visible. The
¹⁷ effect of cheated disambiguation is clear in Figure 4.10b, where the dip in efficiency as a
¹⁸ function of θ at fixed $\phi = \frac{\pi}{2}$ is completely removed. The same is not true however, for the
¹⁹ dip in efficiency as a function ϕ at fixed $\theta = \frac{\pi}{2}$, though the reduction in efficiency was not
²⁰ as severe as that seen for fixed values of $\phi = \frac{\pi}{2}$. The effect of “cheated disambiguation”
²¹ can still be seen though, as the reconstruction efficiency in Figure 4.10b can be seen to
²² improve for values of ϕ . There are still however, noticeable decreases in the reconstruction
²³ efficiency for values of ϕ close to 0 or π , when using Pandora. The improvement in the
²⁴ performance of the reconstruction algorithms that comes from “cheating” the reconstruction,
²⁵ is part of the motivation for the wire pitches in the DUNE FD being 36° , as opposed to the
²⁶ $45 \pm 0.7^\circ$ used in the 35 ton. This is because, as discussed in Section 2.5, the shallower wire
²⁷ pitch makes disambiguation easier. Though disambiguation will be easier in the different
²⁸ geometry, further efforts to improve disambiguation are still required, as are continued efforts
²⁹ to reconstruct the shortest tracks.

³⁰

31 4.4 Performing particle identification

³² Being able to perform reliable Particle IDentification (PID) is a key requirement for the
³³ DUNE experiment, and so efforts have been made to establish a procedure by which this can
³⁴ be achieved. The predominant method of performing PID in LAr is to use the relationship

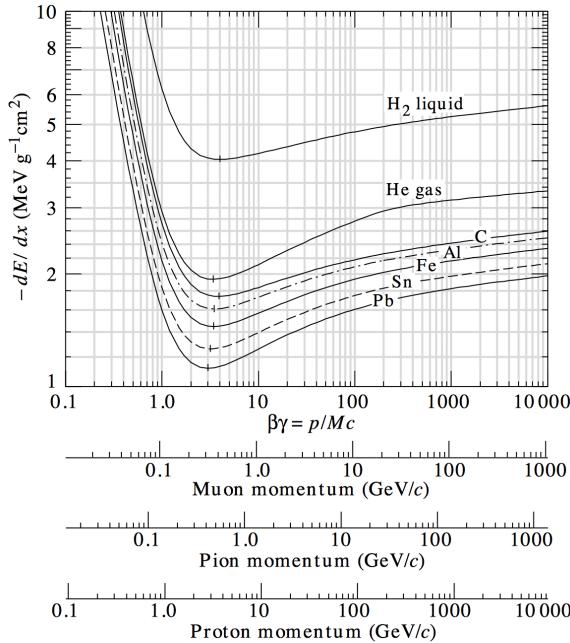


Fig. 4.11 The mean energy loss per unit track length of different particle masses in different materials [26]. The $\frac{Z}{A}$ of liquid argon is slightly less than that of carbon.

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. This relationship is observed to be dependent on particle mass, and is quantified by the Bethe-Bloch equation [132, 133] which is shown in Figure 4.11, and presented in Equation 4.3.

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

where $K = 4\pi N_A r_e^2 m_e c^2$, N_A is Avogadro's number, r_e is the 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 kinetic energy, I is the mean excitation energy, and $\delta(\beta\gamma)$ is the density effect correction to ionisation energy loss.

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 momenta at which the increase in $\frac{dE}{dx}$ occurs is very different for muons and protons. However, it can be seen the momenta 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

4.4 Performing particle identification

Table 4.1 Stopping power parameterisation for various particle types in liquid argon. The table is taken from [122].

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

1 from a muon track, though it may be difficult to discern a muon track from a pion track.

2

3 The particle mass dependence can be seen by plotting the $\frac{dE}{dx}$ against the residual range
4 of the particle on a log-log plot, as shown in Figure 4.12. A power law dependence is found
5 to describe the relationship [122], as shown in Equation 4.4.

$$6 \quad \frac{dE}{dx} = AR^b \quad (4.4)$$

7 The dependence on b is found to be weak, and so can be set to -0.42 for all particle masses.
8 This means that the main discriminant used is the A parameter, which has a strong depen-
9 dence on particle mass. Table 4.1 shows the values of A and b which are calculated from
10 Figure 4.12. It is found that the error introduced by fixing the b parameter is small compared
11 to the error from ionisation fluctuations [122].

12

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

$$16 \quad A_i = \left(\frac{dE}{dx} \right)_i \times (R_i)^{0.42} \quad (4.5)$$

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

23



Fig. 4.12 Stopping power for different particle masses as a function of residual range in liquid argon. The figure is taken from [122].

The PIDA method was tested in [122], 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 cleanly separated in the truth study. The author notes of [122] that an incorrect tuning of the electron recombination effects will cause the distributions in Figure 4.13 to become broader. The dependence of $\frac{dE}{dx}$ on the recombination effects ($Recomb_{A/B}$) were presented in Equations 2.1 and 2.2. The author of [122] also notes that an incorrect calibration of the detector will introduce a systematic shift in the expected values of PIDA, 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 particle types in cosmic 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

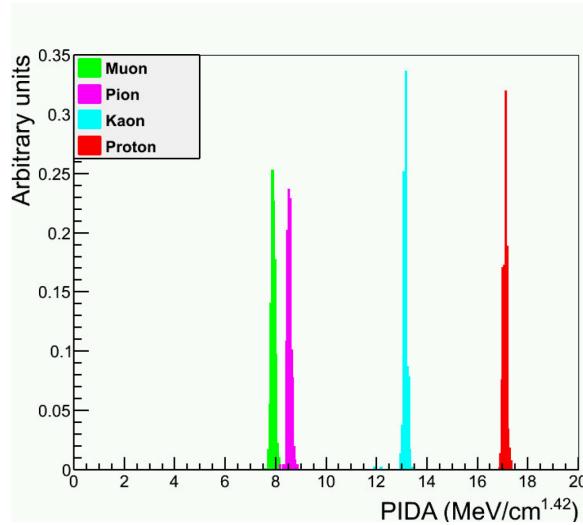


Fig. 4.13 The distribution of PIDA values, calculated using Monte Carlo truth, for different particle masses. The figure is taken from [122].

¹ of protons in the CRY sample (Figure 4.14).

²

³ Figure 4.14 shows that the average reconstruction efficiency for PMTrack is higher than
⁴ that for Pandora when considering protons. This can be easily seen in Figure 4.14c, where
⁵ the efficiency for PMTrack is roughly 10% higher than that of Pandora, for all values of
⁶ θ . The reconstruction efficiency is still much lower than the overall efficiency shown in
⁷ Figure 4.8, for both the positive muon and CRY samples though. This shows that the overall
⁸ reconstruction efficiency for protons is quite low. Comparing Figures 4.6b, and 4.14a, it is
⁹ evident that the reconstruction efficiency for protons with track lengths of more than 10 cm is
¹⁰ reasonably similar to that of the overall reconstruction efficiency for the CRY sample, when
¹¹ using PMTrack. However, the reconstruction efficiency is significantly lower for protons
¹² with tracks of less than 10 cm. When using Pandora to reconstruct protons, the reconstruction
¹³ efficiency is lower for all track lengths. It is found that 60% of simulated protons have
¹⁴ track lengths of less than 1 cm, and that none of these particles are reconstructed. It is this
¹⁵ large number of very short particles which causes the overall reconstruction efficiency to be
¹⁶ relatively low. When particles with track lengths of less than 1 cm (10 cm) are removed, the
¹⁷ average reconstruction efficiency for PMTrack rises to 37% (58%). This shows that, when
¹⁸ the shortest tracks are not counted, the reconstruction performs reasonably well.

¹⁹

²⁰ It is also useful to produce samples where the primary particle is a single muon, or proton,
²¹ located in the active volume of the detector. This allows for a sample of isolated tracks to
²² be made, upon which the capabilities of the PIDA metric can be tested. It also allows the

4.4 Performing particle identification

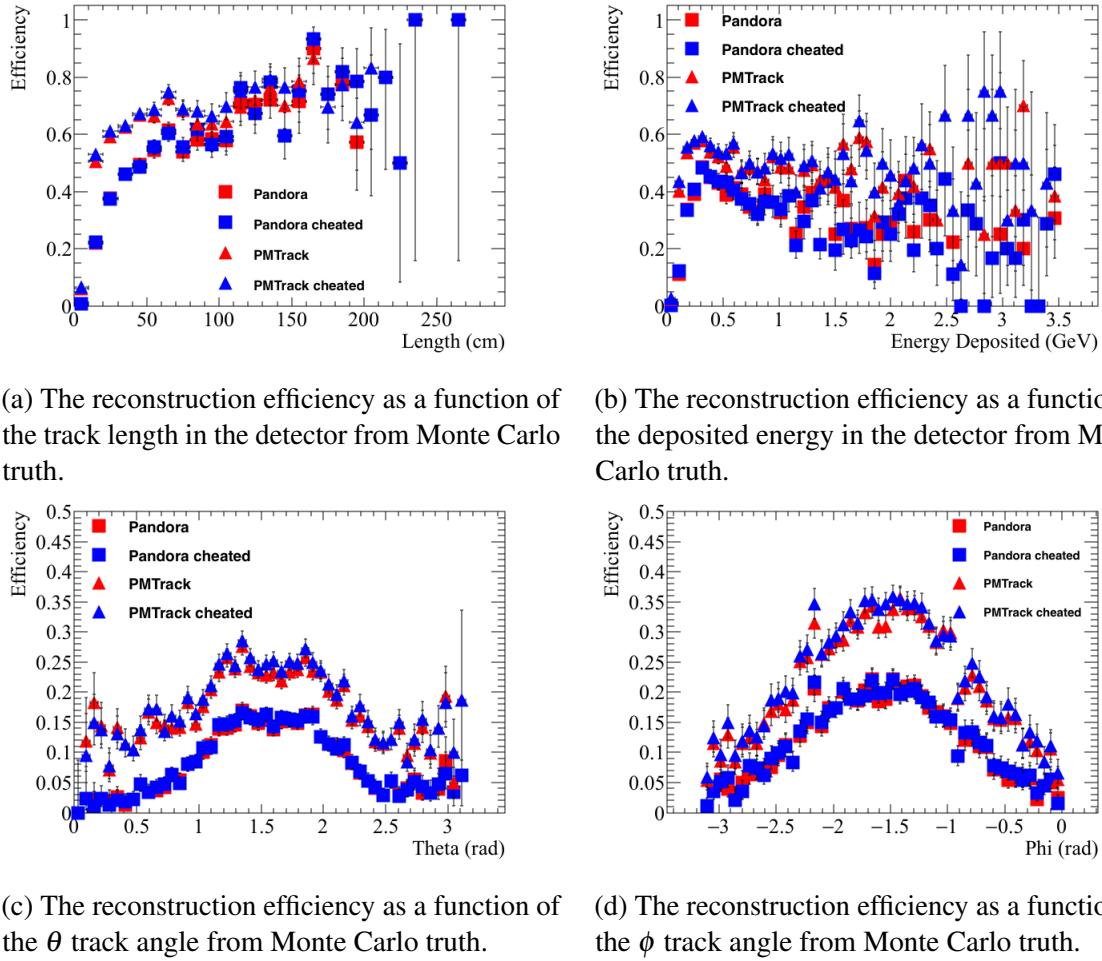


Fig. 4.14 The reconstruction efficiencies for protons in a sample generated using CRY. Top left: the efficiencies as a function of the track length in the detector from Monte Carlo truth. Top right: the efficiencies as a function of the deposited energy from Monte Carlo truth. Bottom left: the efficiencies 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 [118] (squares) and PMTrack [84] (triangles).

4.4 Performing particle identification

102

Table 4.2 The properties of initial particles simulated in the muon and proton samples. The angles θ_{xz} , and θ_{yz} , are defined as the angle that a vector makes in the xz , and yz planes, respectively.

	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

1 reconstruction efficiency to be found for particles in isolation. The properties of the generated
 2 particles are illustrated in Table 4.2. The values of the simulated quantities were found by
 3 changing the given parameters by an amount taken from a random sampling of a Gaussian
 4 distribution of width equal to the error listed. These simulation parameters were chosen to
 5 produce samples which would contain both exiting, and stopping particles, whilst generating
 6 the particles in the LAr would ensure that there should always be a reconstructable track in
 7 the detector. The reconstruction efficiencies when using the PMTrack reconstruction method
 8 are shown for the simulated particles in Figure 4.15.

9

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

23

24 As the increase in $\frac{dE}{dx}$ is only visible when the particle stops in the detector, it is necessary
 25 to remove exiting particles from the sample. This is done by applying a fiducial cut on the
 26 end point of the reconstructed track. It is important to only place this on the end point of the
 27 track, as one does not want to remove particles which enter the detector and then stop. When

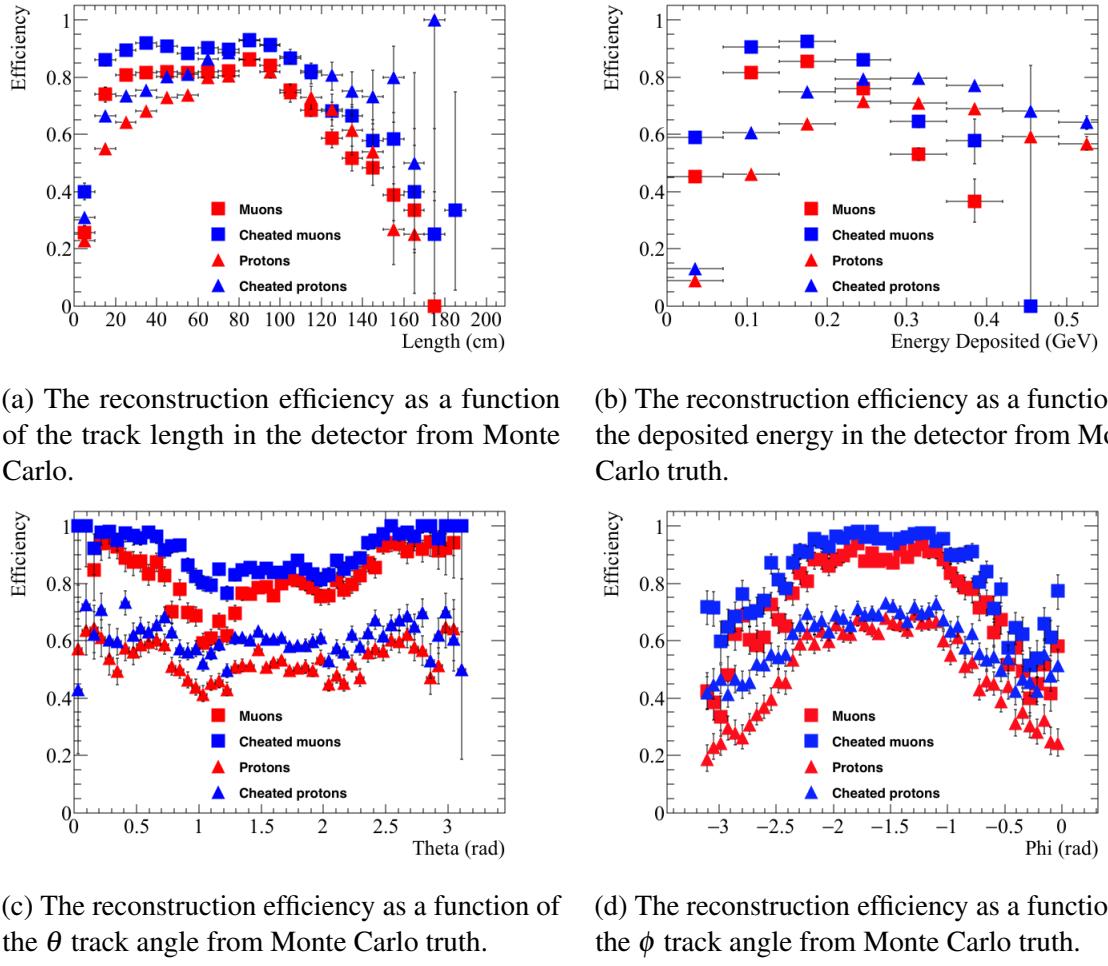


Fig. 4.15 The reconstruction efficiencies for the simulated isolated muon and proton samples in the 35 ton detector. Top left: the efficiencies as a function of the track length in the detector from Monte Carlo truth. Top right: the efficiencies as a function of the deposited energy from Monte Carlo 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 [118] (squares) and PMTrack [84] (triangles).

4.4 Performing particle identification

104

1 calorimetry is performed, the end point of the track is determined using, among other metrics,
 2 the increase in $\frac{dE}{dx}$, and so the residual range of the track (a stored data member of the track
 3 object), and so should always refer to the distance to the end of the particles trajectory. For
 4 this study, a fiducial cut of 5 cm is used. This means that any track with hits within 5 cm of
 5 the edge of the detector volume is discarded, and counted as an exiting particle. This should
 6 mean that very few tracks due to exiting particles are identified as stopping in the detector,
 7 as it would require the reconstruction algorithms to miss a large section of the track. This
 8 will mean that some stopping particles are incorrectly assigned as exiting particles, causing
 9 the identification efficiency to drop, but it is necessary to ensure that exiting particles are
 10 not included in the final distributions. A further cut that is applied, is the requirement that
 11 the track contains a minimum of 10 collection plane hits, this is to ensure that an adequate
 12 number of points are taken upon which to find an average value of PIDA for the track. Similar
 13 cuts are described in [122], and the resulting distributions of PIDA values for the isolated
 14 muon, and isolated proton samples, are shown in Figure 4.16.

15

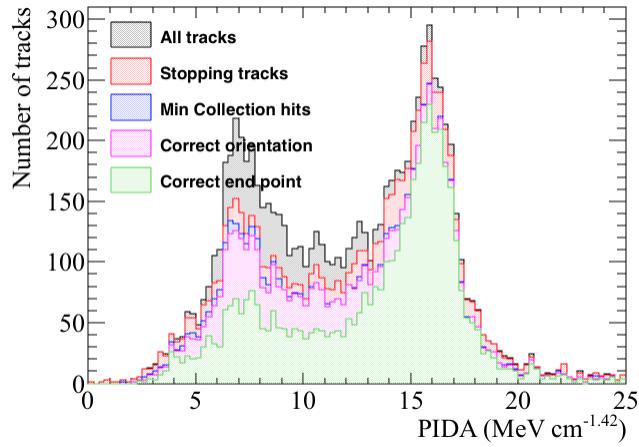
16 As can be seen from Figure 4.16, using Monte Carlo truth information can make the
 17 distributions much cleaner, particularly when discounting particles for which the reconstruc-
 18 tion algorithms do not track to their end point. A track is identified as having a correctly
 19 reconstructed end point, if the reconstructed end point is within 2.5 cm of the end point of
 20 the particle from Monte Carlo truth. It is reassuring to see that few tracks are reconstructed
 21 backwards, as if this were not the case then performing particle identification would be very
 22 difficult. This is because, it would indicate that the calorimetry and tracking algorithms are
 23 not performing well. However, improvements can still be made, as both plots in Figure 4.16
 24 contain many tracks which do not extend to the end points of the particles from Monte Carlo
 25 truth. This can be seen as when the requirement that the reconstructed track end point is
 26 consistent with the end point from Monte Carlo truth, the low tails of the PIDA distributions
 27 are significantly reduced. This is most noticeably the case in Figure 4.16b, where the peak at
 28 low values of PIDA is significantly reduced. It is observed that the PIDA distributions are
 29 cleaner when information from all three wire planes are used, as opposed to only using the
 30 collection plane, and so this is presented here. This shows how important it is to calibrate the
 31 electronics responses of all three wire planes, and how additional wire planes can improve
 32 calorimetry, as well as the accuracy of reconstruction algorithms [134].

33

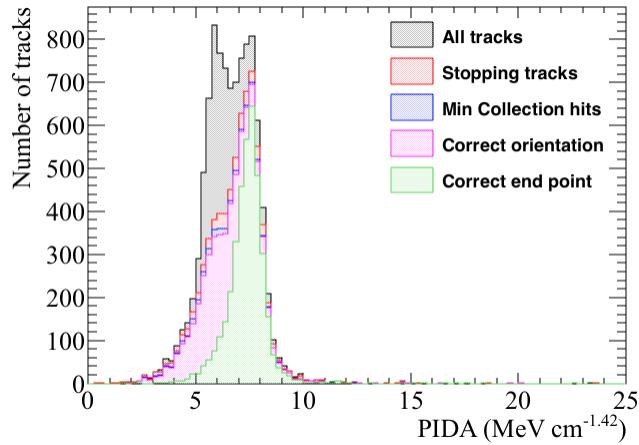
34 Figure 4.17 shows the relationship between the $\frac{dE}{dx}$ and residual range of a track, for both
 35 protons and muons. The much steeper increase in $\frac{dE}{dx}$ at low residual ranges for protons,
 36 compared to muons, is clearly visible when comparing Figures 4.17a and 4.17b. The contam-

4.4 Performing particle identification

105



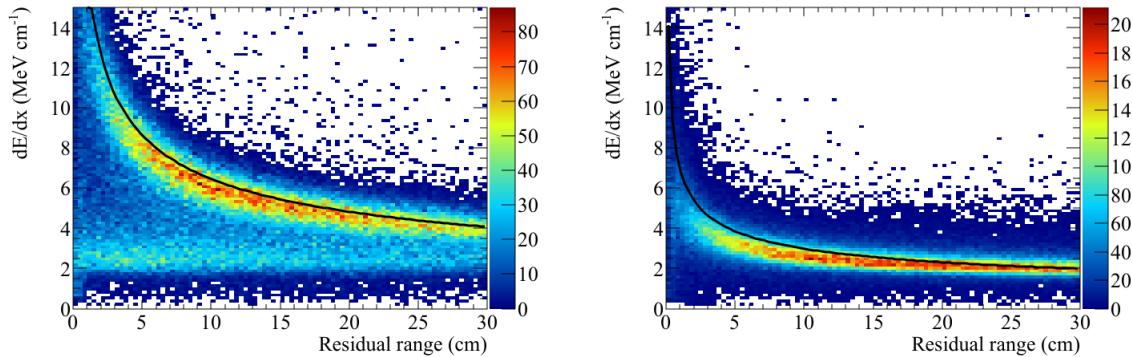
(a) The PIDA values calculated for the isolated proton sample.



(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.

4.4 Performing particle identification



(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. 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.

ination 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 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. 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 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

4.4 Performing particle identification

107

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%

are not due to stopping particles, the initial purity is low, though this increases markedly after the fiducial cut is applied. The PIDA ranges referred to are $14\text{-}18 \text{ MeV}\cdot\text{cm}^{-1}$,⁴², and $5\text{-}9 \text{ MeV}\cdot\text{cm}^{-1}$,⁴², for protons and muons respectively, as these ranges cover the peaks of the distributions shown in Figure 4.17, and are centred on the peaks in Figure 4.13.

As can be seen in Table 4.3, the efficiency upon 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 secondary protons, as the efficiencies shown in Table 4.3 are lower than those in Table 4.5. It

4.4 Performing particle identification

108

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%

¹ is thought that this is due to the low reconstruction efficiency for particles with the shortest
² track lengths, as many of the secondary protons will have short track lengths in the detector
³ from Monte Carlo truth, as discussed in Section 4.3. A similar table is not produced for
⁴ primary muons, as there were no secondary muons produced in the isolated muon sample,
⁵ and so Table 4.4 is itself the efficiency with which the primary muons can be identified.

⁶

⁷ Upon verifying that the PIDA metric can reliably determine particle type when they are
⁸ simulated in isolation, the next step is to observe the accuracy with which particles can be
⁹ identified in a CRY sample. The sample used here differs from the CRY sample used earlier,
¹⁰ in that only events which contain a proton track in the detector from Monte Carlo truth are
¹¹ reconstructed. This is done to reduce simulation time and storage space, as this cut will still
¹² provide a substantial number of muons, whilst ensuring that a large proton sample can be
¹³ reconstructed. This sample is thus called the “proton enriched CRY sample.”

¹⁴

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, 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 incorrectly determining the interaction time, and performing incorrect 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.

From Figure 4.18, it can be seen that the proton tracks which are reconstructed in the proton enriched sample, look very similar to the proton tracks which were 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 $8 \text{ MeV}\cdot\text{cm}^{-1}$ after the application of all cuts, this cannot be seen before the last cut is applied. This is because of two large peaks with PIDA values of around $3 \text{ MeV}\cdot\text{cm}^{-1}$ and

4.4 Performing particle identification

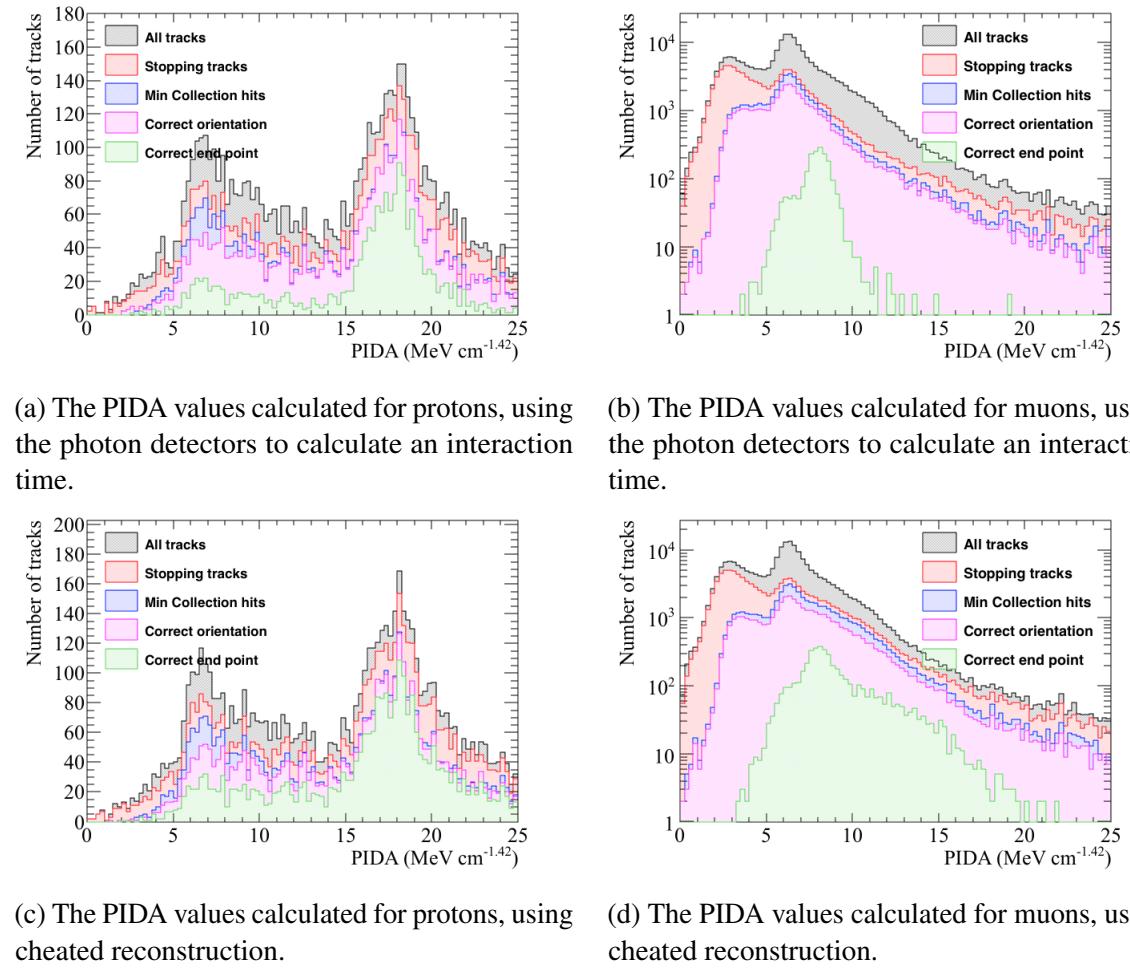


Fig. 4.18 The calculated PIDA values for the simulated proton enriched sample in the 35 ton detector. Top left: the PIDA values calculated for protons using the photon detectors to measure the interaction time. Top right: the PIDA values calculated for muons using the photon detectors to measure the 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.

6 MeV·cm⁻¹. A large peak at a PIDA value of 6 MeV·cm⁻¹ 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 MeV·cm⁻¹, 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 MeV·cm⁻¹. 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 associated with particles which have track lengths in the detector from Monte Carlo truth, which are much longer than themselves. Placing a cut at a minimum 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 Figures 4.19a, and 4.19b, that though some of the proton tracks which have PIDA values within the expected range will also be removed. It may therefore be necessary to develop a more sophisticated cut to remove these delta rays, should the identification of protons be 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 cosmic ray data collected by the 35 ton detector, is presented here though.

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

4.4 Performing particle identification

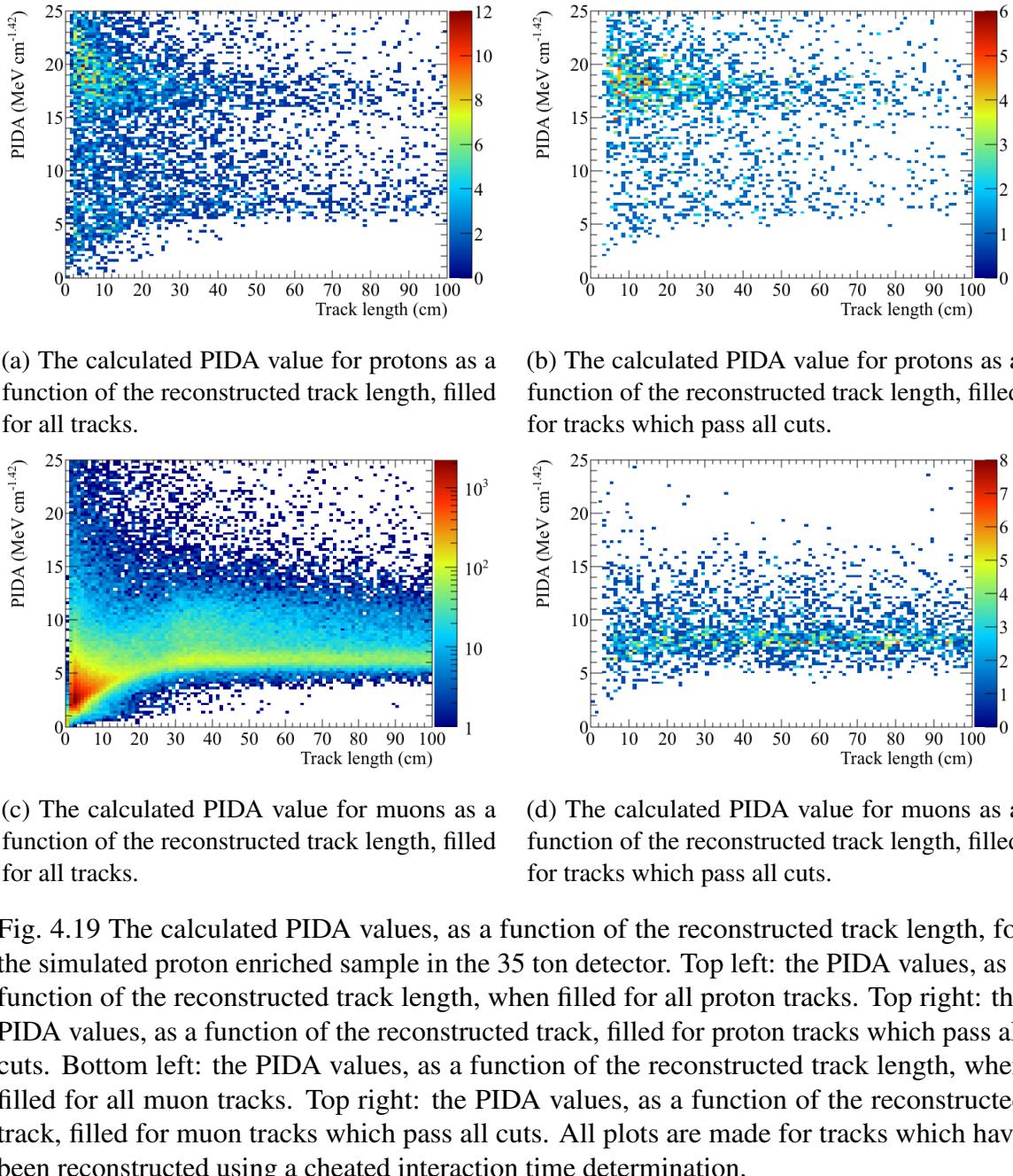
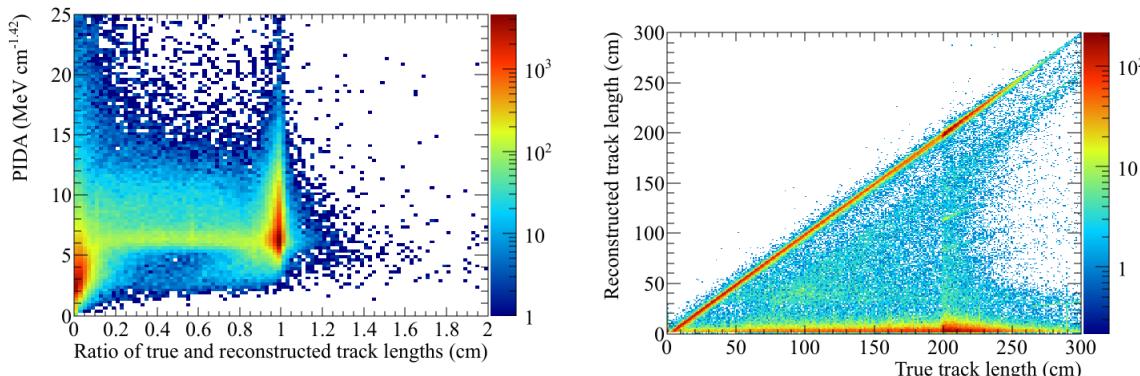


Fig. 4.19 The calculated PIDA values, as a function of the reconstructed track length, for the simulated proton enriched sample in the 35 ton detector. Top left: the PIDA values, as a function of the reconstructed track length, when filled for all proton tracks. Top right: the PIDA values, as a function 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.

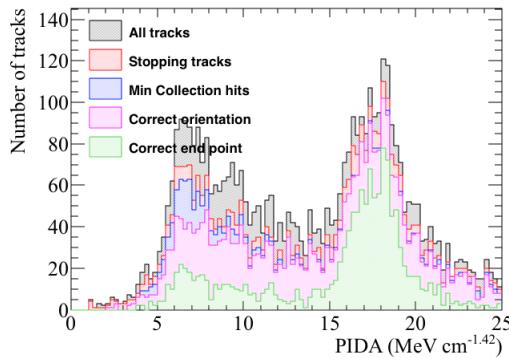
(b) The relationship between reconstructed track length, and the track length from Monte Carlo truth, filled for all tracks.

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.

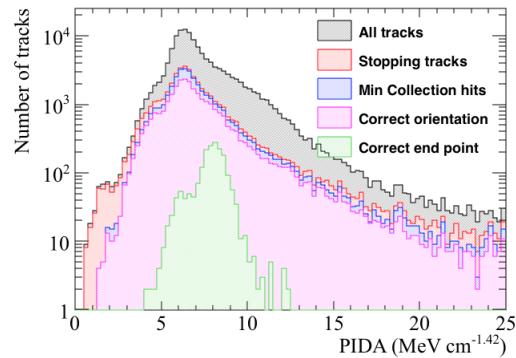
10 cm is applied.

When comparing Figures 4.18, and 4.21, the removal of muon tracks with PIDA values of $3 \text{ MeV}\cdot\text{cm}^{-1}$ 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 $8 \text{ MeV}\cdot\text{cm}^{-1}$ 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 reliably identified as such.

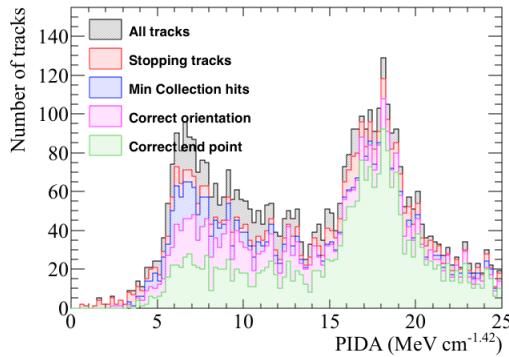
As outlined above, partially reconstructed muon tracks have PIDA values of $6 \text{ MeV}\cdot\text{cm}^{-1}$, and so in order to remove the large peak of these tracks still seen in Figure 4.21, some form of track stitching is required in order to remove these tracks. It is possible that a large number



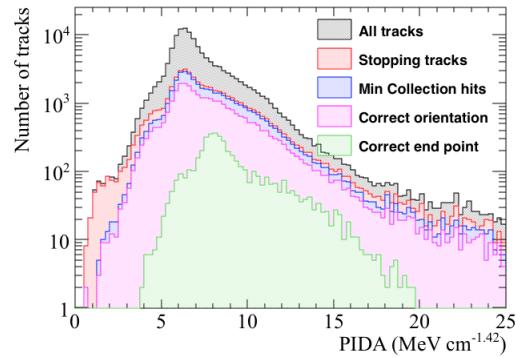
(a) The PIDA values calculated for protons, using the photon detectors to calculate an interaction time.



(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.

4.4 Performing particle identification

115

of these tracks are due to muons which pass through the detector, as many cosmic muons
1 will do, but have been reconstructed as at least two separate tracks. Should this be the case, at
2 least one of the tracks would be considered to “stop” in the detector. The effect of removing
3 these split tracks can be seen in Figure 4.22, where all particle which do not stop in the
4 detector, from Monte Carlo truth, have been removed. This effectively “cheats” the stitching
5 which is required to remove the majority of split muon tracks, though any muons which stop
6 in the detector, but have been split into two or more tracks by the reconstruction, will still be
7 present in Figure 4.22. It can be seen that these tracks are still present, as there is still a peak
8 in PIDA values of $6 \text{ MeV}\cdot\text{cm}^{-1}$, though these tracks are now overshadowed by tracks with
9 PIDA values of $8 \text{ MeV}\cdot\text{cm}^{-1}$.
10

The number of muon tracks with high PIDA values in all plots is concerning, though this
11 is particularly true for those shown in Figure 4.22, as these tracks are due to particles which
12 stop in the detector, and are accurately reconstructed. This is because the contamination
13 caused by these tracks, would mean that it would be difficult to ascertain whether a track
14 with a PIDA value of around 18 is actually a proton.
15

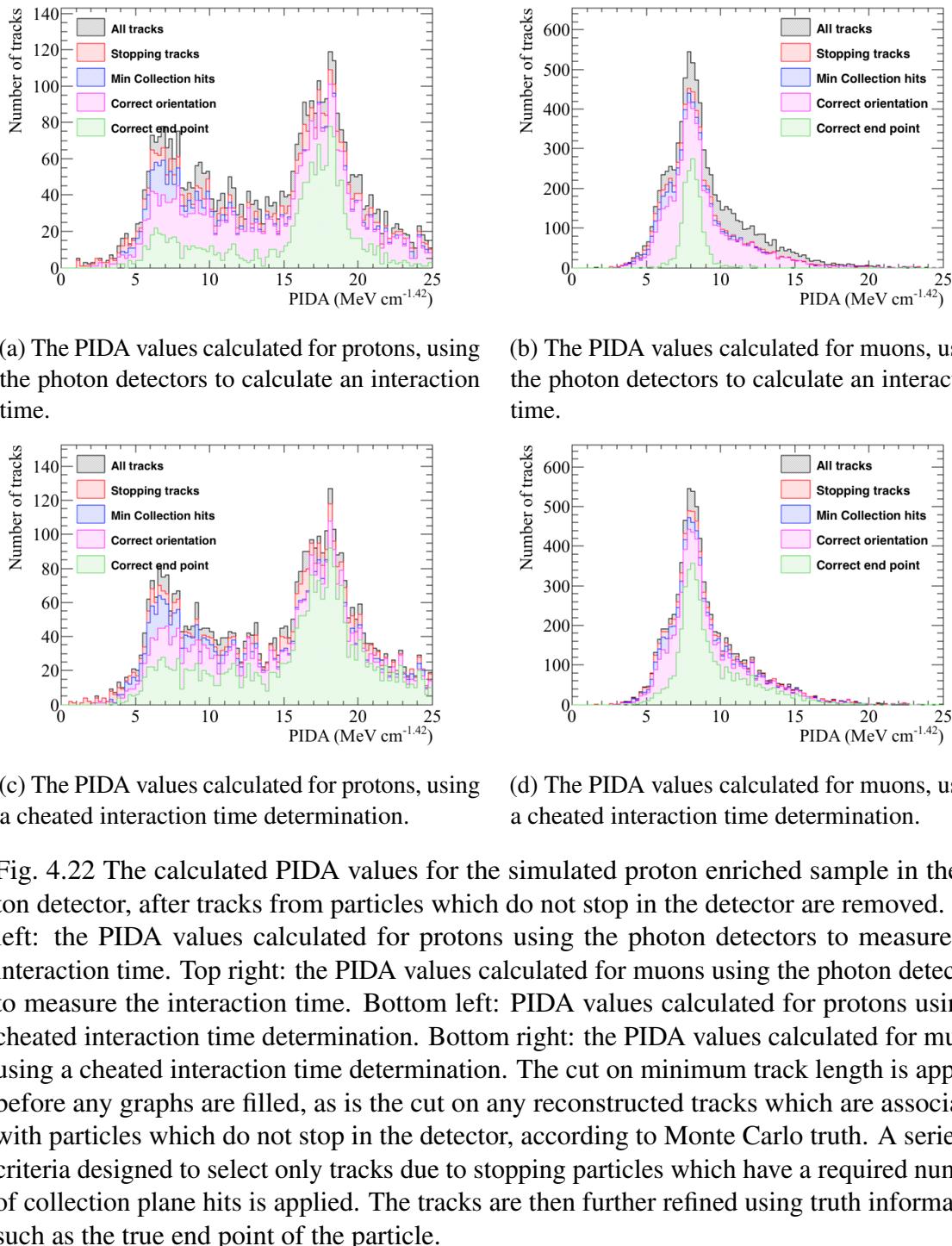
However, the basic framework by which PID can occur has been outlined here, and
16 is shown to be very effective when particles are simulated in isolation. It has also been
17 shown that a clear separation between muons and protons can be seen in cosmic rays, though
18 there is a non-negligible contamination of the proton track PIDA range by muon tracks.
19 This separation was only possible when using information from Monte Carlo truth though,
20 and so a more sophisticated method of identifying through-going muons is required, should
21 the analysis be performed on the cosmic ray data, such as that collected by the 35 ton detector.
22

23

24

25

4.4 Performing particle identification



(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

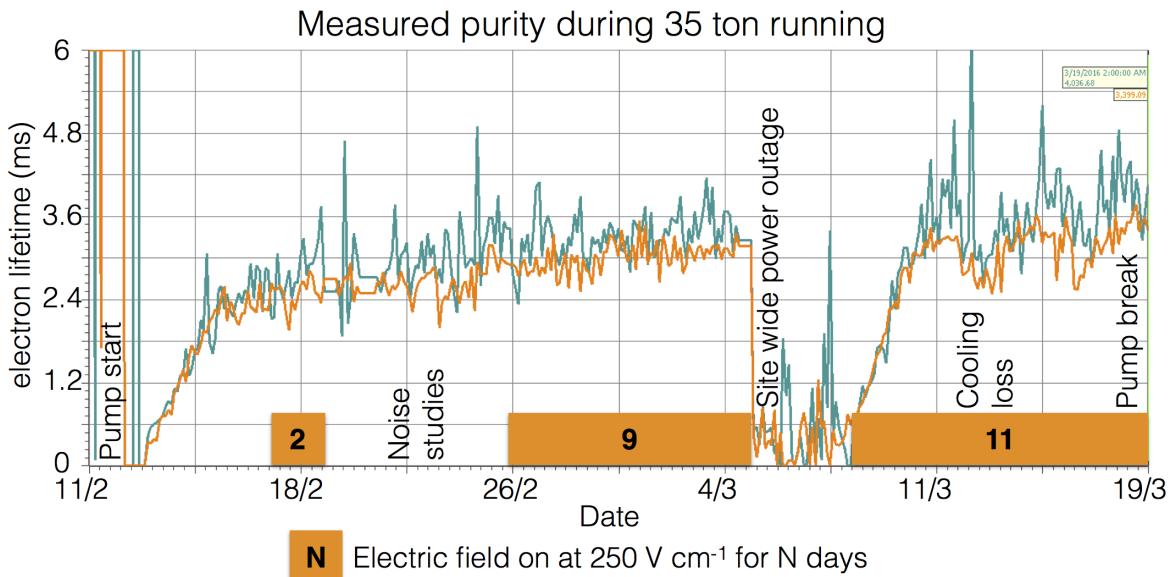


Fig. 5.1 Timeline showing the data collected during the 35 ton Phase II run once the purification pumps were turned on.

¹ rays. The DAQ combined these three data streams into synchronous events in time and saved
² them as LArSoft data objects. These data objects would later have to be converted to the
³ offline data products, which the reconstruction tools developed on simulation used, this is
⁴ discussed in Section 5.2. This section describes the structure of the data objects in the raw
⁵ form.

⁶

⁷ During operations the DAQ was configured to maximise data throughput, and physics
⁸ potential. This meant recording different lengths of times for each of the three sub-systems,
⁹ as the data volumes and length of physics information were significantly different. For
¹⁰ example, due to the emission of prompt light, the physics information from the SSPs is of a
¹¹ much shorter length of time than the physics information from the RCEs, where data has to be
¹² recorded whilst the electrons drift through the LAr. During the running period the recorded
¹³ data was triggered by through-going muons which produced coincidences on the CRCs on
¹⁴ opposite side of the cryostat. A coincidence is defined as two CRC modules recording a hit
¹⁵ within 32 ns of each other. The system used to collect the CRC data was also responsible
¹⁶ for generating the triggers, and so this meant that the trigger rate could be suppressed to
¹⁷ approximately 1 Hz, by only producing triggers every N times a coincidence occurred, where
¹⁸ N was a tuneable variable. A trigger rate of 1 Hz was used as the maximum speed at which
¹⁹ data could be written to disk was approximately $60 \text{ MB} \cdot \text{s}^{-1}$, which is roughly equal to the
²⁰ 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. 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 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),

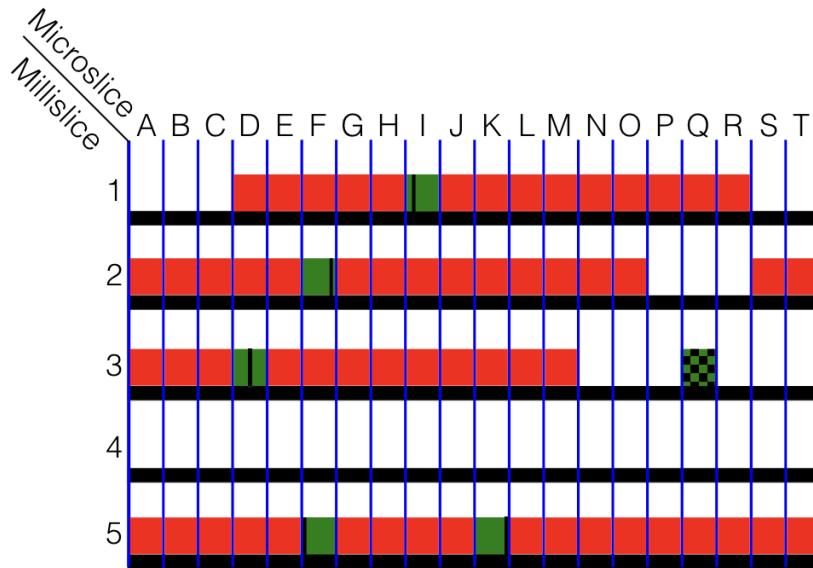


Fig. 5.2 A diagram of possible millislice structures for the TPC data recorded by the 35 ton detector. Each row represents a millislice (numbered 1 to 5), whilst each box represents a microslice (labelled A to T). The vertical blue lines delineate each microslice, giving 20 microslices per millislice. 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 millislice 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.

¹ and the start of physics event 3 (2S).

²

³ During normal data taking microslices were buffered in the RCEs, so that if a trigger was
⁴ issued they could be accessed before being deleted. As the data was buffered in the form of
⁵ microslices, previous microslices could only be accessed as a whole. This meant that a whole
⁶ number of microslices had to be loaded before the trigger, so when a trigger was issued part
⁷ way through a microslice, the previous X microslices were sent to the event builders, where
⁸ X was typically 5. As a result, there are always a minimum number of ticks both before
⁹ (5,000 ticks) and after (9,000 ticks) the trigger, but the exact numbers can change by up to
¹⁰ 1,000 ticks for a given event, depending on where in the microslice the trigger came. The
¹¹ result of this is that it is impossible to know the number of ticks before/after a given counter
¹² 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 flight, 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

5.3 Observations on data quality and noise mitigation

122

¹ record so that it can be accessed later if required.

²

³ It is important to integrate flexibility at all points in this process, so that the user can
⁴ choose the length of events, which sub-systems are in the events, and what the conditions
⁵ are for making events. It is also important for users to be able to run the service on already
⁶ formatted events, as the unpacking services are the major overhead in running the interface. It
⁷ is also conceivable that users would want to reformat Monte Carlo events so as to centre them
⁸ around their chosen conditions, and so the use of the unpacking algorithms was determined
⁹ by the interface depending on the format of the input file.

¹⁰ **5.3 Observations on data quality and noise mitigation**

¹¹ Reformatting the online data to the offline format was an important step in maintaining
¹² data quality, as subsequently there was no access to the raw data due to the framework of
¹³ the 35 ton software. Some of the important checks which are performed are outlined in
¹⁴ Figure 5.3. If any of these issues are present in a given physics event it is discarded as
¹⁵ the integrity of the data cannot be guaranteed. It was decided that these events would be
¹⁶ discarded as non-synchronous events would lead to hits in the detector being at incorrect
¹⁷ times, and padding empty events with pedestals could mean that tracks seem to disappear as
¹⁸ they travel through the detector, only to reappear at a later time in a different detector location.

¹⁹

²⁰ Another example of an inconsistent event is when the sub-systems are not synchronised
²¹ with each other. This is normally caused by one of the sub-systems missing a clock increment
²² from the master timing unit, due to the data trigger being issued close to an increment from
²³ the master unit. This misalignment causes an incorrect time sample being read out, and so
²⁴ the data from each sub-system within a millislice is not consistent. The result of this is that
²⁵ the event will fail the timestamp check, and so won't be added to the event record. To avoid
²⁶ incomplete events, these physics events are also discarded when observed.

²⁷

²⁸ The electronic noise in the 35 ton was higher than anticipated, with the RMS of the RCE
²⁹ ADC being approximately 30 counts, compared to an expected thermal noise of around
³⁰ 2.5 ADC counts. Many sources contributed to this elevated noise, some of which are ex-
³¹ plained below.

³²

³³ Though not directly affecting the noise issues “stuck ADC codes” were a feature of the
³⁴ data which had to removed. “Stuck ADC codes” were caused by bit level corruption where

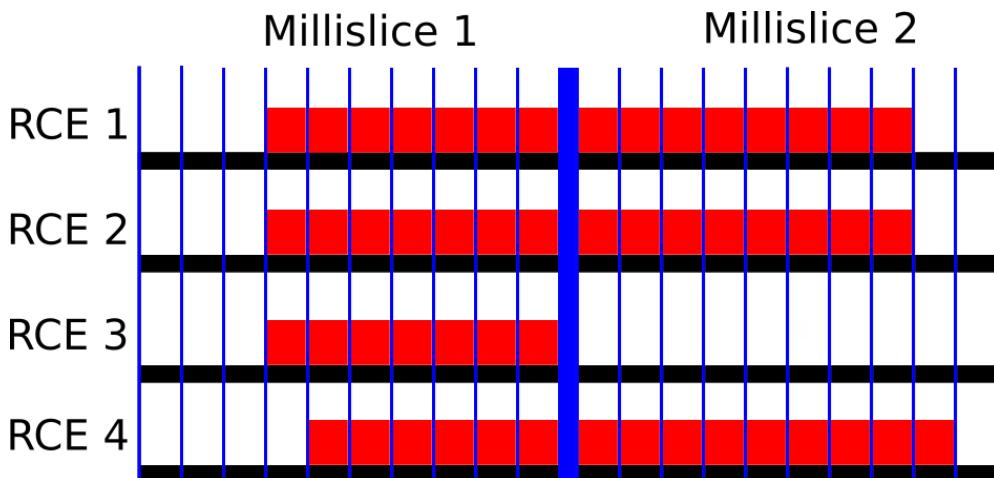


Fig. 5.3 A diagram of TPC microslices within millislices in the 35 ton data stream. Two millislices are shown, each containing 10 microslices. One physics event straddling the millislice boundaries is shown, and 4 RCEs representing each row are read out. The vertical blue lines delineate each microslice (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.

5.3 Observations on data quality and noise mitigation

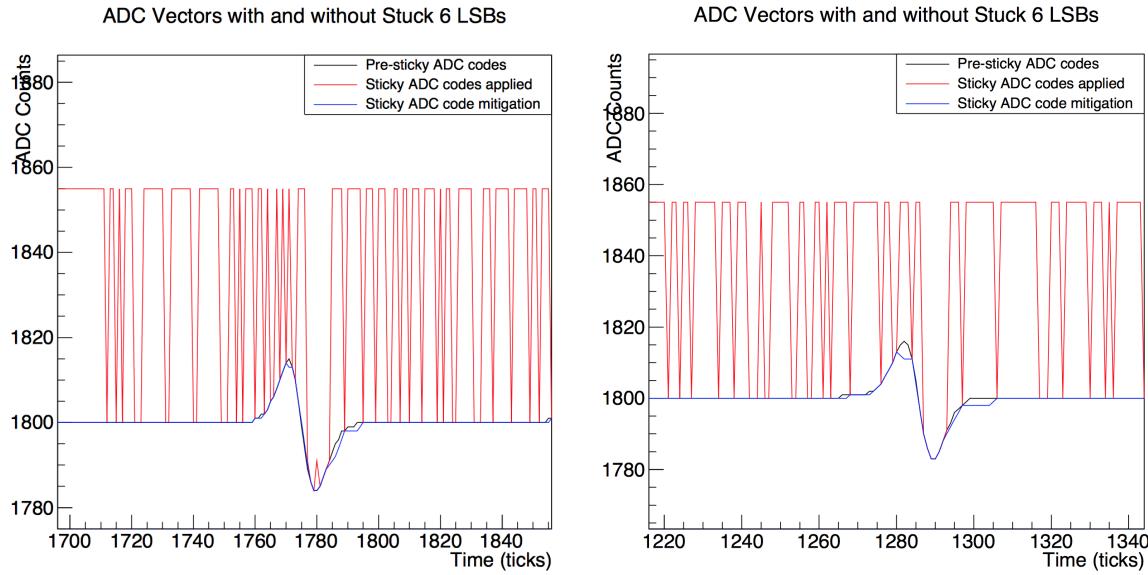


Fig. 5.4 Two Monte Carlo spectra showing the effect of the introduction and removal of stuck bits on a simulated signal. The black line shows the simulated signal on a wire, which is then modified by adding the effects of “stuck ADC codes,” shown by the red line. The “stuck ADC codes” are then removed, 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 [135].

1 the lowest 6 bits in the ADC became frozen to either 0x0 or 0x3f. This was observed during
 2 the first stages of commissioning, and an algorithm to remove them was developed and tested
 3 on Monte Carlo [135]. In simulations it was observed that the signal could be recovered with
 4 minimal losses, as shown in Figure 5.4, where the signals after stuck code removal (blue
 5 lines) are seen to closely match the signals before stuck codes were added (black lines).

6

7 A significant portion of the noise was correlated between groups of 32 channels, where
 8 the ADCs would coherently oscillate. To remove these coherent shifts, ADC baselines
 9 were calculated for these groups of 32 channels at each tick, and then subtracted from the
 10 measured ADC values. This was found to be an effective method of removing coherent noise
 11 in MicroBooNE [136]. The effect of removing coherent noise is shown in Figure 5.5, where
 12 the signal peak becomes much easier to discern after noise removal, and a coherent noise
 13 peak around tick 6030 is removed. An issue with removing coherent noise in this way is that
 14 events which are parallel to the Anode Plane Assembly (APA) frames will produce signals at
 15 common times, and across adjacent wires, and these signals may be removed along with the
 16 coherent noise. This will cause a reduction in the hit reconstruction efficiency. The only way
 17 to prevent this is to “protect” potential signal regions from the coherent noise removal, as is

5.3 Observations on data quality and noise mitigation

125

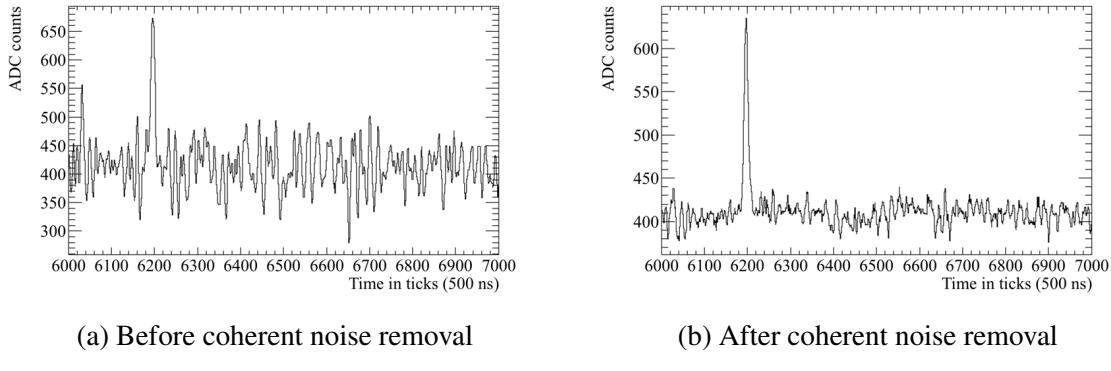


Fig. 5.5 The effect of coherent noise removal on a 35 ton signal event. Left: the signal before coherent noise is removed. Right: the signal after the coherent is removed. The signal peak around tick 6200 is much clearer after coherent noise removal, meaning that hit reconstruction becomes much simpler.

done in MicroBooNE [136].

After performing a Fast Fourier Transform (FFT) [137] 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 [138]. 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. 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 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 [139].

An example of the effect of the noise mitigation steps is shown in Figure 5.7a. 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

5.3 Observations on data quality and noise mitigation

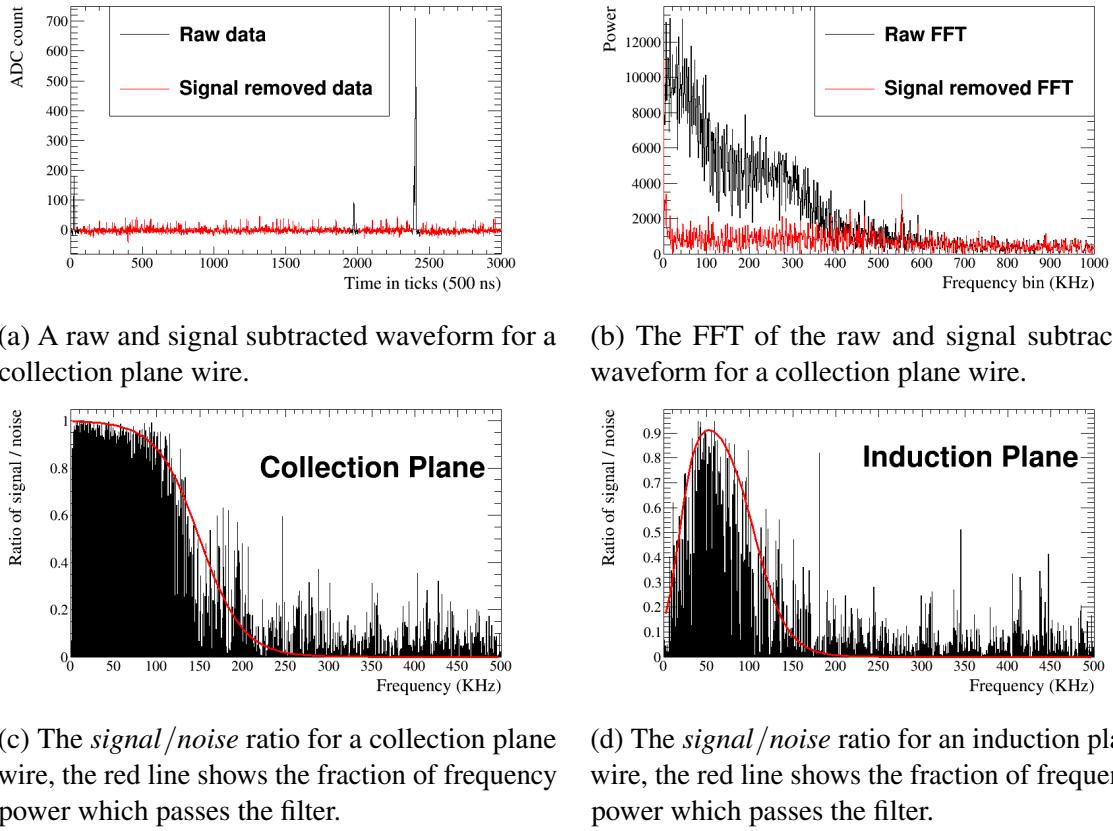


Fig. 5.6 The application of Wiener filters to the 35 ton data. Top left: a waveform from a collection plane wire which is then signal suppressed. Top right: the FFT of both the raw and signal suppressed waveforms. Bottom left: the *signal/noise* ratio for the collection plane waveform, a sigmoid function 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*.

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 [139]. 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 [140].

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.

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.

5.4 Performance of reconstruction algorithms

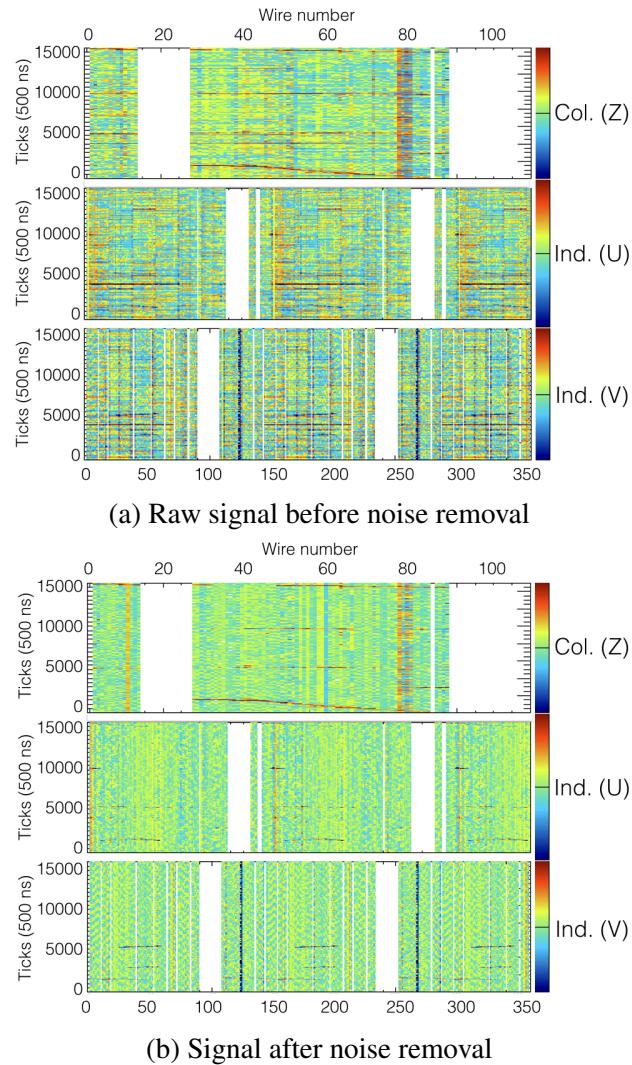


Fig. 5.7 Event displays showing the effect of the noise removal algorithms on data from the 35 ton detector. The event displays show the signals in the collection (Z), and induction planes (U and V). The plots show wire number, and time in ticks, on the x and y axes respectively. The charge is shown 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? 11
 - This is because when there is a signal on an induction plane there should also be signals on the collection wires. 12 13
- Are there adjacent wires which have hits at a similar time? 14
 - This is because one would expect a track to deposit energy on multiple adjacent wire segments. 15 16
- Which hit lies closest to the line defined by unique collection plane hits in the xz plane? 17
 - 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. 18 19 20 21 22
 - This can be changed to consider the line defined by previously selected hits in the given TPC and plane where the hit choices are. 23 24

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 standard reconstruction has been used in the discussions to follow [141]. There has since been an effort to improve the counter shadowing hit disambiguation to remove the outlying collection plane hits using the MLESAC method [142], whereby points which are far away from a best fit are ignored. These studies are still ongoing [143]. 25 26 27 28 29 30 31

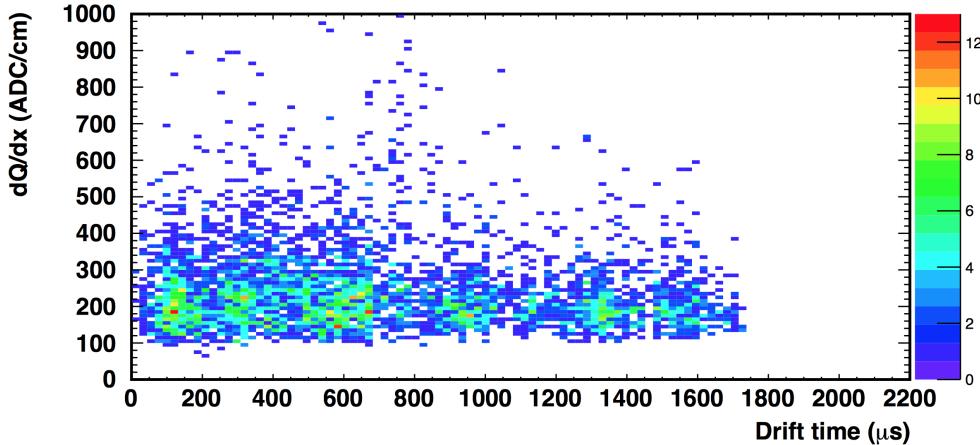


Fig. 5.8 The dQ/dx values for a sample of muon collection plane hits, note the cutoff at $100 \text{ ADC}\cdot\text{cm}^{-1}$ due to the hit finding threshold. The figure is taken from [145].

1 A symptom of the elevated noise state is that signals are often dropped on one of the
 2 induction planes, this means that the tracking algorithms often have to combine clusters in
 3 only two of the three planes. Reconstruction using two planes was shown to be effective
 4 by the ArgoNeuT collaboration [144], so the loss of signal in one of the three planes is not
 5 prohibitive to track reconstruction. Another consequence of the elevated noise level is that
 6 even when the counters are used to seed hit finding, the hit finding threshold is too high to re-
 7 construct the very lowest hits. This causes the plot of dQ/dx for muons, shown in Figure 5.8,
 8 to look flat, due to a cutoff at $100 \text{ ADC}\cdot\text{cm}^{-1}$, below which no hits are reconstructed. The
 9 inability to reconstruct the lowest energy hits means that calorimetry is all but impossible
 10 on the 35 ton dataset, even though the tracking algorithms perform relatively well. The
 11 inability to perform reliable calorimetry en masse means that the only particles which can be
 12 assuredly identified are the muons which triggered the counter coincidences. This means that
 13 performing the analysis that was presented in Section 4.4 on the 35 ton dataset is extremely
 14 difficult, if not impossible.

15

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

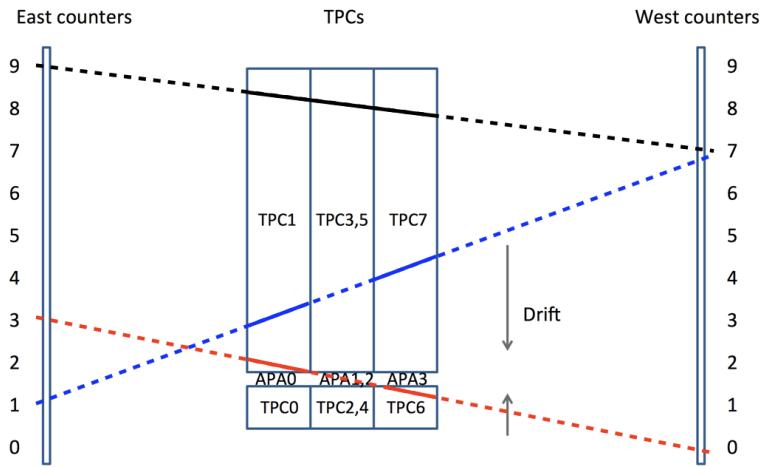


Fig. 5.9 The numbering scheme for the east - west counters in the 35 ton. The counters have been numbered from 0 to 10 depending on their position from the end of the short drift volume. This is different to the LArSoft numbering scheme shown in Figure 2.16 where they go from 6-15 and 28-37 for the east 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 time, and any tracks reconstructed from the noise will have random directions. This is shown in data where if multiple tracks pass the dot product cut they are co-linear, and are not randomly orientated, as shown in Figure 5.11.

1
2
3
4
5
6
7
8
9
10
11
12
13
14

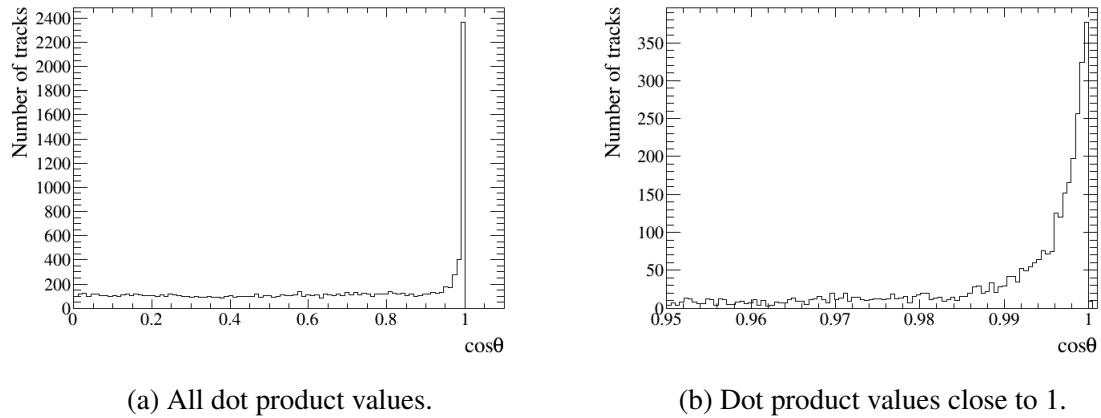


Fig. 5.10 The dot product of the track and vector joining the centres of the coincidence counters in the yz plane. The number of tracks with a given dot product is plotted on the y axis. A threshold value of 0.98 is required for a track to be considered to be due to the counter coincidence. It can be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

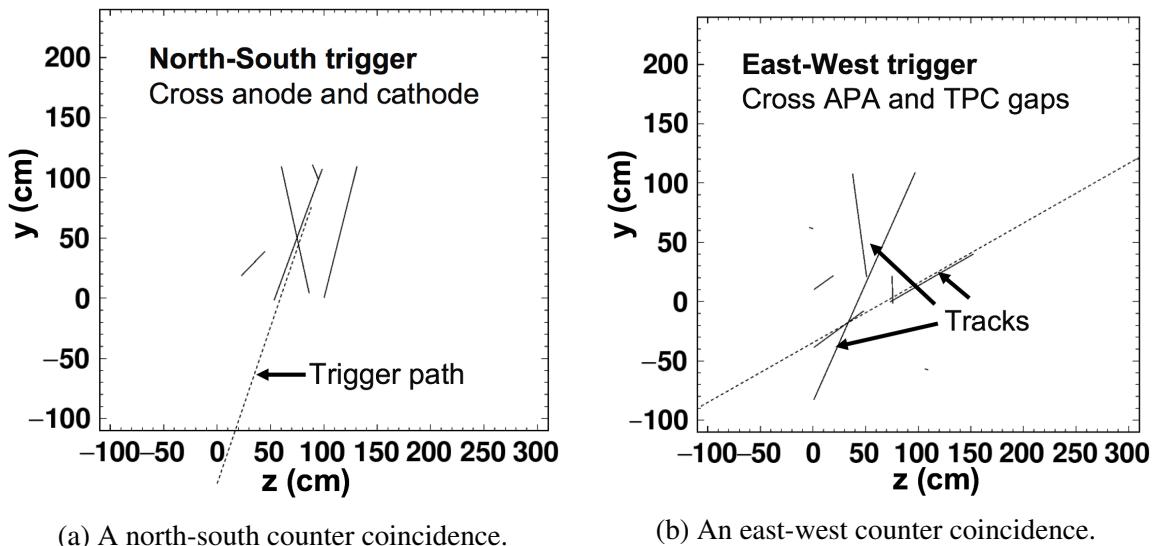


Fig. 5.11 The alignment of reconstructed tracks with the vectors joining the centres of the coincidence counters. The dashed lines show the vectors joining the centres of counters hit in the coincidence, whilst the solid lines show the reconstructed tracks. Left: the alignment of tracks with a north-south 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 [141].

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.

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 5.10, 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 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

Table 5.1 The angles which tracks, with given counter differences, have relative to the APA frames. Though the east and west counters have a width in the y (vertical) direction, this is much less than their extent in the z direction. The depth of the counters, their extent in x , is negligible compared to the separation of the east-west counters. The counters have identical widths in both the y and z directions. The angles are calculated using the difference in the centres of the counters in the z direction, divided by the separation of the east and west counters in z .

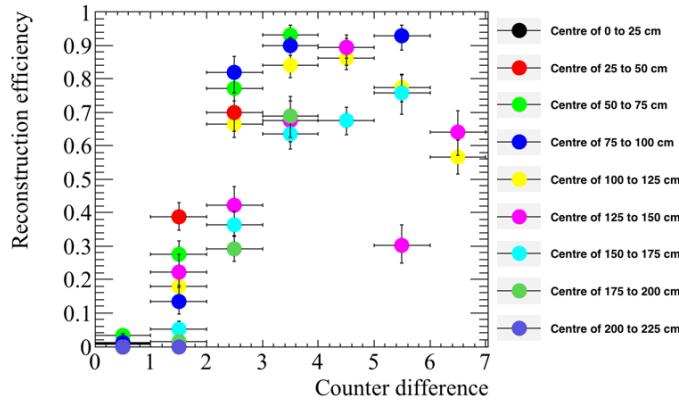
Absolute counter difference	Approximate angle ($^{\circ}$)
0	0 ± 2.1
1	4.2 ± 2.1
2	8.4 ± 2.0
3	12.5 ± 2.0
4	16.5 ± 2.0
5	20.3 ± 1.9
6	23.9 ± 1.8
7	27.3 ± 1.7
8	30.7 ± 1.6
9	33.5 ± 1.5

¹ following discussion. Only coincidences which would produce tracks that are contained
² within the long drift volume are considered here, hence there being no negative x positions.

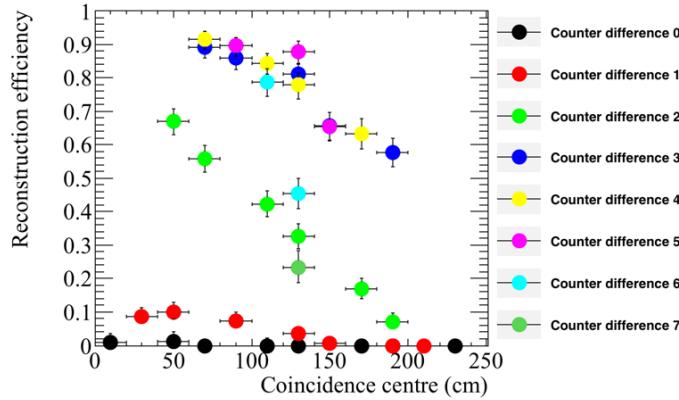
³

⁴ From Figure 5.12a, it is evident that the reconstruction efficiency for tracks with shallow
⁵ angles relative to the APAs is extremely poor, with the efficiency for tracks aligned with
⁶ counter differences of 0 or 1 never rising above 10%. This is due to the coherent noise
⁷ removal, where hits which are correlated in time will be removed as they will be perceived as
⁸ being due to the coherent noise, as opposed to being due to real signals. As the difference in
⁹ counter number increases, the efficiency is seen to increase, though the rate of this increase
¹⁰ is seen to depend on the “coincidence centre”. The effect of increasing “coincidence centre”
¹¹ can be seen more clearly in Figure 5.12b, where the efficiency for each counter difference
¹² as a function of “coincidence centre” is plotted. Here, it can be seen that the reconstruction
¹³ efficiency decreases for coincidences that are centred further away from the APAs. This is
¹⁴ due to the fact that when an energy deposition has further to drift, it will induce a smaller
¹⁵ pulse on the wires, meaning that it is more likely to be below the hit threshold. Figure 5.12c
¹⁶ combines Figures 5.12a and 5.12b, to show how the reconstruction efficiency for increasing
¹⁷ “coincidence centre” changes, with increasing counter difference. It can be seen that tracks
¹⁸ with counter differences of between 3 and 5, where the “coincidence centre” is between
¹⁹ 60 cm and 140 cm away from the APAs, are the best reconstructed coincidences. Finally,

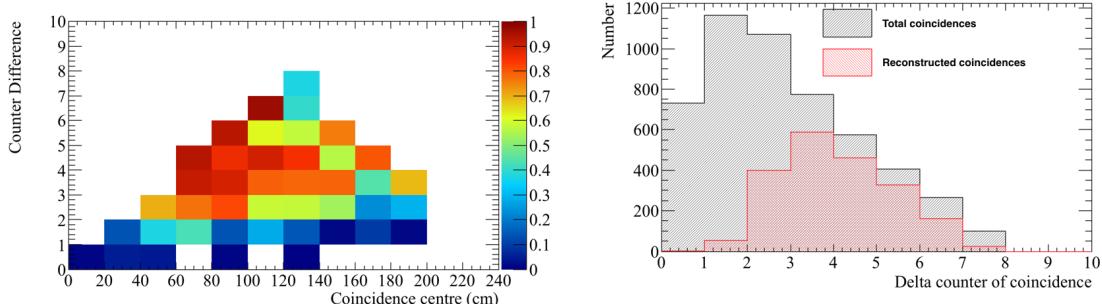
5.4 Performance of reconstruction algorithms



(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.

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

8 **5.5 Measuring interaction times using electron diffusion**

9 As electrons drift from the interaction point to the wire planes they become spread out in
10 both time and space, this effect is known as diffusion, and is an important property of elec-
11 tron transport in LAr, which must be well understood. The mechanism by which diffusion
12 occurs in LAr was first discussed in [146–149], and has since been developed to consist of a
13 complete set of measurements for electric fields between 100 and 2000 V·cm⁻¹ [150]. The
14 diffusion of electrons is rarely isotropic, and so the component that is transverse to the drift
15 field, and the component that is parallel to the drift field, are normally measured separately.
16 Diffusion parallel to the drift field is called longitudinal diffusion, and is generally smaller
17 than the component of diffusion that is transverse to the drift field. Figure 5.13 shows how
18 diffusion can smear the electrons collected on a set of wires when the electrons are initially
19 highly correlated in time and space.

20
21 Longitudinal diffusion has the effect of spreading the drifting electrons out in time, caus-
22 ing signals to become wider in time, and smaller in height, as the total charge is conserved.
23 The increasing hit width can be measured for increasing drift times (distances), provided the
24 hits do not fall below a hit finding threshold. Transverse diffusion causes drifting electrons to
25 spread out in space, changing the amount of charge deposited on a wire, and reducing the
26 charge resolution of the detector. Transverse diffusion is measured by discerning how the
27 width of the hit charge distribution changes for increasing drift distances [150].

28
29 Through-going particles make ideal tracks to study diffusion as they are minimally ion-
30 ising, and so have roughly constant energy depositions along their tracks. The tracks that
31 they produce can also cover a wide range of drift distances, if they are not parallel to the
32 APAs. The drift distances of hits within a track can be determined by matching the track with
33 a counter coincidence as discussed at the end of Section 5.4. The *x* positions of the hits can

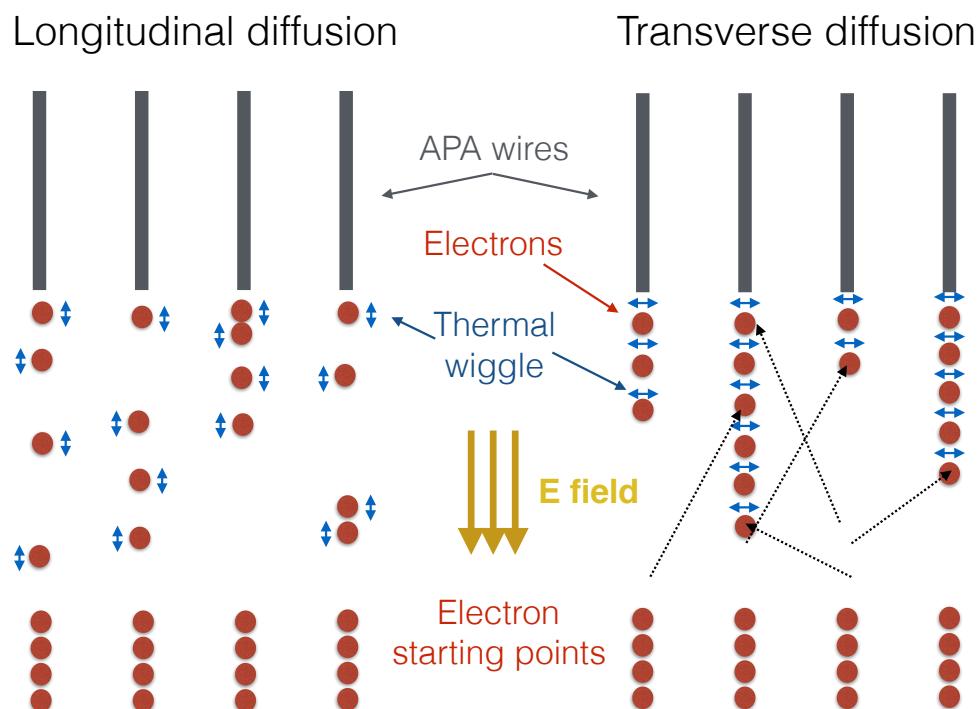


Fig. 5.13 A schematic showing the longitudinal diffusion (left), and the transverse diffusion (right), of electrons. In both cases, four electrons are initially shown below four wires, and are allowed to diffuse in either the drift direction, or perpendicular to the drift direction, in the longitudinal 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 [151].

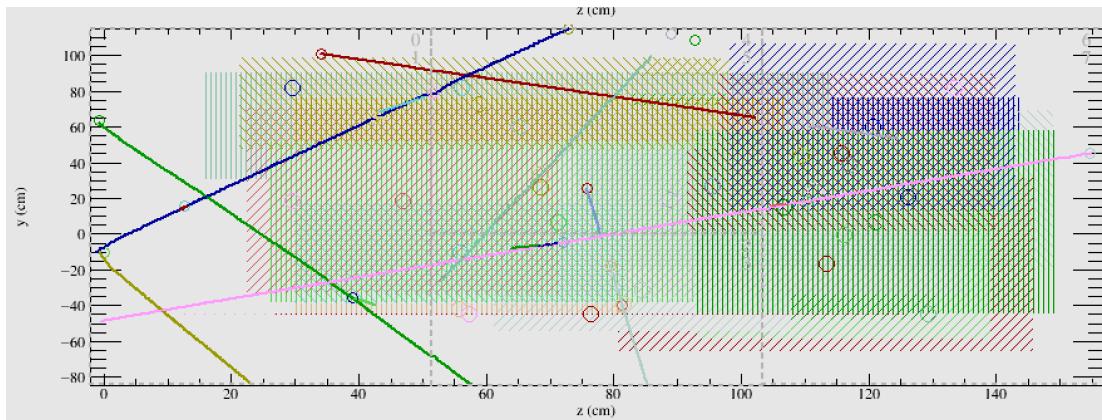


Fig. 5.14 A simulated event display showing multiple tracks and flashes to be assigned to each other in the 35 ton detector, in the yz plane. The coloured lines represent reconstructed tracks, whilst the coloured dashed boxes represent flashes.

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

2

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

16

17 The 35 ton dataset is ideal for testing this hypothesis, as the counters are able to provide
18 a sample of tracks with known angles and interaction times, which can be used to tune
19 interaction time determination metrics. These metrics can then be applied to another sample
20 of tracks, where the interaction time is known but not used, so that the accuracy of the
21 calculated interaction times can be found. As longitudinal diffusion is the dominant effect
22 that increases the hit width, transverse diffusion will not be directly considered further. The
23 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 rate 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 as the noise level in the detector increases, 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. Tracks are considered en masse, and so the hits for every track are separated into 10 cm

5.5 Measuring interaction times using electron diffusion

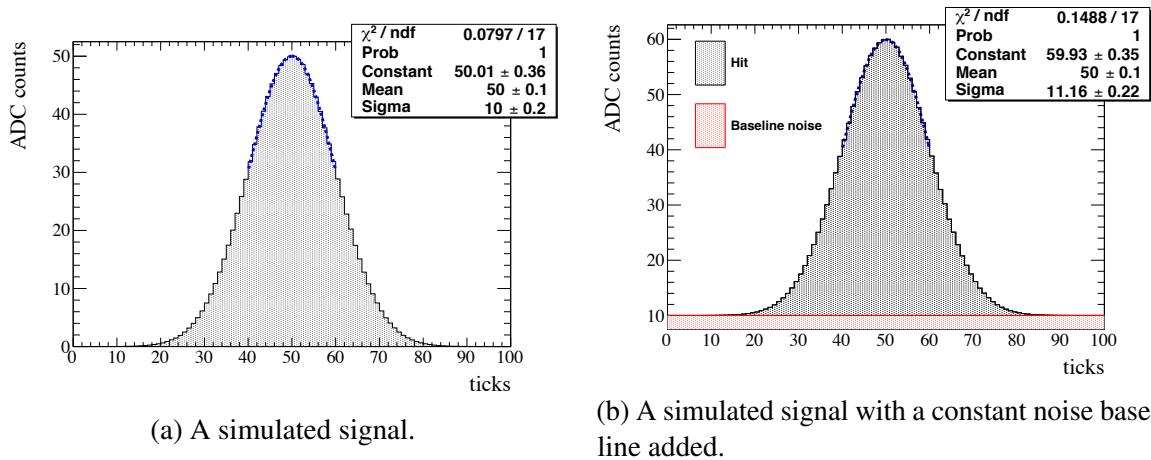
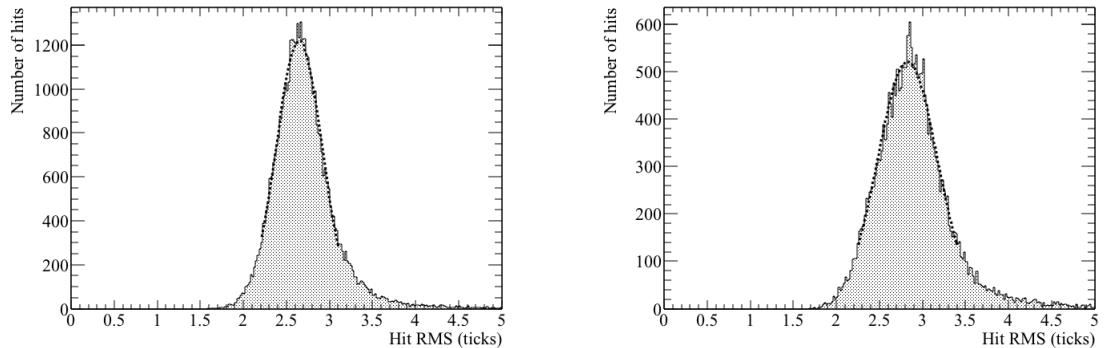


Fig. 5.15 ADC counts as a function of time for a simulated signal with a width of 10 ticks, and an amplitude of 50 ADC counts, both before and after, a constant noise baseline of 10 ADC counts is added. The simulated ADC value is shown on the y axis, and the time, in ticks, is shown on the x axis. In reality the noise would fluctuate with time. 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 noise can cause the measured width of a hit to increase. The figure is taken from [151].

1 regions of increasing drift distance from the APAs. The following quantities are calculated
 2 for each 10 cm drift region:

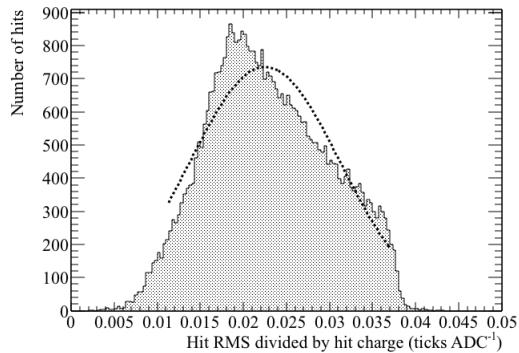
- 3 • The hit *RMS* - the most direct way to measure transverse diffusion.
- 4 • The hit *RMS/Charge* - an attempt to incorporate the effect of impurities in the LAr for
 5 relatively low purity data, as this will have a drift distance dependence.
- 6 – The charge of a hit is calculated by integrating the ADCs of the reconstructed hit
 7 over time.
- 8 Fitting Gaussian functions around the peaks of the distributions will yield the most probable
 9 values for the drift regions, as is shown in Figure 5.16.

10
 11 From Figure 5.16 it is clear the the width of the hit *RMS* distribution increases for
 12 hits which are further from the APAs. However, the width of the hit *RMS/Charge* is seen
 13 to decrease, though this is due to a sharp cut-off at a hit *RMS/Charge* equal to roughly
 14 0.038 ticks·ADC⁻¹. The reason for this cut off was shown in Figure 5.8. It can also be seen
 15 that the most probable values of both the hit *RMS* and hit *RMS/Charge* increases with drift
 16 distance.

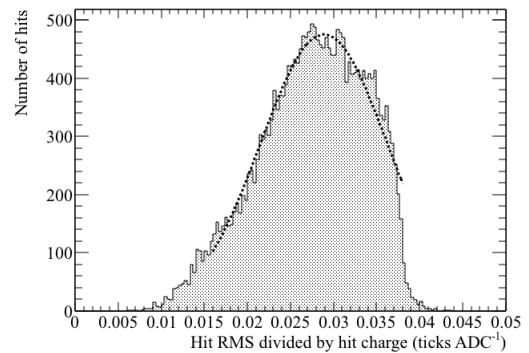


(a) The distribution of hit *RMS* value for hits between $x = 20$ cm and $x = 30$ cm.

(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.

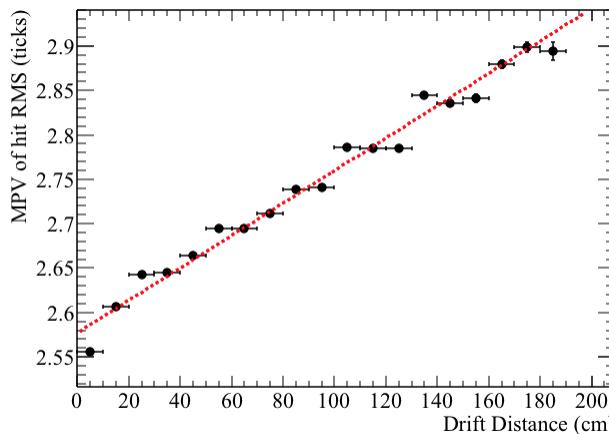


Fig. 5.17 The most probable values of hit RMS as a function of drift distance, for the hits within a track, associated with a coincidence that had a counter difference of 4.

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

7

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

15

16 To consider single tracks, the best line fits for the counter differences for a large sample
 17 of tracks, such as in Figure 5.17, are required. These best line fits can then be used to predict
 18 the position you would expect a hit to originate from, given values for the hit RMS and
 19 hit $RMS/Charge$, and the angle of the track to which it belongs. The predicted positions
 20 can then be compared to the known position from the counter coincidence to determine the
 21 accuracy of the prediction.

22

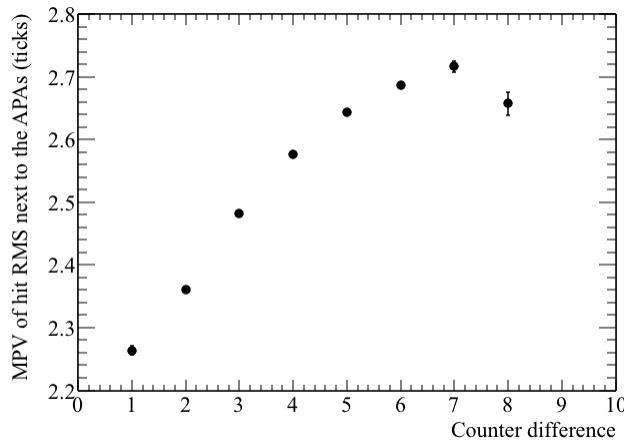


Fig. 5.18 The most probable values of hit *RMS* within 10 cm of the APAs, as a function of the counter difference of the coincidence, that the track, to which the hits belong, was associated with.

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 normalised hit charge, whereby the most probable hit charge has a value of 1. A conservative cut on normalised hit charges of 0.25 is made, so that it can be guaranteed that the tails 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). The difference in the predicted and reconstructed hit times should then be centred around the interaction time, and be distributed much more uniformly. This will mean that the interaction time can be determined much more accurately.

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

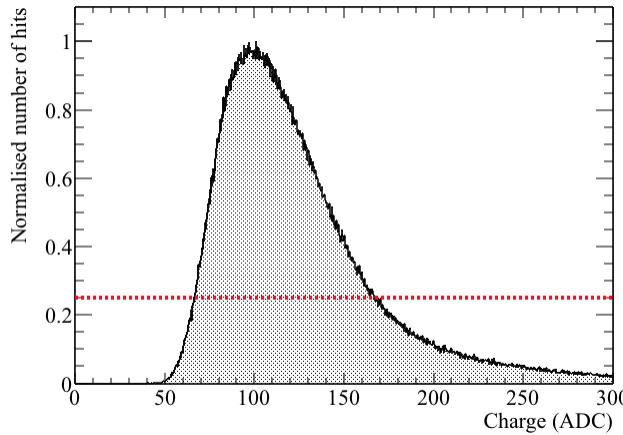


Fig. 5.19 The distribution of normalised hit charge, shown in units of ADC, in the 35 ton dataset. The number of hits with the most probable hit charge has been normalised to a value of 1. A cut on the normalised number of hits being greater than 0.25 is shown, the aim of this cut is to remove the tails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

¹ problems in the 35 ton detector, very few tracks will have hits on all of these collection
² wires. However, requiring at least 100 collection plane hits is not unreasonable, and would
³ correspond to a reconstructed track length of at least 50 cm. The difference between the
⁴ predicted and reconstructed hit time for each hit is shown in Figure 5.20, for both the hit
⁵ *RMS* and hit *RMS/Charge* metrics.

⁶
⁷ Figure 5.20 shows that both distributions are centred around a time difference of 0 μs
⁸ in the 35 ton dataset. This is encouraging as it shows that the method has potential. The
⁹ width of the distribution for the *RMS/Charge* metric is smaller, and the peak higher, so it is
¹⁰ expected that this will provide the more robust metric. This is because these features show
¹¹ that the predicted hit times are likely to be close to the reconstructed hit times. The peaks are
¹² centred around a time difference of 0, as the hit times had previously been corrected using
¹³ the measured interaction time from the counter coincidence. This was done so as to avoid
¹⁴ the uncertainty which would arise from allowing the coincidences to remain at random times
¹⁵ between ticks 5000 and 6000, in the 15000 tick event. For an explanation as to why this
¹⁶ occurs, see the discussion concerning Figure 5.2.

¹⁷
¹⁸ It is interesting to observe the non-symmetric nature of the hit *RMS/Charge* distribution
¹⁹ which is seen in Figure 5.20. This is caused by there being a tendency for the prediction
²⁰ metric to overestimate the *x* positions of the track. However, it is clearly evident that many
²¹ of the predicted *x* positions are highly accurate, as there is a definite peak around a difference

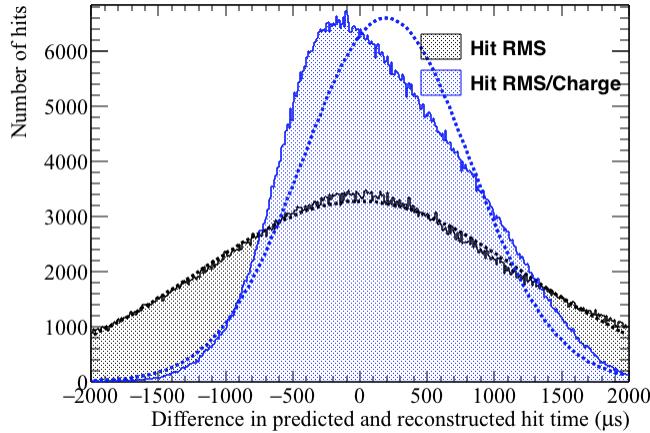


Fig. 5.20 The difference between the predicted and reconstructed hit times in the 35 ton dataset. The differences in time when the hit *RMS* metric is used are shown in black, whilst the differences in time when the hit *RMS/Charge* metric is used are shown in blue.

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 [152]. 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

5.5 Measuring interaction times using electron diffusion

146

1 information can be confidently used.

2

3 Figure 5.21 shows that using the effects of diffusion, and the hit *RMS*, the interaction
4 time and central *x* position of a track, can be reliably predicted in the 35 ton dataset. The
5 accuracy in determining the interaction time is found to be 240 μs , where the distribution
6 has a FWHM of 281 μs . When this is converted into the difference in central *x* position of
7 the track, the accuracy is found to be 27.7 cm, with a FWHM of 32.1 cm.

8

9 Figure 5.22 shows that using the effects of diffusion, and the hit *RMS/Charge*, the inter-
10 action time and central *x* position of a track, can be reliably predicted in the 35 ton dataset.
11 The accuracy in determining the interaction time is found to be 171 μs , where the distribution
12 has a FWHM of 210 μs . When this is converted into the difference in central *x* position of
13 the track, the accuracy is found to be 18.5 cm, with a FWHM of 23.0 cm.

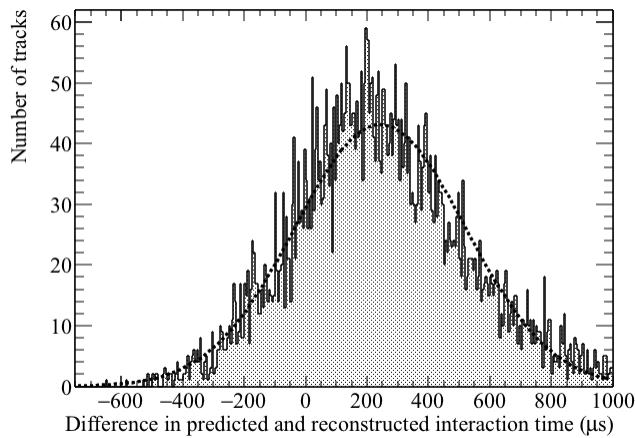
14

15 The resolutions found are quite impressive, as given that the total drift time for electrons
16 through the whole 35 ton detector volume of 223 cm is roughly 5200 ticks, it means that
17 tracks can be cleanly distinguished throughout the detector volume. Though the resolution
18 of the interaction time determination is impressive, it is concerning that there appears to
19 be a systematic offset which has been introduced. This is seen by the distributions not
20 being centred around a difference in predicted and reconstructed times of 0 μs , as would
21 be expected from the discussion concerning Figure 5.20. As discussed earlier, the issues
22 with noise in the 35 ton dataset affect the accuracy with which tracking and calorimetry can
23 be performed, and so it is reasonable to expect that the effectiveness of the interaction time
24 determination was also affected. Therefore, it is prudent to repeat the study on a Monte Carlo
25 sample, with the same detector conditions, but a much lower level of detector noise. This is
26 presented in Section 5.5.2.

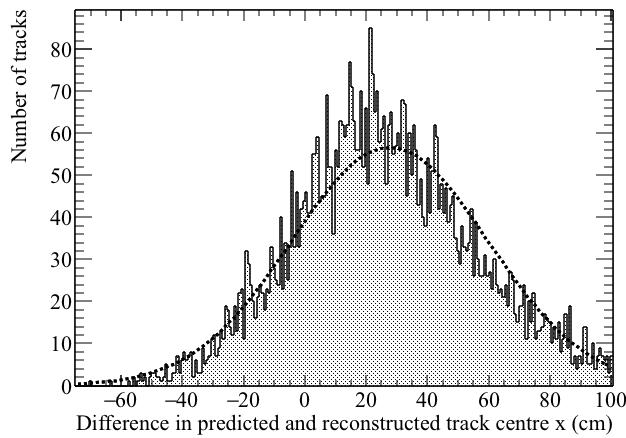
27

28 **5.5.2 Determining interaction times in a low-noise detector using Monte Carlo, and differences with data**

29 When determining interaction times in Monte Carlo simulations, exactly the same criteria
30 are applied to the hits. This is because δ -rays would still change the measured hit width, and
31 will be present in any sample. In a low noise detector it is expected that few wires would
32 be removed due to being noisy, but for consistency there is no danger in applying this cut.
33 Imposing a minimum number of collection plane hits is again important to ensure that the

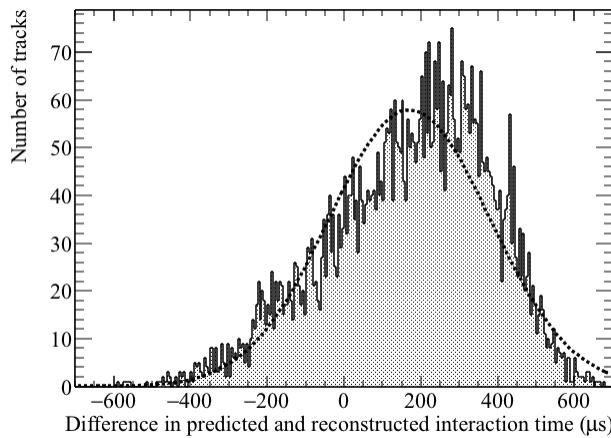


(a) The average difference in interaction times using the hit *RMS* metric.

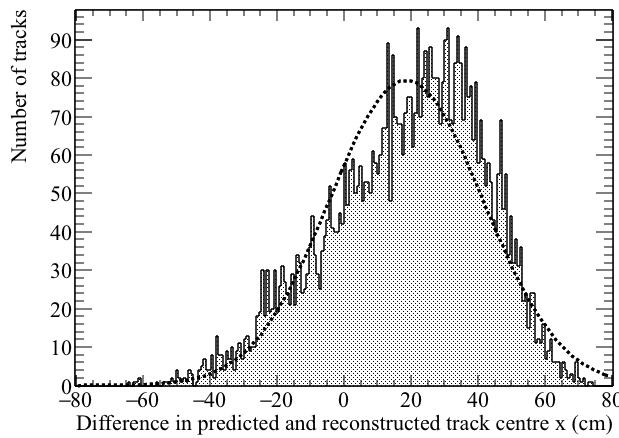


(b) The average difference in the central *x* position of a track using the hit *RMS* metric.

Fig. 5.21 The accuracy of the hit *RMS* method in the 35 ton dataset. Top: the accuracy to which interaction times can be determined in μ s. 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.

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 more tightly distributed 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 ticks·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.

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

5.5 Measuring interaction times using electron diffusion

150

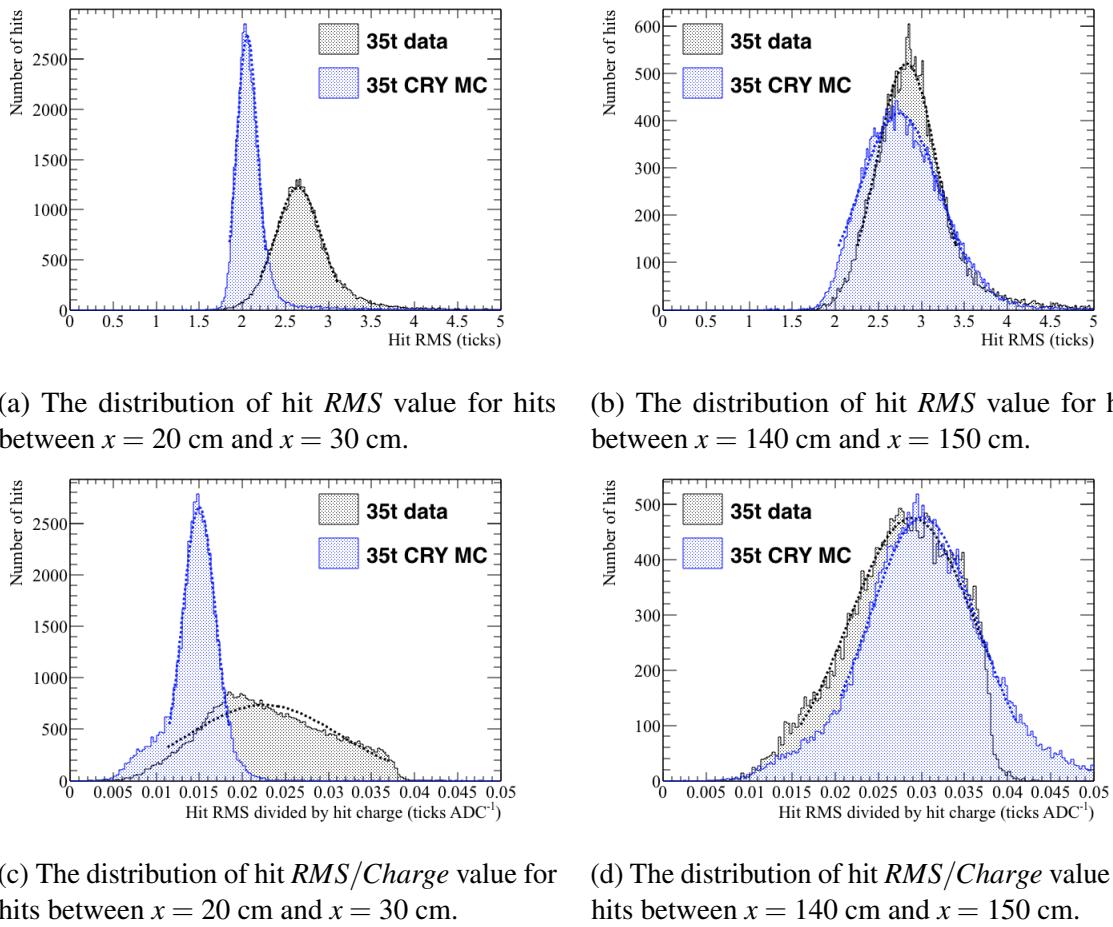


Fig. 5.23 The distributions of hit *RMS* (top), and hit *RMS/Charge* (bottom), for points between 20 and 30 cm from the APAs (left), and points between 140 and 150 cm from the APAs (right), for tracks associated with coincidences that have a counter differences of 4. The most probable values hit *RMS* 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.

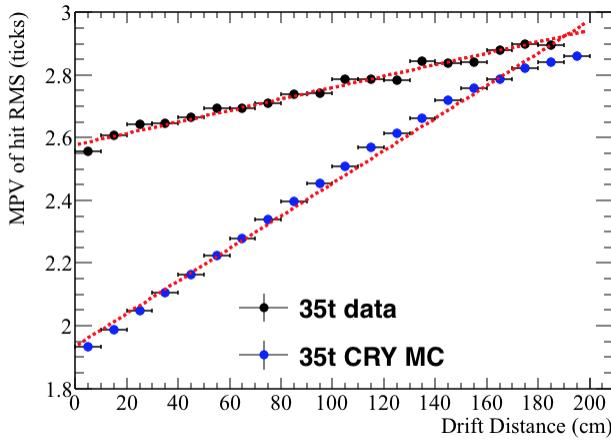


Fig. 5.24 The most probable values of hit RMS as a function of drift distance, for tracks associated with a coincidence that had a counter difference of 4. The distribution for the 35 ton dataset is shown in black, whilst the distribution for the Monte Carlo simulation is shown in blue.

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 prescribed to the increase in hit width caused by the elevated noise 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

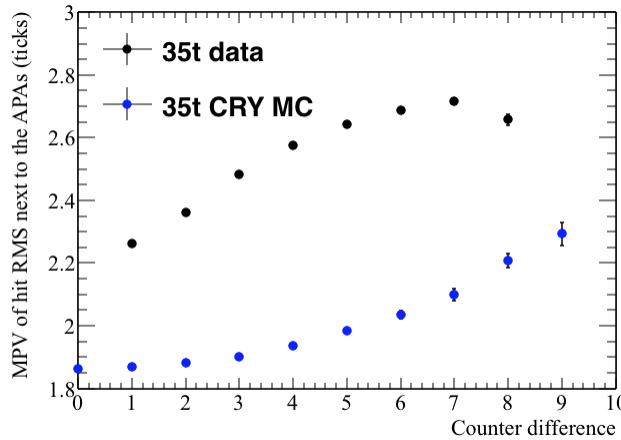


Fig. 5.25 The most probable values of hit *RMS* within 10 cm of the APAs, as a function of the counter difference of the coincidence, that the track was associated with. The distribution for the 35 ton dataset is shown in black, whilst the distribution for the Monte Carlo simulation is shown in blue.

1 case for the 35 ton dataset.

2

3 Figure 5.27 compares how reliably the interaction time, and central *x* position of a track,
4 can be predicted using the effect that diffusion has on the hit *RMS*, in the 35 ton dataset and a
5 low-noise Monte Carlo sample. As was the case when considering the 35 ton dataset, the av-
6 erage time difference (*x* position) is calculated by taking the sum of individual hit differences
7 for every hit in the track, and dividing this by the number of hits in the track. The accuracy
8 in determining the interaction time in Monte Carlo (data) is found to be, 108 (240) μ s, where
9 the distribution has a FWHM of 98 (281) μ s. When this is converted into the difference in
10 central *x* position of the track, the accuracy is found to be, 11.8 (27.7) cm with a FWHM of
11 10.9 (32.1) cm.

12

13 Figure 5.28 compares how reliably the interaction time, and central *x* position, of a
14 track can be predicted using the effect that diffusion has on the hit *RMS/Charge*, in the
15 35 ton dataset and a low-noise Monte Carlo sample. The accuracy in determining the
16 interaction time in Monte Carlo (data) is found to be, 3 (171) μ s, where the distribution
17 has a FWHM of 114 (210) μ s. When this is converted into the difference in central *x* po-
18 sition of the track, the accuracy is found to be, 0.4 (18.5) cm with a FWHM of 12.6 (23.0) cm.

19

20 The hit *RMS/Charge* metric appears to be able to more accurately predict interaction
21 times, as was seen in when considering the only 35 ton dataset. This is again due to the

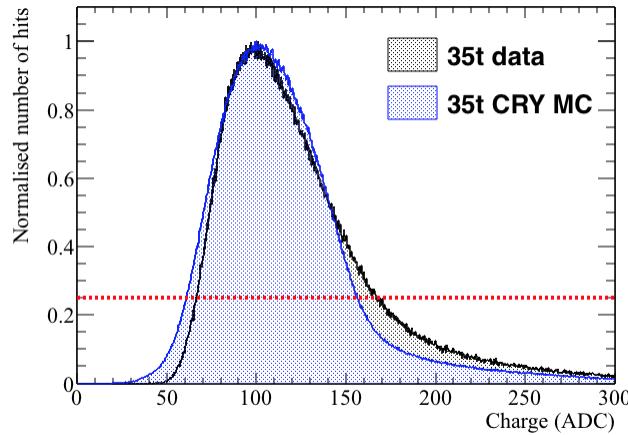
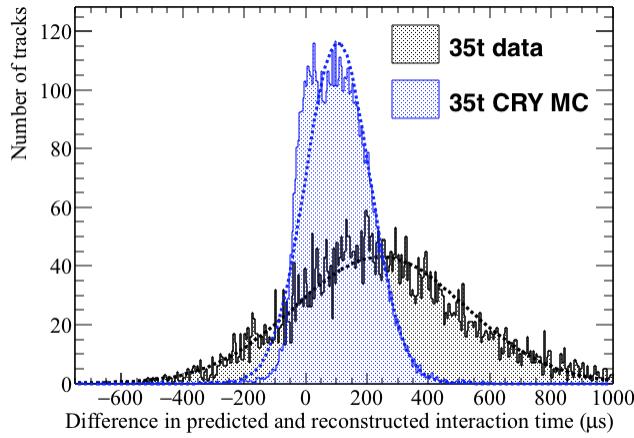


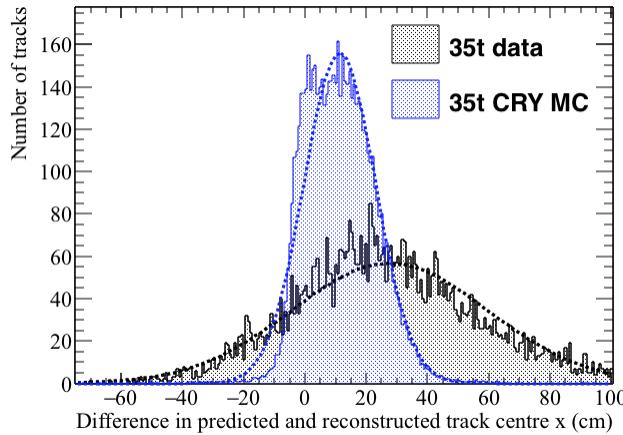
Fig. 5.26 The distribution of normalised hit charge, shown in units of ADC, in the 35 ton dataset and Monte Carlo sample. The number of hits with the most probable hit charge has been normalised to a value of 1. A cut on the normalised number of hits being greater than 0.25 is shown, the aim of this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

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 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

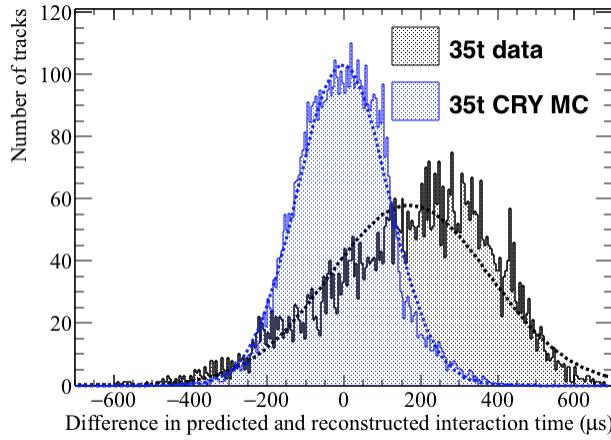


(a) The average difference in interaction times using the hit *RMS* metric.

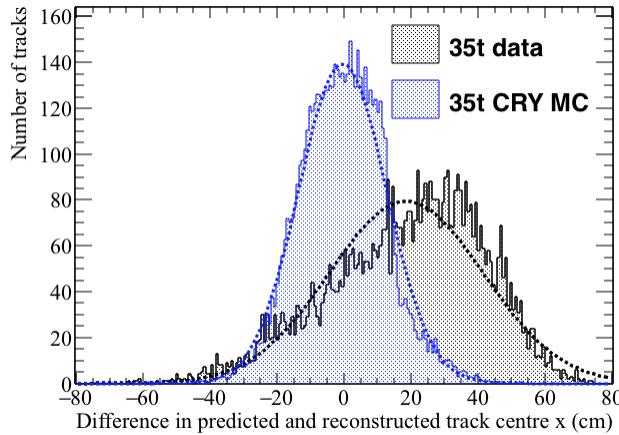


(b) The average difference in the central *x* position of a track using the hit *RMS* metric.

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 shows the accuracy to which interaction times can be determined in μs . Bottom shows 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

5.5 Measuring interaction times using electron diffusion

156

1 on the hit *RMS* distribution, and so this could explain the decreased accuracy using this metric.

2

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

10

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

15

16 **5.5.3 Impact of changing detector properties using Monte Carlo sam- 17 ples**

18 Much has been made of the difficulty that the noise level in the 35 ton dataset introduces,
19 when performing reconstruction and analysis of the data. It is necessary to verify this claim,
20 and so a sample of Monte Carlo events with increasing noise levels is produced, and analysed
21 below. The noise level in the Monte Carlo samples is increased from the low-noise state used
22 in the previous section, to a level more similar to that which is seen in the 35 ton dataset. If
23 the claim that the noise level made reconstruction more difficult is correct, then the accuracy
24 with which the interaction time can be determined should be seen to anti-correlate with the
25 noise level of the simulated detector. In addition to varying noise levels, the electron lifetime,
26 the electric field, and the constant of longitudinal diffusion are varied. All samples have
27 used the same initial muons, this is done so that the only difference between the different
28 samples are the detector conditions. Only one detector condition is varied at a time, so that
29 the effect of each detector condition can be studied in isolation. As only one detector property
30 is changed between samples, there is one sample that is consistent to all sample sets. This
31 is when the RMS of the noise is 2.5 ADCs, the electron lifetime is 3 ms, the electric field
32 is $500 \text{ V}\cdot\text{cm}^{-1}$, and the coefficient of longitudinal diffusion is $6.2 \times 10^9 \text{ cm}^2\cdot\text{ns}^{-1}$. When
33 presenting the studies with changing detector conditions, only the accuracy with which the
34 interaction time and central *x* position of a track can be predicted, is shown here. A more

5.5 Measuring interaction times using electron diffusion

157

robust collection of figures can be seen in Appendix A.

1

2

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.

3

4

5

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.

6

7

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.

8

9

10

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.

11

12

13

14

15

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.

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

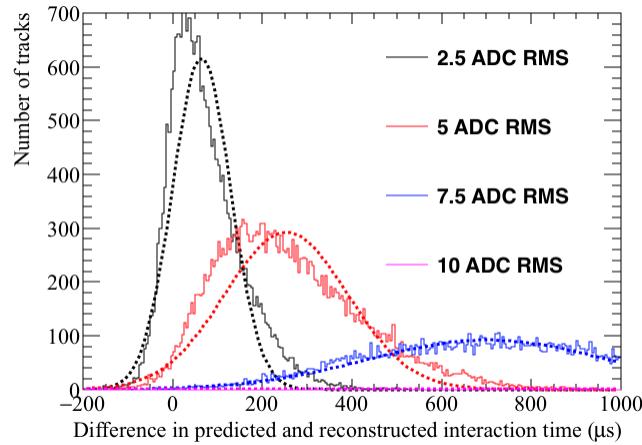
No noise mitigation algorithms have been applied to the increased noise samples shown here. Instead, the threshold that the hit finder uses has been increased to the level that was necessary for a reasonable number of hits to be reconstructed. A reasonable number of hits simply means, not reconstructing such a large number of noise hits that they outweigh the

33

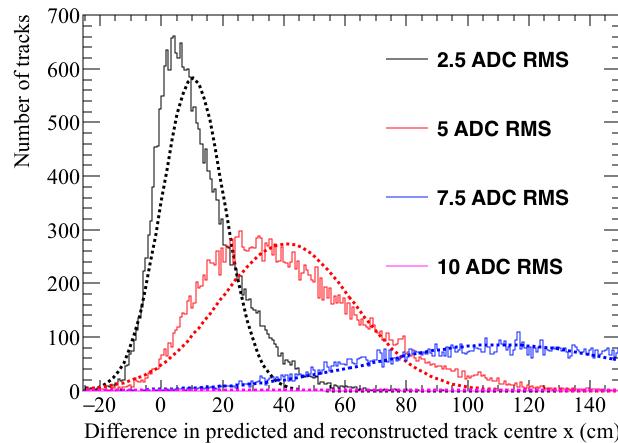
34

35

36

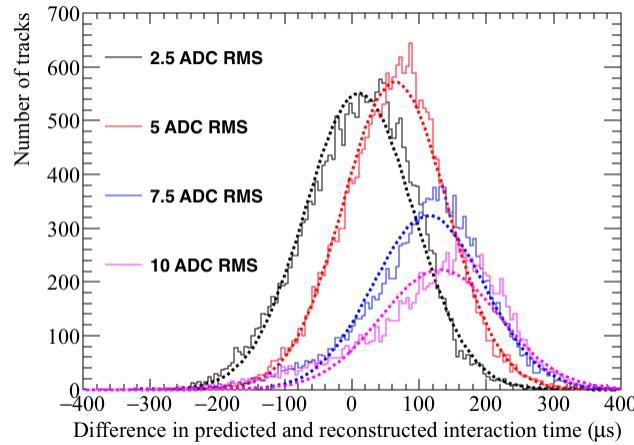


(a) The average difference in interaction times using the hit *RMS* metric.

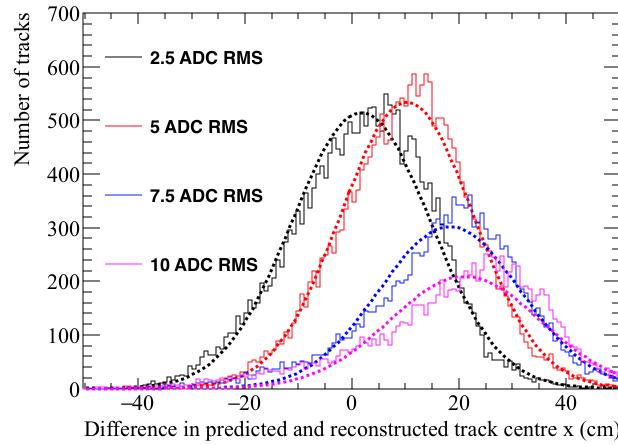


(b) The average difference in the central x position of a track using the hit *RMS* metric.

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. Gaussian functions are fitted to the distributions so that any offset in the predicted times or positions can be discerned.



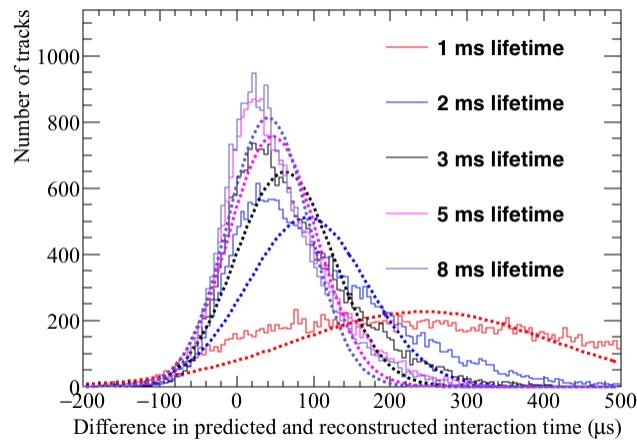
(a) The average difference in interaction times using the hit *RMS/Charge* metric.



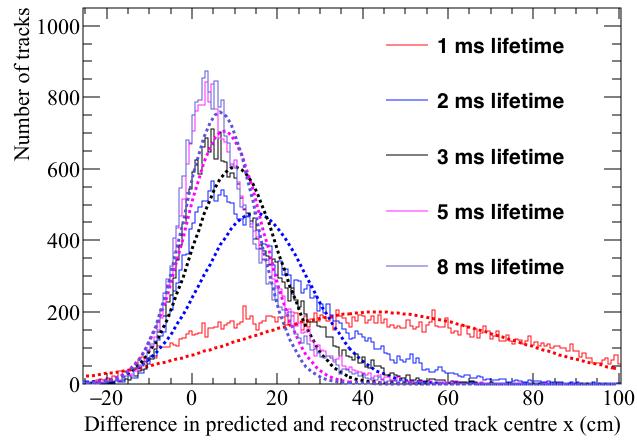
(b) The average difference in the central *x* position of a track using the hit *RMS/Charge* metric.

Fig. 5.30 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. Gaussian functions are fitted to the distributions so that any offset in the predicted times or positions can be discerned.

5.5 Measuring interaction times using electron diffusion

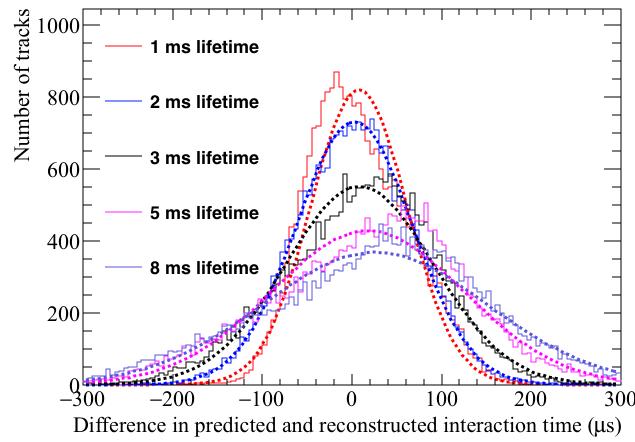


(a) The average difference in interaction times using the hit *RMS* metric.

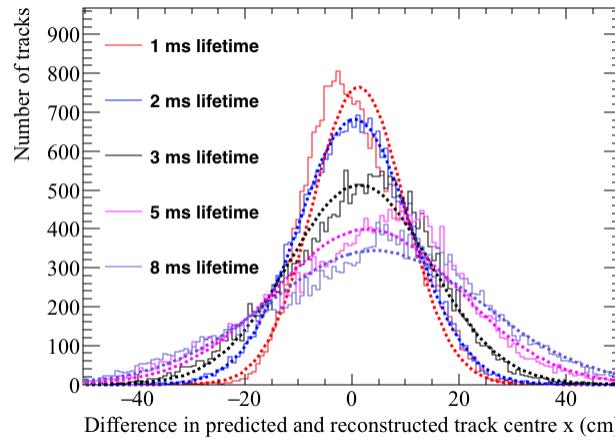


(b) The average difference in the central x position of a track using the hit *RMS* metric.

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. Gaussian functions are fitted to the distributions so that any offset in the predicted times or positions can be discerned.

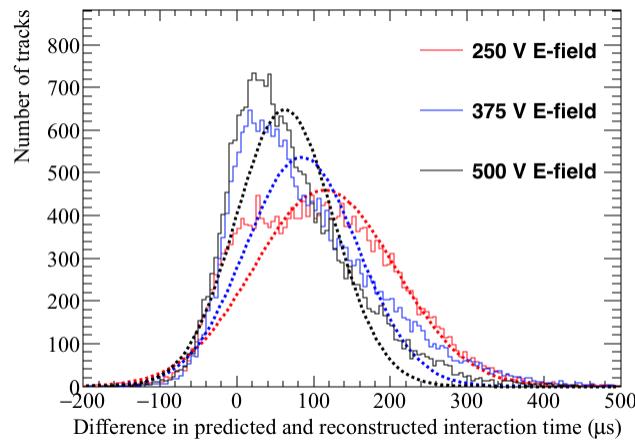


(a) The average difference in interaction times using the hit *RMS/Charge* metric.

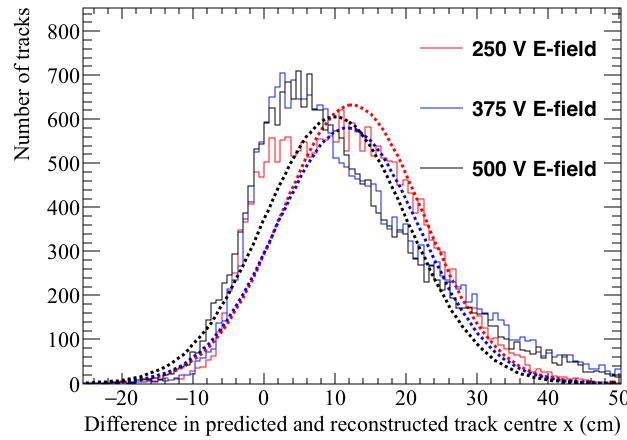


(b) The average difference in the central x position of a track using the hit *RMS/Charge* metric.

Fig. 5.32 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. Gaussian functions are fitted to the distributions so that any offset in the predicted times or positions can be discerned.

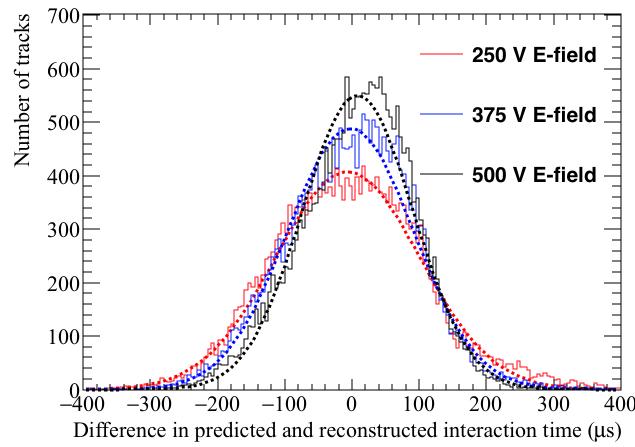


(a) The average difference in interaction times using the hit *RMS* metric.

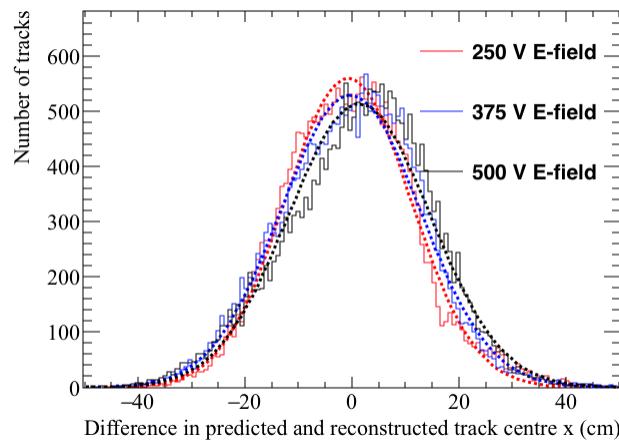


(b) The average difference in the central x position of a track using the hit *RMS* metric.

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. Gaussian functions are fitted to the distributions so that any offset in the predicted times or positions can be discerned.

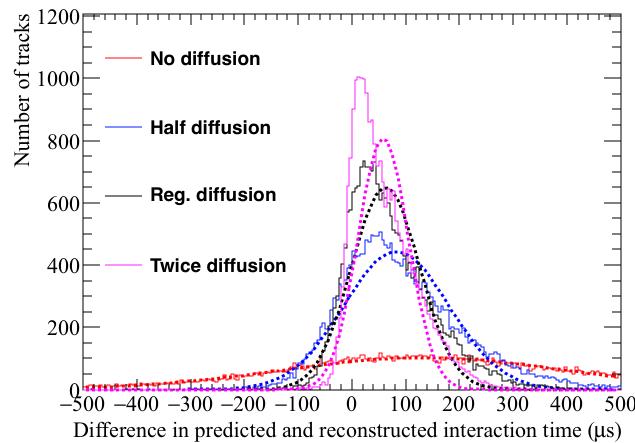


(a) The average difference in interaction times using the hit *RMS/Charge* metric.

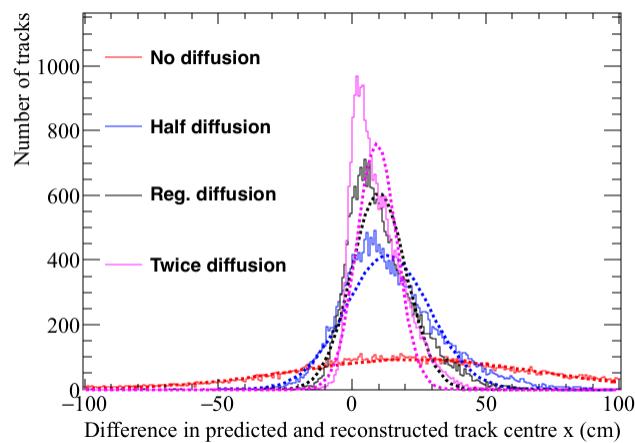


(b) The average difference in the central x position of a track using the hit *RMS/Charge* metric.

Fig. 5.34 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. Gaussian functions are fitted to the distributions so that any offset in the predicted times or positions can be discerned.



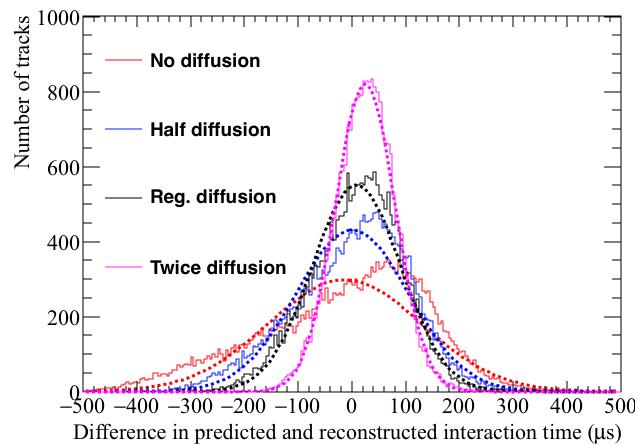
(a) The average difference in interaction times using the hit *RMS* metric.



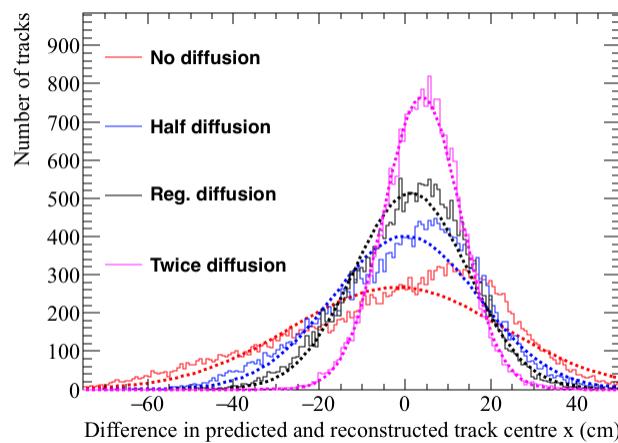
(b) The average difference in the central x position of a track using the hit *RMS* metric.

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. Gaussian functions are fitted to the distributions so that any offset in the predicted times or positions can be discerned.

5.5 Measuring interaction times using electron diffusion



(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. Gaussian functions are fitted to the distributions so that any offset in the predicted times or positions can be discerned.

5.5 Measuring interaction times using electron diffusion

166

1 number of true signals from tracks. The required hit threshold was determined by looking at
2 the deconvoluted signal, and choosing a threshold which was above the majority of the noise
3 signals. The hit thresholds used for each noise level are summarised below:

- 4 • Noise level of 2.5 ADC RMS - hit threshold of 6 ADC
5 • Noise level of 5 ADC RMS - hit threshold of 10 ADC
6 • Noise level of 7.5 ADC RMS - hit threshold of 15 ADC
7 • Noise level of 10 ADC RMS - hit threshold of 20 ADC

8 This means that the main effect of increasing the noise level is to remove the low charge hits,
9 as they will fall below the hit threshold, as it is increased to compensate for the increased
10 noise level.

11
12 When considering Figure 5.29, it can be seen that the accuracy to which interaction times
13 can be determined rapidly decreases as the noise level increases. This is partly due to the fits
14 used to make the prediction metrics not converging for counter differences of 1, 2, 3 and 4,
15 as the MPV of hit *RMS* is not seen to increase for increasing drift distances. For evidence of
16 this, see the Figures in Appendix A. Though this is the extreme case, it can be seen that the
17 validity of the hit *RMS* for increasing drift distances becomes less predictable as the noise
18 is increased. The result of this is a less accurate prediction metric, which leads to the large
19 offsets and widths of the distributions that are shown in Figure 5.29. This is particularly true
20 for the sample which has a noise level of 10 ADC RMS, where the accuracy of the time
21 determination is so bad that it is not contained on the plot.

22
23 The most striking feature of Figure 5.30 is the decrease in statistics seen for the increasing
24 noise levels. This shows the effect that increasing the noise level, and hence hit threshold has.
25 This is because fewer tracks in total are reconstructed, and those that are reconstructed are
26 less likely to meet the criteria about the number of collection hits required to make predictions.

27
28 Figure 5.31 shows that with an electron lifetime of 1 ms, the hit *RMS* metric is very
29 inaccurate, this is likely due to hits which are a large distance away from the APAs being
30 very difficult to reconstruct, because of the extremely poor lifetime. For this reason, the
31 accuracy to which the hit *RMS* metric predicts the interaction time improves as the electron
32 lifetime increases, though this increase is small between the 3 ms, 5 ms and 8 ms samples.
33 Figure 5.32 shows the opposite effect, the accuracy to which the interaction time can be
34 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 position 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 dependant 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

5.5 Measuring interaction times using electron diffusion

168

¹ issues which need to be overcome, some of these are briefly discussed below.

²

³ Many of the figures shown still have slight offsets even though the tails of the hit charge
⁴ distributions have been removed. However, these offsets are generally confined to detector
⁵ conditions which would not be considered optimal, such as very low electron lifetimes (1 ms)
⁶ or high detector noise. The latter is seen to be the case when considering the 35 ton dataset,
⁷ where the high noise scale can be seen to affect the accuracy to which the interaction time,
⁸ and central x position, can be determined. A potential solution to reduce these offsets, and
⁹ also to reduce the width of the distributions, is to perform the interaction time determination
¹⁰ twice. The result of the first run, which is what is shown in this thesis, would then be used
¹¹ to select only hits which lie within the expected regions of hit RMS and hit $RMS/Charge$.
¹² This would be possible as the initial interaction time determination could be used to work out
¹³ the rough x positions of hits, and then only hits which lie within a given region of the MPV
¹⁴ would be used to determine the interaction time from the second pass. When performing the
¹⁵ second pass of interaction time determination, the size of this region would be a user defined
¹⁶ parameter, but would need to be small enough so as to exclude the beginnings of the tails of
¹⁷ the hit RMS , and hit $RMS/Charge$, distributions.

¹⁸

¹⁹ An important improvement to the method would be to expand it to include the induction
²⁰ plane wires, as this will greatly increase both the number of wires which can be used, and
²¹ the range of track angles whose interaction times can be predicted. The angular range of
²² the method would increase, since, when using only collection plane wires, it is impossible
²³ to reconstruct enough hits for nearly vertical muons, as very few wires would be hit. This
²⁴ was discussed in Section 4.3. This was not attempted here, as the electronics noise in the 35
²⁵ ton data was too large to able to reliably reconstruct hits on the induction planes, without
²⁶ reconstructing many noise hits. This meant that the hit threshold on the induction planes was
²⁷ very high.

²⁸

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 [153], 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 [153]. The other work presented was performed for the DUNE collaboration, in conjunction with work done by Vitaly Kudryavtsev and Matthew Robinson, both of the University of Sheffield. This work was performed with the aim of producing muon-induced background limits to nucleon decay in the DUNE FD.

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 [154]. 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

6.1 Simulations of the LBNE surface detector

170

¹ review of the simulations, see [153].

²

³ Simulations have been built using GEANT4 versions 9.4 and 9.6. Initial simulations
⁴ were performed using version 9.4, before being rebuilt to use version 9.6. A study showing
⁵ that the background rate did not change with the newer version of GEANT4 was performed,
⁶ in order to verify that the two sets of simulations were consistent. In all simulations the
⁷ physics list “Shielding” was used [155]. Shielding allows particles to undergo Compton
⁸ scattering, inelastic scattering with nuclei and nucleons, and includes all electromagnetic
⁹ (EM) processes for all charged particles.

¹⁰

¹¹ The authors work built on initial simulations which were performed using a simplified
¹² geometry, referred to as the *simple* geometry, which consisted of a single box of LAr
¹³ measuring $30 \times 15 \times 16 \text{ m}^3$. This detector was enclosed in rock measuring $(5 \times 10^3) \times$
¹⁴ $(5 \times 10^3) \times 22 \text{ m}^3$, in the x , y , and z coordinates respectively. The detector was positioned
¹⁵ in the rock so as to have a 3 m overburden of rock. As simulations were not performed in
¹⁶ LArSoft, the co-ordinate system used was defined as follows;

¹⁷ • x - parallel to the beam direction.

¹⁸ • y - perpendicular to the beam direction.

¹⁹ • z - vertical direction.

²⁰ • The co-ordinate system is centred on the middle of the detector volume in x and y , and
²¹ on the surface of the rock which was above the detector in z .

²² The results of these initial simulations, discussed in detail in [153], are not shown here, but
²³ will be summarised in Table 6.2.

²⁴

²⁵ The improvements which the author made to these initial simulations were two fold.
²⁶ Firstly a more detailed geometry, referred to as the *complex* geometry, was included, this
²⁷ had a much more realistic detector design [156]. The *complex* geometry had two identical
²⁸ cryostats, each containing 120 TPC cells measuring $2.52 \times 2.28 \times 7.00 \text{ m}^3$. These cells
²⁹ each contain an active volume of LAr measuring $2.27 \times 2.25 \times 6.30 \text{ m}^3$. The TPC cells
³⁰ are arranged such that are 10, 6 and 2 TPC cells in the x , y and z directions, respectively.
³¹ This gives a total volume of LAr in each cryostat measuring $28.20 \times 13.95 \times 15.0 \text{ m}^3$, with
³² a mass of 5.35 kton. This gives a total fiducial mass of 10.7 kton of active LAr in the two
³³ cryostats. Running vertically between the TPC cells are anode plane assemblies (APAs) and
³⁴ cathode plane assemblies (CPA’s), all of which are embedded within the larger blocks of LAr.

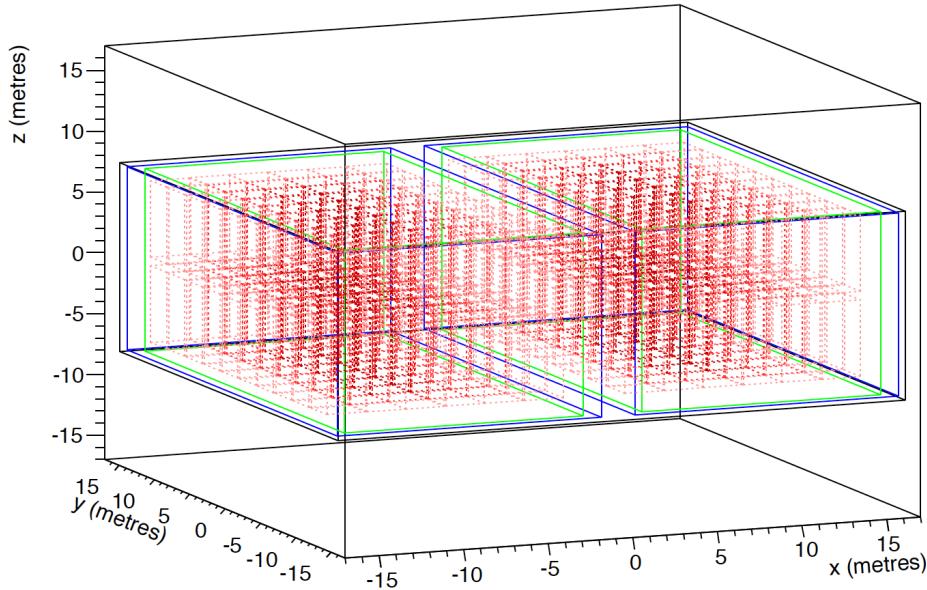


Fig. 6.1 The *complex* detector geometry used in the LBNE surface detector simulations. The TPC cells are shown in red, and are orientated such that there are 10, 6, and 2 TPCs in the x , y , z coordinates of each cryostat respectively. The steel walls of the cryostats are green, whilst the outer edges of polyurethane are blue, and the concrete enclosures are black. Figure taken from [153].

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 20×20 km² is used to generate the accurate surface profile. The map is sampled in bins measuring 5×5 m². To include the effects that the surrounding hills would have on the muon flux, muons are initially sampled 600 m above the surface, and are then stepped through the surface profile, until they reach a box measuring $80 \times 80 \times 36$ m³ 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

6.1 Simulations of the LBNE surface detector

172

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

5

6 It should be emphasised that all work presented here has been done using Monte Carlo
7 truth, and so no reconstruction has performed. It also means that at all times only Monte
8 Carlo truth information is used. Therefore, any positions or trajectories which are referred
9 to are those which are recorded by GEANT4 [104]. This information saved by GEANT4 is
10 “smeared” before analysis is performed, in attempt to take into account of some of the detector
11 effects which were not simulated. This smearing is particle type and energy dependant, and
12 is performed on the particles position, trajectory, and energy [153, 157].

13

14 **6.1.1 Classifying signal and background events**

15 Before the simulations can be described in detail, it is necessary to classify what is classified
16 as signal and background event. In the LBNE and DUNE FD, a ν_e appearance signal event
17 would occur when a ν_e undergoes inelastic scattering with either an electron or nucleon. This
18 interaction will produce an electron, which will in turn produce an EM shower. It is this EM
19 shower which will be identified as a neutrino appearance signal. The electron track which is
20 produced in these interactions will not necessarily be isolated, and may be accompanied by
21 hadronic debris.

22

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

34

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, have 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 [153].

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 signal events.

As charged particles which enter 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 Point of Closest Approach (PoCA), 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 signal events, and also has a lower limit of 2 cm, meaning that any photons which have PoCAs 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 would have a charged particle very close to it, meaning that were this lower limit not in place, it 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 [158]. This means that it is possible to use the angle of a shower with respect to the beam, to distinguish between signal and backgrounds 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 [159]. An example of how the PoCA, and angle of a

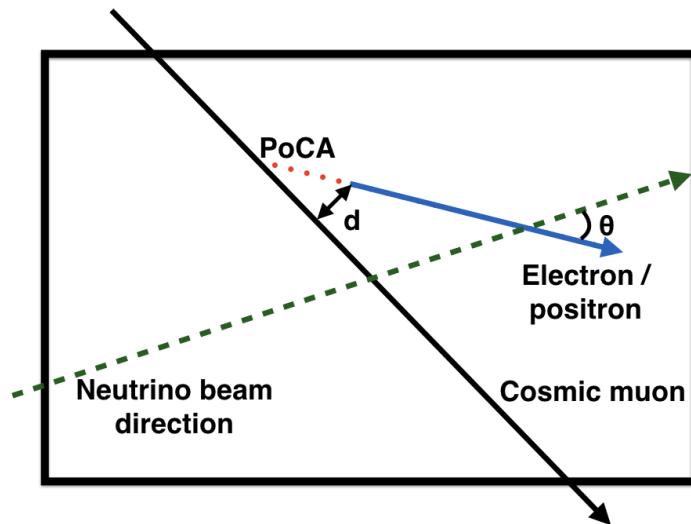


Fig. 6.2 A diagram of the PoCA, and the angle w.r.t the beam calculations in the LBNE surface detector simulations. A cosmic muon which passes through the detector is shown as a black line, whilst an electron/positron that is produced is shown as a blue line. The minimum distance between the muon and 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 [153].

¹ photon with respect to the beam axis are calculated is shown in Figure 6.2.

²

³ It is envisioned that not all of the instrumented LAr will be able to used to identify
⁴ neutrino events. This is because a fiducial cut will be applied around the detector edges, in
⁵ order to ensure that the entire track is contained within the active volume of the detector.
⁶ For the *complex* detector, this fiducial cut is only applied to the outward facing edges of the
⁷ active volumes. This is because it is assumed that tracks passing between cells, within a
⁸ given cryostat, will be able to be stitched together, as is done in LArSoft.

⁹

¹⁰ Additionally, signal events should be able to be distinguished from background events
¹¹ based on the energy deposition measurements. These will come from identifying the start of
¹² an EM shower as coming from either an electron, or a photon, in the case of a signal, and
¹³ background event, respectively. Studies have shown that the failure rate of this separation
¹⁴ tails off at 10% for showers with energies above 0.5 GeV [160]. Therefore, a flat reduction
¹⁵ of 1/10 is applied to all surviving simulated background events.

¹⁶

Finally, it is envisioned that an efficient photon detection system will be used, which uses the scintillation light emitted by excited argon [161]. 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 [153], and so they are not discussed here.

The muons used in the simulations are generated using a Gaisser's parameterisation [162], which has been modified for large zenith angles, and muon decay [163]. 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 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 $\frac{1030 \times 2.06 \times 10^{-3}}{\cos\theta - 120}$, R_c is the ratio of prompt muons to pions, and p_d is the probability for a muon to not decay in the atmosphere [164]. 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 $A = 0.1$ and $\gamma = 2.7$ are used [162]. When considering larger depths, as will be done in Sections 6.2 and 6.3 where MUSUN [110, 111] is used, typical values are $A = 1.84$ and $\gamma = 2.77$.

The protons and neutrons used in the simulations are generated using CRY [107, 108], 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 se-

6.1 Simulations of the LBNE surface detector

176

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 [153].

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

1 rious concern. The particles are generated at an altitude of 2100 m above sea level. However,
 2 the LBNE surface detector would have had an altitude of 1505 m above sea level, and so the
 3 particle fluxes generated by CRY have to be corrected to the particle fluxes expected at 1505
 4 m above sea level [153]. The fluxes generated by CRY are also subject to a further correction,
 5 as CRY underestimates the cosmic ray flux by as much as 70% [165]. This is because CRY
 6 only considers protons striking the Earth's atmosphere. As the muon flux is calculated at
 7 sea level, this flux also has to be corrected so as to be the flux at 1505 m above sea level [153].
 8

9 To reduce simulation time, only particles which will cause background events are simu-
 10 lated. For example, it is found that 92.7% of the proton induced background events are due
 11 to protons with energies of 10 GeV or more, yet these particles represent only 0.76% of the
 12 total proton flux. This means that by not simulating protons with energies below 10 GeV, the
 13 vast majority of background events will be simulated, but this will require less than 1% of
 14 the simulation time. The background events which are not simulated, can be accounted for
 15 by correcting the background seen from simulations, by the proportion of background events
 16 which were not simulated [153]. The minimum energies of the simulated muons, protons,
 17 and neutrons, along with the percentage of background events which they cause, and the
 18 percentage of the particle fluxes above this energy, are shown in Table 6.1.

19
 20 From Table 6.1 it can be seen that in the case of only simulating protons above 10 GeV,
 21 7.3% of the background were not simulated, and so the background rate from simulations
 22 should be scaled by a factor of 1.0787 ($\frac{1}{0.927}$).
 23

Table 6.2 The normalised annual background rate 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 [153].

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

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 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 where the annual background when using the *complex* geometry and simple surface profile are shown in [153]. However, as discussed in [153], these background rates were found to be consistent with the background rates for the *complex* geometry and accurate

Table 6.3 The normalised annual background rate, 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 dependant 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, unless the simulated annual background rate drops to 0, in which case an upper limit at 90% confidence level [166] 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 annual background rate. Table is taken from [153].

	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
$Ext \rightarrow \gamma$	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
$Other \rightarrow \gamma$	7.87×10^4	333 ± 68	97 ± 37	0 – 15.20	0 – 11.43	0 – 0.11	0 – 0.002

1 surface profile.

2

3 Table 6.3 shows the background rate for muons which enter the active volume of detector,
4 as sequential cuts are applied, for the *complex* geometry, and accurate surface profile. A total
5 of 2×10^8 muons with energies greater than 10 GeV are generated, representing 0.1003
6 years worth of detector live time. The overall number of background mimicking events is
7 seen to increase when using the *complex* detector geometry, and the accurate surface profile,
8 as opposed to using the *simple* detector geometry, and simple surface profile. It is observed
9 that the expected annual background rates for the *complex* detector geometry in the flat,
10 and accurate surface profiles, are consistent. This means that including the accurate surface
11 profile does not have a significant effect on the background rate.

12

13 The effectiveness of the PoCA cut with respect to the initial muon is obvious, as the
14 annual background event rate is reduced by over 99%. The rejection rate is observed to be
15 100% when considering the $\mu \rightarrow \gamma$ photon ancestry. The rejection rate is also very high in
16 the $Ext \rightarrow \gamma$ photon ancestry where it approaches 100%, this is because when the photons
17 trajectories are extrapolated backwards, they are seen to be very close to the muon track
18 outside the detector. The remaining background is dominated by photons with π^0 parents.
19 Around half of the surviving events are removed by the application of the angular cut, but

6.1 Simulations of the LBNE surface detector

Table 6.4 The normalised annual background rate, 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 [153].

	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
$Ext \rightarrow \gamma$	16500 ± 500	2858 ± 199	998 ± 118	748 ± 102	75 ± 10	0.53 ± 0.07
$Other \rightarrow \gamma$	28 ± 5	28 ± 5	28 ± 5	$0 - 20$	$0 - 2.01$	$0 - 0.01$

more than 80% of these events are removed by the application of the all charged PoCA cut. This again shows how powerful the use of the PoCA cut is when rejecting background events. Whilst the fiducial cut 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, and the use of an efficient photon detection system 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 muons used in Table 6.3 are used here, meaning that a detector live time of 0.1003 years has been simulated. The overall number of background mimicking events is seen to increase substantially when using the *complex* detector geometry, and the accurate surface profile, as opposed to using the *simple* detector geometry, and simple surface profile. This is to be expected, as there is a much larger external surface area, and the vertical gaps produce internal gaps between the TPCs. It is observed that the expected annual background rates for the *complex* detector geometry in the flat, and accurate surface profiles, are consistent, though the total number of background events is less when the accurate surface profile is used.

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 pass through gaps between TPCs, and produce secondaries which are far away from 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.

6.1 Simulations of the LBNE surface detector

180

Table 6.5 The normalised annual background rate, 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 [153].

	E_γ	$D > 30$ cm	$\theta_{beam}(E)$	PoCA _{all}	$e - \gamma(E)$	γ detection
<i>Total</i>	1.55×10^4	$10559 \pm$	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
$Ext \rightarrow \gamma$	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	$0 - 1.04$	$0 - 0.10$	$0 - 0.001$

¹ 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.

³

⁴ Table 6.5 shows the background rate for protons, as sequential cuts are applied, for the ⁵ *complex* geometry, and accurate surface profile. A total of 1×10^7 protons with energies ⁶ greater than 10 GeV are generated, representing 2.482 years worth of detector live time.

⁷ As the protons are generated on a plane which measures 50×50 m², and the accurate ⁸ surface profile was modified to have a flat area above the detector measuring 100×100 m², ⁹ it is not necessary to propagate protons through the accurate surface profile. As discussed ¹⁰ earlier, protons are generated on a plane measuring 50×50 m² as protons will not induce ¹¹ background events at the high inclinations which muons do. The overall number of back- ¹² ground mimicking events is seen to decrease substantially when using the *complex* detector ¹³ geometry, as opposed to using the *simple* detector geometry.

¹⁴

¹⁵ It can be seen from Table 6.5 that the cut with respect to the beam angle, and the PoCA ¹⁶ calculation with respect to all charged particles, are very effective as they remove 67% and ¹⁷ 90% of all remaining background events respectively. That the all charged PoCA cut is so ¹⁸ effective is not surprising, as many of the photons produced will be close to the initial proton ¹⁹ that was simulated. After all cuts are applied, the annual number of background events is ²⁰ seen to decrease by over a factor of 10 when using the *complex* geometry. This decrease is ²¹ attributed to the additional shielding which is present in the *complex* geometry, as concrete, ²² insulation and inactive regions of LAr have been added.

²³

²⁴ Table 6.6 shows the background rate for neutrons, as sequential cuts are applied, for the ²⁵ *complex* geometry, and accurate surface profile. A total of 1.1×10^8 neutrons with energies ²⁶ greater than 1 GeV are generated, representing 0.653 years worth of detector live time. The ²⁷ neutrons are generated on a plane which measures 50×50 m², and so as seen with the

Table 6.6 The normalised annual background rate, 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 [153].

	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
$Ext \rightarrow \gamma$	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	$0 - 3.68$	$0 - 0.37$	$0 - 0.003$

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.

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 increase in surface area of the detector is found to cause the number of background events caused by muons which both enter and miss the active volume of the detector 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.

Table 6.7 The normalised annual background rate 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 [153].

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

¹ Following the simulation of large samples of muons, protons, and neutrons, the expected
² background for the LBNE surface detector is found when considering a *complex* detector
³ geometry, and an accurate surface profile. This is compared to the expected background rate
⁴ when using a *simple* detector geometry and flat surface profile [153]. It is observed that the
⁵ effect of the *complex* detector geometry is to reduce the overall background rate, primarily
⁶ due to the additional shielding which it provides. This additional shielding causes the proton
⁷ and neutron induced backgrounds to decrease substantially, as was shown in Table 6.7. It
⁸ is however found that there will be a significant source of backgrounds from muons which
⁹ do not enter the active volume of the detector. This is due to the presence of vertical gaps
¹⁰ between the TPC cells. The effect of incorporating the accurate surface profile is found
¹¹ to be negligible, as the hills only offer minimal amounts of shielding. These simulations
¹² provide a relatively accurate estimate of the expected background for the LBNE surface
¹³ design, although it must be stressed that these studies have been performed using Monte
¹⁴ Carlo truth information, and so have not used reconstruction.

¹⁵

¹⁶ 6.2 The use of MUSUN in LArSoft

¹⁷ The primary muons in the following discussions are all generated using MUSIC [110, 112,
¹⁸ 113] and MUSUN [110, 111], and so a brief overview of them is required. MUSIC first
¹⁹ propagates muons through a medium, defined by the user, for given initial energies. A range
²⁰ of energies between 10^2 GeV and 10^7 GeV are considered, and their energy distributions are

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

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.

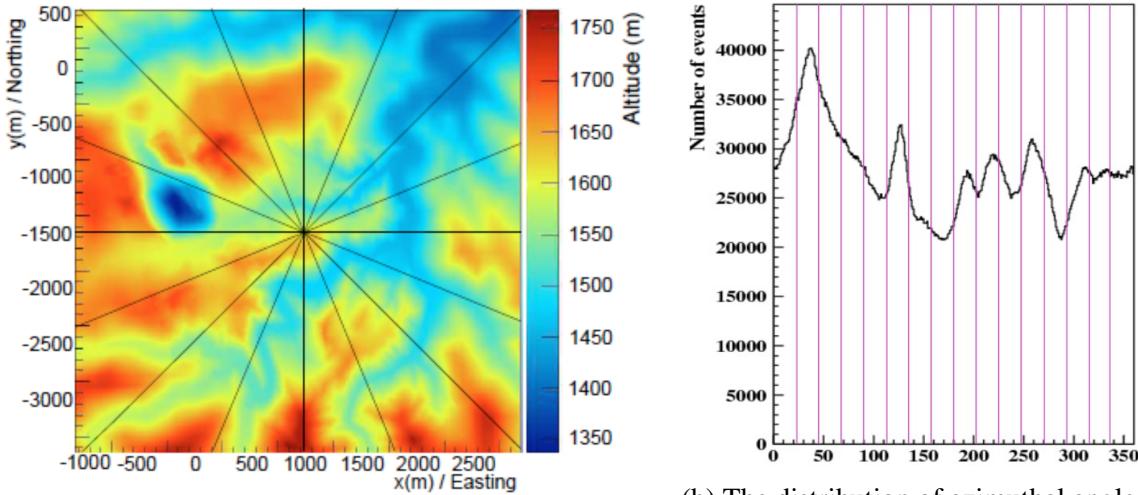
The location of the DUNE far detector, near the Ross shaft at SURF, has global coordinates of $44^\circ 20' 45.21''$ North, $103^\circ 45' 16.13''$ West. The rock composition is assumed to be, $<Z> = 12.09$ and $<A> = 24.17$. The density is assumed to be $2.70 \text{ g}\cdot\text{cm}^{-3}$ [167], though some measurements suggest that it may closer to $2.90 \text{ g}\cdot\text{cm}^{-3}$ [168, 169]. The vertical flux calculated by MUSIC/MUSUN 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 $(5.38 \pm 0.07) \times 10^{-9} \text{ cm}^{-2}\cdot\text{s}^{-1}\cdot\text{sr}^{-1}$ [170]. Given the small differences in these values, and another measurement by the Majorana demonstrator [171], the systematic uncertainty in calculating the muon flux is estimated to be 20% [172].

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.

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

6.2 The use of MUSUN in LArSoft



(a) The surface profile of the DUNE far detector site at SURF [173].

(b) The distribution of azimuthal angles of muons at the DUNE far detector site at SURF.

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.

1 in these cascades which enter the detector, in the absence of the initial muon, are of particular
 2 interest, as some of them could be mistaken for nucleon decay events. The study of these
 3 nucleon decay mimicking events is discussed in Section 6.3. The size of the box which the
 4 muons are sampled from is $74.43 \times 29.54 \times 30.18 \text{ m}^3$, compared to the simulated cryostat
 5 which has dimensions, $61.62 \times 14.94 \times 13.58 \text{ m}^3$. The dimensions are given as length
 6 \times width \times height, using the LArSoft coordinate system which was defined at the start of
 7 Section 2.5. The muons are sampled according to their energy spectrum, for a given zenith
 8 and azimuthal angle, using the angular distribution obtained with MUSUN.
 9

10 However, before this could be done MUSUN had to be incorporated into the DUNE
 11 software framework, as it had previously been maintained in FORTRAN as an external
 12 package. Before simulations in LArSoft were performed, it was ensured that the muon
 13 distributions produced by the ported LArSoft code were identical to the original distribu-
 14 tions produced by the FORTRAN code. The distributions produced by the DUNE software
 15 framework are shown in Figure 6.4, and are consistent with the distributions made for the
 16 LBNE collaboration [174]. Figure 6.4 shows the initial positions of 10,000 muons around the
 17 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 $14.5 \times 12 \times 58$ m³, 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, this causes the fiducial mass to be reduced to 13.8 kt.

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 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.

The DUNE experiment will run for more than 20 years, and 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

6.3 Nucleon decay channels in DUNE

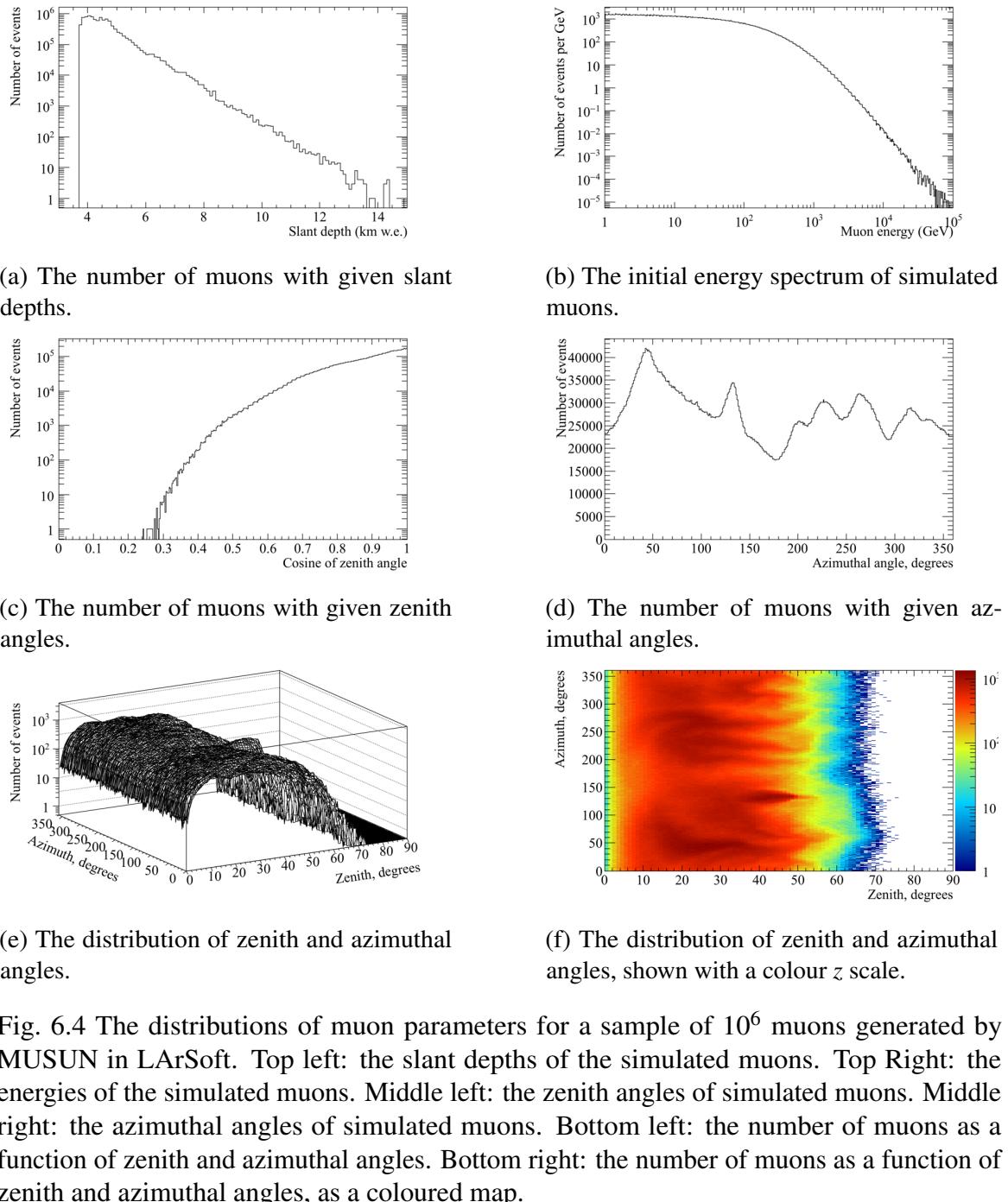


Fig. 6.4 The distributions of muon parameters for a sample of 10^6 muons generated by MUSUN in LArSoft. Top left: the slant depths of the simulated muons. Top Right: the energies of the simulated muons. Middle left: the zenith angles of simulated muons. Middle right: the azimuthal angles of simulated muons. Bottom left: the number of muons as a function of zenith and azimuthal angles. Bottom right: the number of muons as a coloured map.

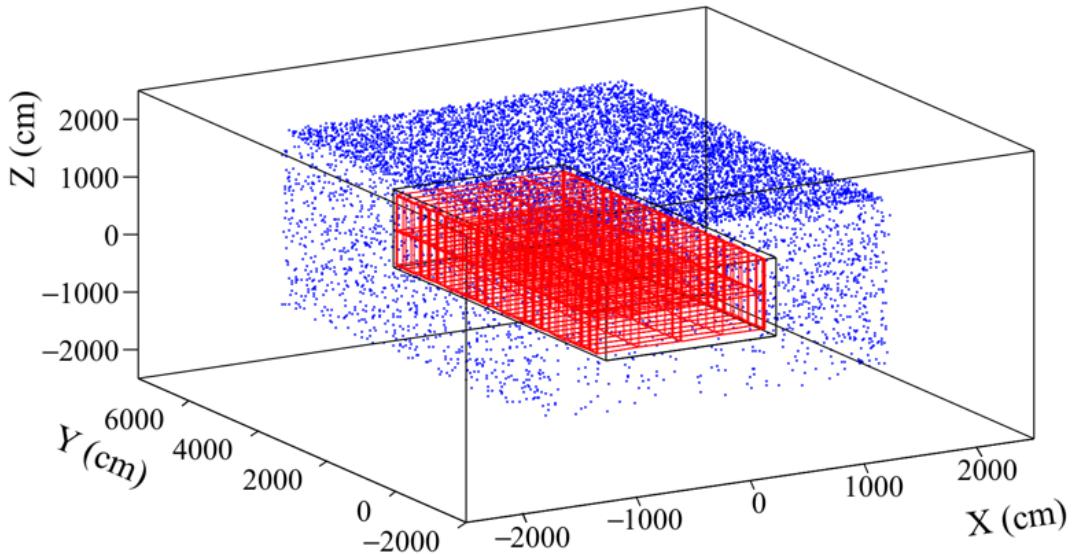


Fig. 6.5 The initial positions of muons generated by MUSUN around a DUNE 10 kt module. The initial positions of the muons are shown as blue points, whilst the cryostat is a single black box and each TPC is a single red box.

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 [104], where events satisfying one of the following cuts are discarded;

- Contain 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

¹ tracks of more than 1 m is clearly less than this. This fraction of dead time is assumed to
² be acceptable. Filtering out events where there are no energy depositions in the detector is
³ clearly acceptable, as there are no energy depositions which could mimic a nucleon decay
⁴ signature.

⁵

⁶ After applying this series of cuts, 98.12% of the initial muon sample are removed, mean-
⁷ ing that the initial sample of 2×10^9 muons is reduced to around 4×10^7 muons. This is a
⁸ much more reasonable sample size to store on disk, and to perform analyses on. It is upon
⁹ this reduced sample of muons that the cosmogenic background analyses are performed. As
¹⁰ discussed in Section 2.3.2, the proton decay channel of $p \rightarrow K^+ + \bar{v}_e$ is referred to as the
¹¹ 'Golden Channel' in LAr, this analysis is discussed in [172]. The related decay of a neutron
¹² in the decay $n \rightarrow K^+ + e^-$ is discussed here. The theoretical motivation for this channel was
¹³ briefly discussed in Section 1.2.

¹⁴

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

¹⁶ As shown in Table 2.3, the predicted sensitivity that DUNE will have to this channel is much
¹⁷ better than that of Super-K. As a result, it is an interesting decay mode to study. As discussed
¹⁸ in Section 1.2.2, the cosmogenic background to nucleon decay is predominantly caused
¹⁹ by neutral particles, such as a K^0 , entering the detector volume, and interacting relatively
²⁰ far from the detector edges. This is particularly true for the 'Golden Channel,' as shown
²¹ in Figure 1.2, but it also holds for other channels. Events like this are the main cause for
²² concern when eliminating all cosmogenic backgrounds. As mentioned in Section 1.2.2, it is
²³ difficult to separate a K^+ from a K^- in a LArTPC, and so any charged kaon is considered a
²⁴ background in the analysis which is presented in this thesis.

²⁵

²⁶ The analysis presented in this thesis has been performed on Monte Carlo truth informa-
²⁷ tion, and so does not contain any reconstructed quantities. Studies involving hit and track
²⁸ reconstruction are in progress [175], though they will not be discussed here. As Monte Carlo
²⁹ truth information has been used, perfect particle identification (PID) has been assumed. There
³⁰ was also no smearing of the energies, locations, or trajectories of any simulated particles,
³¹ meaning that it is assumed that all deposited charge will be reconstructed, and that the
³² detector characterisation is perfect. This is something which will need to be refined in future
³³ analyses, and will be taken into account when the analysis progresses to use reconstructed
³⁴ quantities, as discussed in Section 6.3.5.

³⁵

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 due to kaons and due to electrons.
- 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 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.

¹ – The energy deposited in the detector which does not fit any of the above criteria.

² When performing the analysis it is important to be able to trace the particle ancestry.
³ This is so that energy depositions in the detector can be properly assigned to the relevant
⁴ particles, as cuts are applied. For example, a μ^+ is often produced when a K^+ decays at
⁵ rest, and this muon may travel more than 20 cm. However, the cut on muon track length
⁶ should not be applied to this muon as it was produced by the decay of the kaon. Similarly,
⁷ as the kaon interacts in the detector secondary particles will be produced, which will be
⁸ reconstructed as tracks coming off the main kaon track. The initial kinetic energy of the
⁹ kaon can only be determined by summing the energy depositions due to these secondary
¹⁰ particles, and the energy depositions due to the kaon itself. Correctly calculating the initial
¹¹ kaon kinetic energy is critical when determining if an event is a nucleon decay event. The
¹² reason for this is that nucleon decay events have very specific energy spectra, and so being
¹³ able to correctly assign the ancestry of energy depositions is vitally important. The same
¹⁴ is true for the electron energy, which is calculated by tracing the ancestry of the particles
¹⁵ produced in the EM shower which it produces back to the electron.

¹⁶

¹⁷ As no reconstruction has been performed, the tracks referred to here are different from
¹⁸ those in Chapters 4 and 5. The definition of a track used here, is that the particle in question
¹⁹ has energy depositions, on simulated wires, which are directly associated with it. These
²⁰ simulated wires are not the same as the wires which have been considered in Chapter 4,
²¹ as the simulated signals have not been digitised. This distinction is important, as it allows
²² the energy depositions directly from GEANT4 to be used, whilst also allowing for LArSoft
²³ methods concerning whether depositions are within TPC boundaries to be utilised.

²⁴

²⁵ The simulated electrons may begin showering immediately, or they may produce a short
²⁶ “track like” segment before beginning to shower. To ensure that every electron shower can be
²⁷ identified, electrons are not required to produce a short “track like” segment. This means that
²⁸ all electrons are assumed to begin showering immediately, and it is also assumed that all of
²⁹ the energy in the shower can be identified as coming from a single electron. This definition
³⁰ of shower energy is the one that is used when the showering algorithms are developed in
³¹ LArSoft. It is hoped that when DUNE begins taking data the showering algorithms will be
³² able to achieve this level of energy reconstruction.

³³

³⁴ When calculating the distance between the start of the kaon track, and the start of the
³⁵ electron track/shower, the energy depositions whose locations are closest to the Monte Carlo
³⁶ truth start points of the particles are used. For particles which are produced within the active

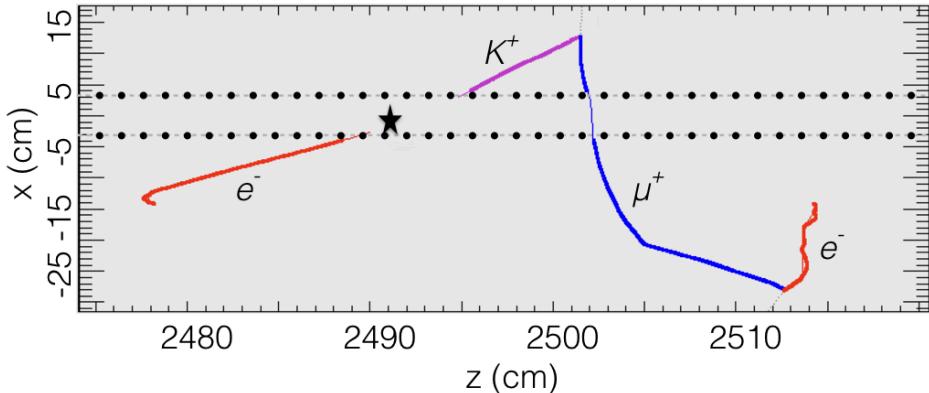


Fig. 6.6 A simulated $n \rightarrow K^+ + e^-$ decay which occurred in a gap between TPCs. The path of the kaon produced in the decay is shown as a purple line. The path of the muon, produced by the decay of the kaon, is shown as a blue line. The paths which the electrons in the event took are shown 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.

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 to be 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 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 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 de-

¹ positions to be collected over approximately 10 collection wires. This is assumed to be
² unlikely during data taking, and cannot happen in the simulations considered here, as Monte
³ Carlo truth information is used. As shown by Figure 6.6, however, it is possible for the
⁴ kaon and electron produced in a signal event to be separated by more than 5 cm. To prevent
⁵ events such as this being missed, a second criteria is applied to events with large separations.
⁶ This criteria is that the “Point of Closest Approach” (PoCA) between the two particles,
⁷ found by extrapolating the kaon track and the electron shower, forwards and backwards,
⁸ is less than 2 cm. This is the same PoCA calculation which was made in Section 6.1.2,
⁹ and it means that events such as the one shown in Figure 6.6 are still identified as signal events.

¹⁰

¹¹ The fiducial cut is only applied to the outer edges of the cryostat, as if it were done
¹² with respect to the edge of every TPC in the far detector the loss of volume would be non-
¹³ negligible. This means that the event shown in Figure 6.6 would not fail the fiducial cut, as
¹⁴ the decay occurred over 6 m away from the edge of the detector, but happened to be in a gap
¹⁵ between two TPCs. The need for a fiducial cut is two fold, firstly the vast majority of cosmi-
¹⁶ cally induced events in the detector will have a charged particle which enters the detector.
¹⁷ Performing a fiducial cut will remove all of these events, and will then mean that the only
¹⁸ cosmic background events which can mimic a signal event would involve either a significant
¹⁹ amount of charge being missed, or a neutral particle entering the detector, and interacting
²⁰ relatively far from the detector walls. Secondly, in order to calculate the kinetic energies of
²¹ the particles produced in the nucleon decay, and also to perform particle identification, they
²² must be fully contained within the detector. As such, if one of the particles produced in the
²³ nucleon decay escapes the detector then its kinetic energy cannot be determined accurately,
²⁴ and if it is the kaon, or its decay products, then the particle cannot be identified using the
²⁵ method discussed in Section 4.4. This will also affect any particles which stop in the gaps
²⁶ between TPCs, as the end point will not be reconstructed, though PID may still be possible
²⁷ as the end point can be estimated. A fiducial cut of 2 cm is used, as the loss of active volume
²⁸ which this causes is negligible in the DUNE FD. A 2 cm fiducial cut also ensures that a
²⁹ significant amount of charge would have to be missed for a particle which enters/escapes the
³⁰ detector to be incorrectly identified as being contained within the detector.

³¹

³² Once both the ancestry of energy depositions in the simulation, and the initial kinetic
³³ energies, have been correctly accounted for and calculated, it is possible to observe the distri-
³⁴ bution of background events, as the cuts outlined above are applied. The energy distribution
³⁵ of background events surviving the application of sequential cuts is shown in Figure 6.7.
³⁶ The energy distribution of background events per MeV of energy deposited, surviving the

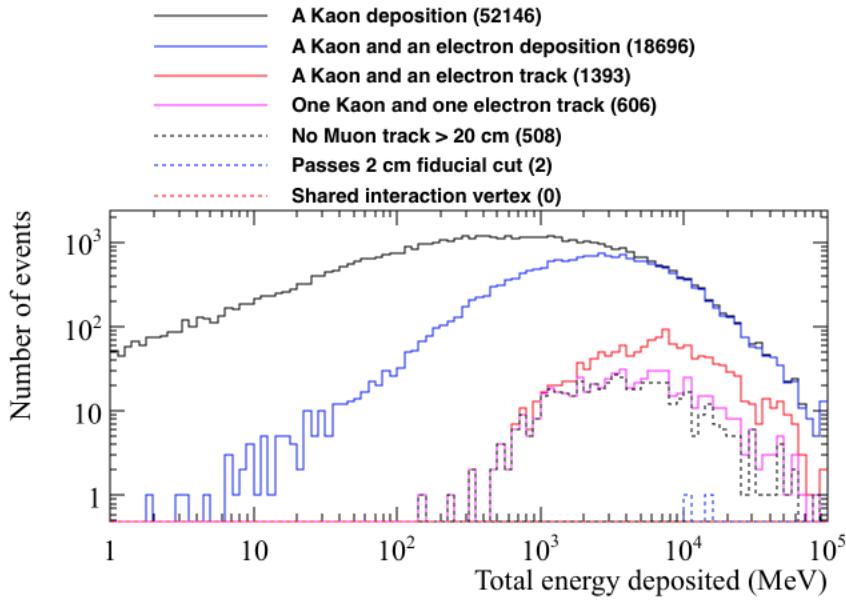


Fig. 6.7 The energy distribution of background events surviving the application of sequential cuts. The total energy deposited in the detector is plotted on the x axis. A sample of 2×10^9 muons, representing 401.6 years of detector for a single FD module is shown.

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 events/Mt/year 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.

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.

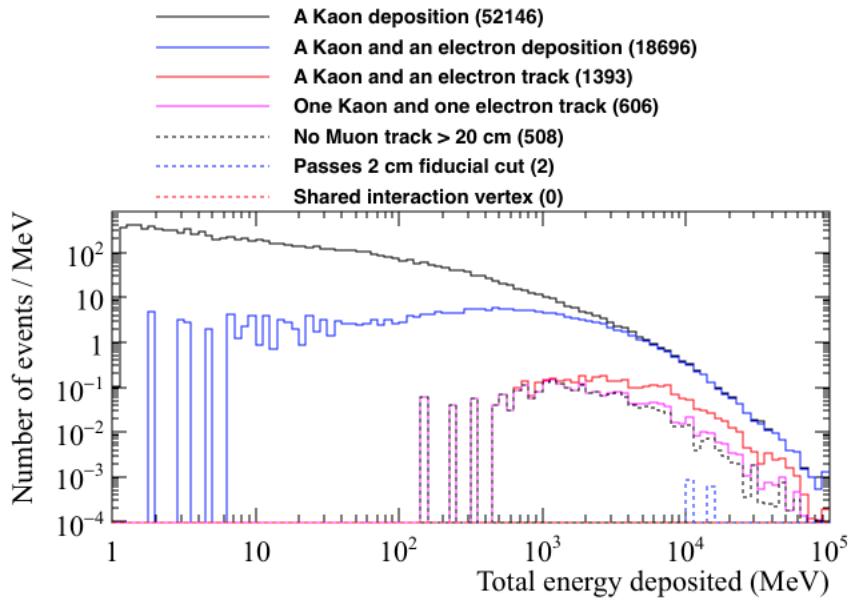


Fig. 6.8 The energy distribution of background events per MeV of deposited energy surviving the application of sequential cuts. The total energy deposited in the detector is plotted on the x axis. This distribution is obtained by dividing the number of events within an energy bin by the bin width. A 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

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 a muon with a track length of more than 20 cm present in the event 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 [105]. 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

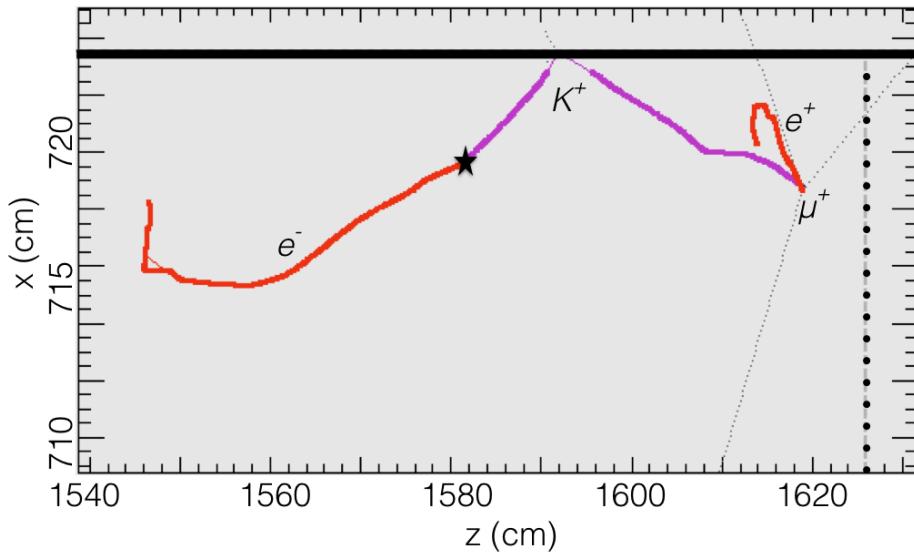


Fig. 6.9 A simulated $n \rightarrow K^+ + e^-$ decay which occurred near the edge of the detector volume. The path of the kaon produced in the decay is shown as a purple line. The path of the muon, produced by the decay of the kaon, is shown as a blue line, though it is very short and so barely visible. 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 spallation neutrons. 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.

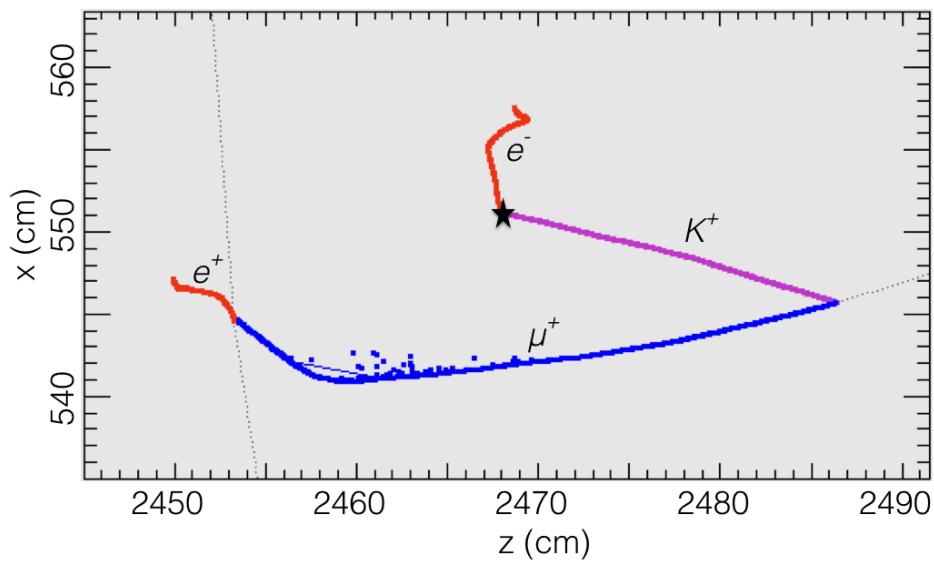


Fig. 6.10 A simulated $n \rightarrow K^+ + e^-$ decay which is fully contained in a single volume. The path of the kaon produced in the decay is shown as a purple line. The path of the muon, produced by the decay of the kaon, is shown as a blue line. The paths which the electrons in the event took are 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 spallation neutrons. 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.

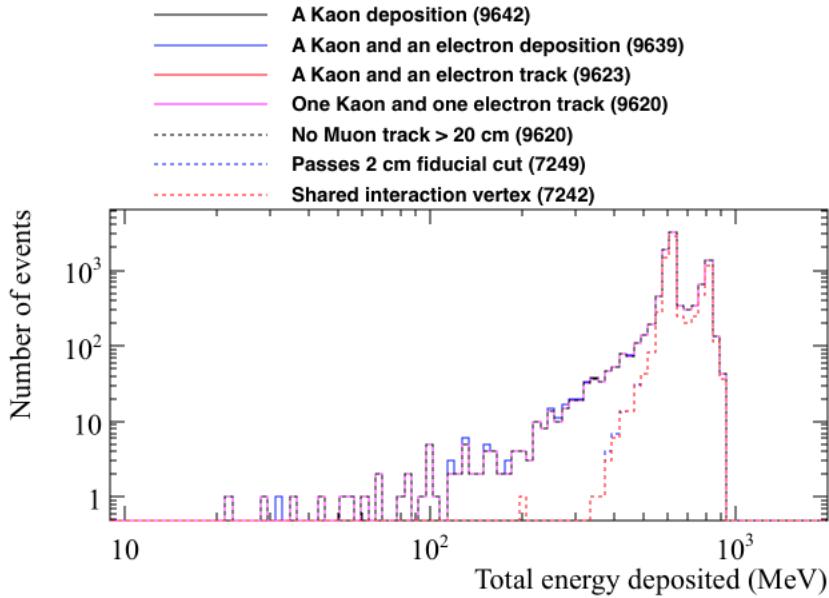


Fig. 6.11 The energy distribution of signal events surviving the application of sequential cuts in the $n \rightarrow K^+ + e^-$ channel. The total energy deposited in the detector is plotted on the x axis.

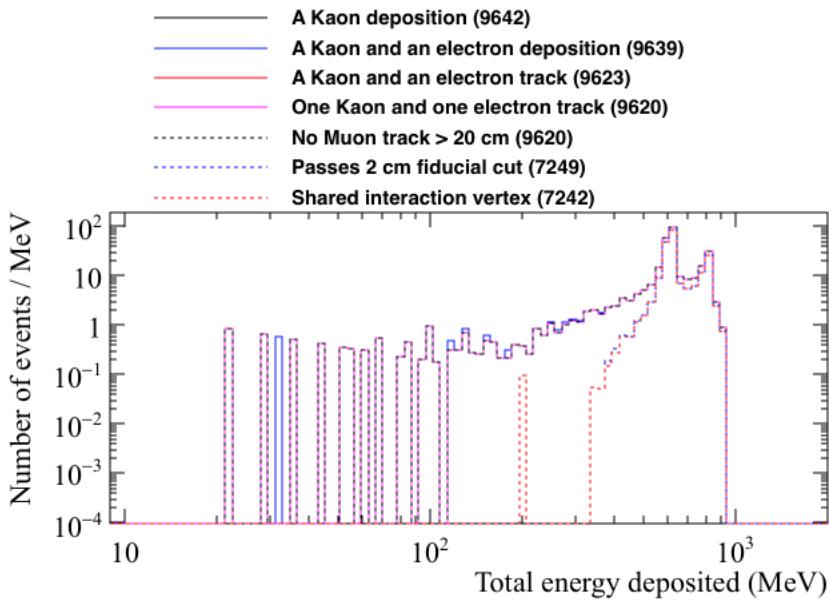


Fig. 6.12 The energy distribution of signal events per MeV of deposited energy surviving the application of sequential cuts. The total energy deposited in the detector is plotted on the x axis. This distribution is obtained by dividing the number of events within an energy bin by the bin width.

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

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.

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. This will mean that a significant proportion of the kaon rest mass energy will not be seen, as the kaon is assumed to decay at rest. Then 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 two body

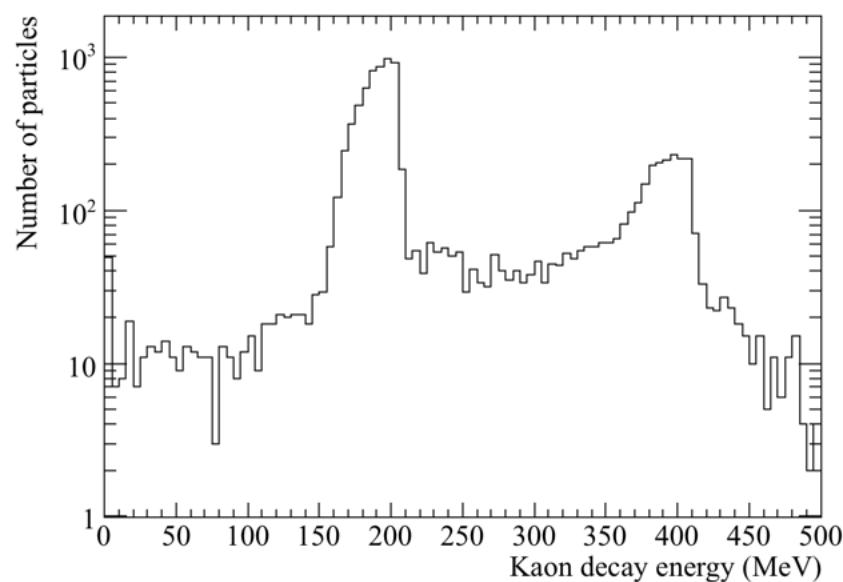


Fig. 6.13 The number of events as a function of the energy deposited by the kaon decay products. The energies deposited due to the $K^+ \rightarrow \mu^+ + \nu_\mu$ and $K^+ \rightarrow \pi^+ + \pi^0$ decay channels can be seen in the peaks at roughly 200 MeV, and 400 MeV, respectively. The underlying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

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 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. 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 the spallation neutrons 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 500 (2,500) 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).

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36

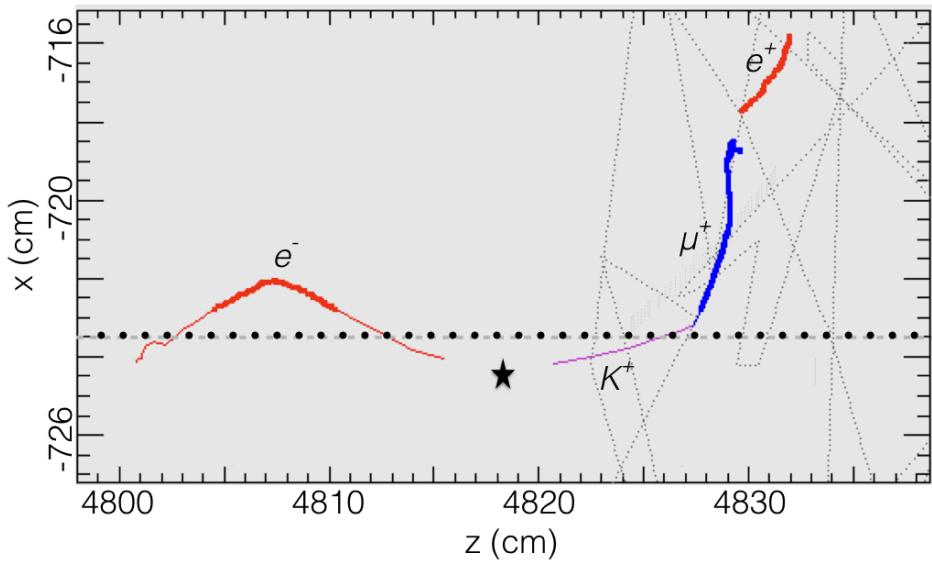


Fig. 6.14 A simulated $n \rightarrow K^+ + e^-$ decay where the kaon did not deposit any energy in the active volume. The path of the kaon produced in the decay is shown as a purple line. The path of the muon, produced by the decay of the kaon, is shown as a blue line. The paths which the electrons in 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.

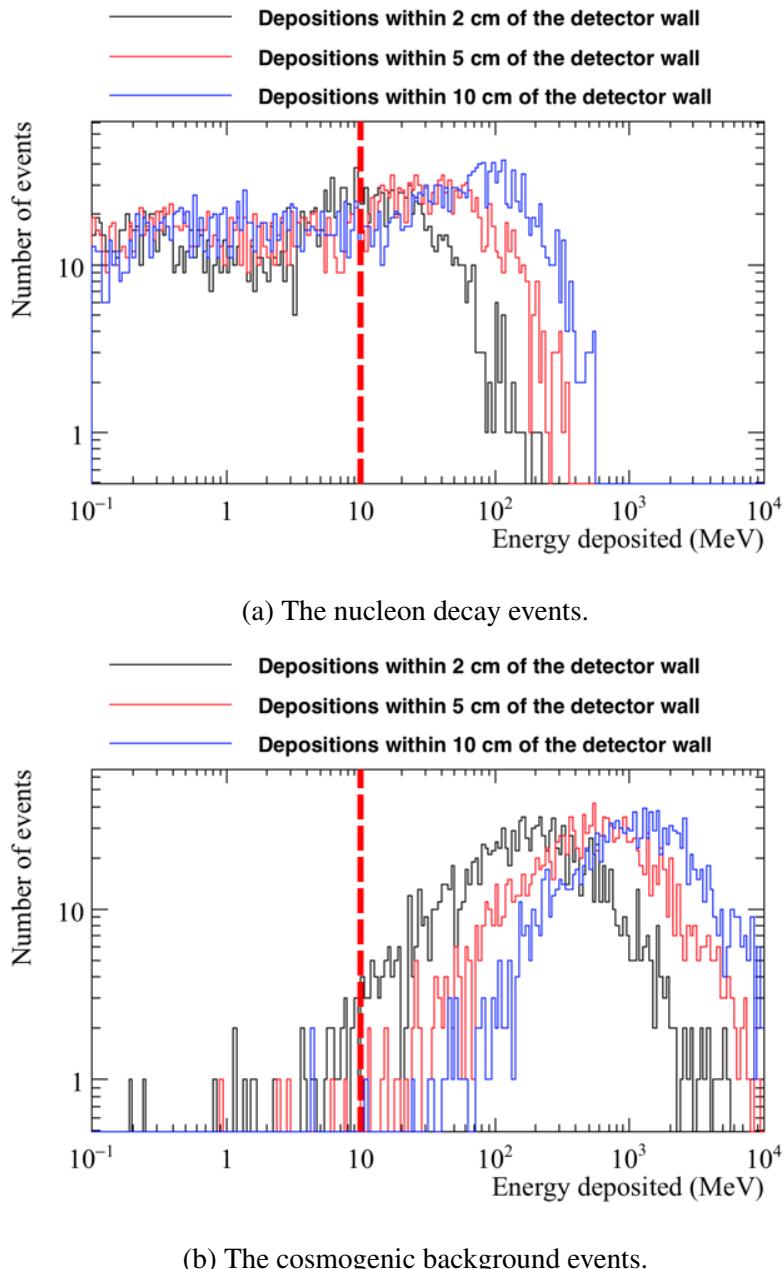


Fig. 6.15 The number of events, as a function of the energy deposited within 2 cm, 5 cm, and 10 cm of the detector edges. Top: energy depositions for the simulated neutron decay events. Bottom: energy depositions for the simulated cosmogenic background events. The histograms are filled after the cut 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.

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

9

10 A cut demanding that there should be no more than 10 MeV of energy deposited within 2
11 cm of the detector edge is used. This cut is designed to preserve as many of the simulated
12 signal events as possible, whilst rejecting as many background events as possible. This
13 cut removes only 683 of the 9,623 signal events, whilst only 39 out of the 1,393 cosmic
14 background events meet this requirement. Whilst this does mean that the fiducial is less
15 effective at removing background events from the cosmic sample, it does not cause the huge
16 loss of signal events which seen when using the hard cut of “no energy depositions within
17 2 cm of the detector edge.” It is also important to remember that the other cuts will still be
18 applied to these 39 events, as the fiducial cut is not applied in isolation. As will be seen in
19 Figure 6.17b none of the muon-induced background events survive the application of all cuts.

20

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

36

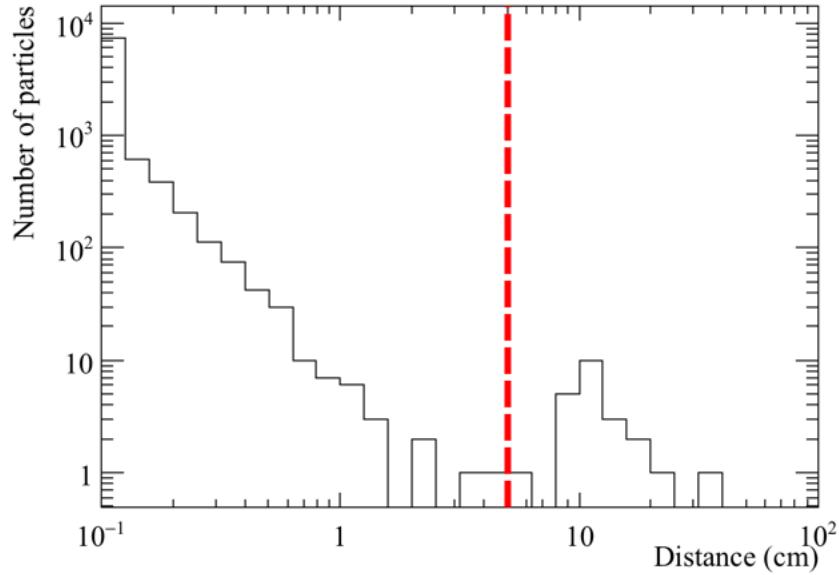
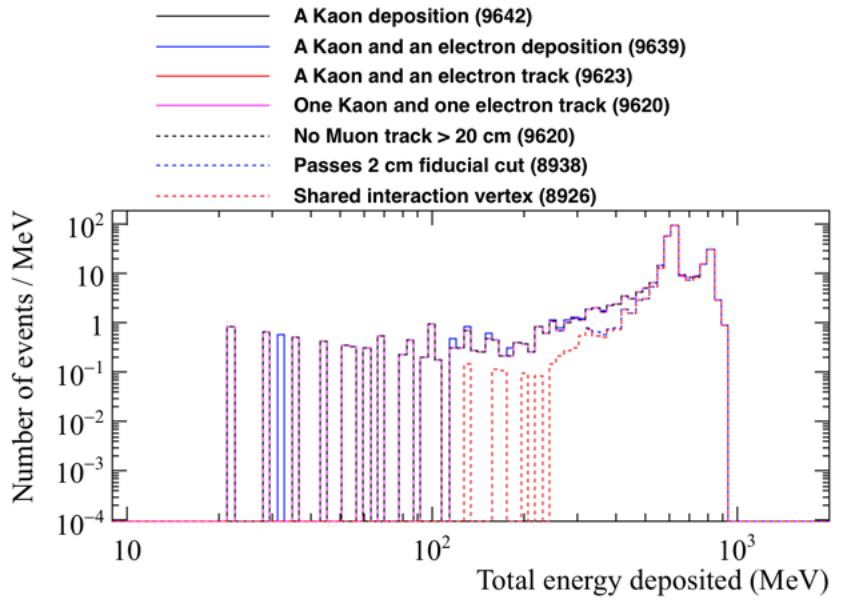


Fig. 6.16 The separation of the kaon and the electron produced in the simulated $n \rightarrow K^+ + e^-$ decays. The dashed red line, drawn at a separation of 5 cm, shows the maximum possible separation a kaon and an electron could have, and still be considered to have a common vertex. A separation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

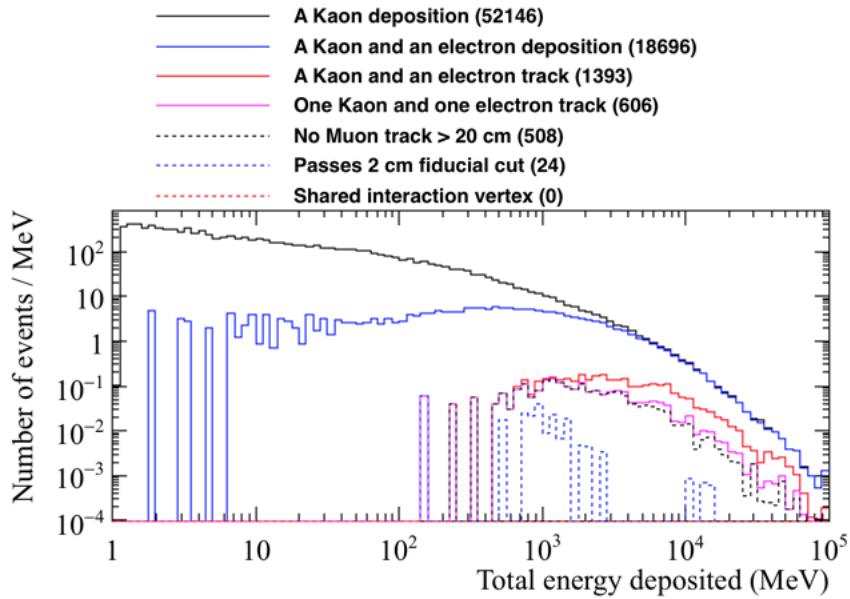
Figure 6.17 shows the energy distributions per MeV of deposited energy for the simulated decay (Figure 6.17a) and cosmogenic background (Figure 6.17b) events, after the fiducial cut is changed to require that there is less than 10 MeV of energy deposited within 2 cm of the detector edge. As before, these 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 (682 events).

It can be seen that none of the cosmic background events pass the application of all cuts, though 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 more than the rest



(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.

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 $\vec{p}_K = -\vec{p}_e$

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

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

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

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

$$\vec{p}_e = \frac{\sqrt{m_n^2 + m_e^2 - m_K^2 - 4m_n^2 m_e^2}}{2m_n} \quad (6.6)$$

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. 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 nucleons which decay in LArTPCs will be contained in argon nuclei, this means that the energies of any particles which are produced will be

¹ smeared by the Fermi motion of the decaying nucleon within the nucleus. Any kaon which
² is produced in the decay is also likely to scatter as it exits the nucleus, further smearing its
³ energy and momenta [176]. This causes the true momenta and total energy of the particles
⁴ produced in the decay to be different from the values which were calculated above. Thus,
⁵ when searching for nucleon decay events, it is necessary to consider energy ranges of a few
⁶ hundred MeV.

⁷

⁸ **6.3.4 Energy constraints on the cosmogenic background to the $n \rightarrow$ ⁹ $K^+ + e^-$ decay channel**

¹⁰ As outlined in Section 6.3.1, it is possible to exclude background events from signal events,
¹¹ using the distribution of energy depositions in the detector. As previously outlined, the energy
¹² depositions were split into several categories:

- ¹³ • The energy directly deposited by the kaon and its secondaries, excluding its decay
¹⁴ products.
- ¹⁵ • The energy deposited by the kaon decay products and any of their secondaries.
- ¹⁶ • The energy directly deposited by the electron and its secondaries.
- ¹⁷ • The energy deposited near the shared kaon and electron vertex.
- ¹⁸ • The energy deposited in the detector which does not fit any of the above criteria.

¹⁹ Following the earlier discussions in Section 6.3.1, it should be clear that the energy directly
²⁰ deposited by the kaon corresponds to the sum of all energy depositions, which are due to the
²¹ kaon or its interaction products. Equivalently, the energy directly deposited by the electron,
²² corresponds to the sum of all energy depositions which are due to the electron as it showers,
²³ and any particles created as a result of the shower. The energy deposited by the kaon decay
²⁴ products would include all depositions by the muon and subsequent electron, in the case that
²⁵ the kaon decayed via $K^+ \rightarrow \mu^+ + \nu_\mu$, and then the muon decayed via $\mu^+ \rightarrow e^+ + \nu_e + \bar{\nu}_\mu$.
²⁶ The energy deposited near the shared kaon and electron vertex, would primarily consist of
²⁷ energy depositions due to spallation products. These depositions would largely be due to
²⁸ protons, though may also include some depositions due to neutrons too, if they deposited
²⁹ energy near the interaction vertex. Any depositions within 5 cm of either the start of the kaon
³⁰ track, or the start of the electron shower, are considered “near” to the interaction vertex. It is
³¹ 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 labelled 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. 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

¹ applied cuts. This is because it is possible for signal events to have no energy identified as
² ‘near’ the common interaction vertex, such as the event shown in Figure 6.6. It is also possi-
³ ble for there to be no energy deposition which would be classed as “other” energy depositions.

⁴

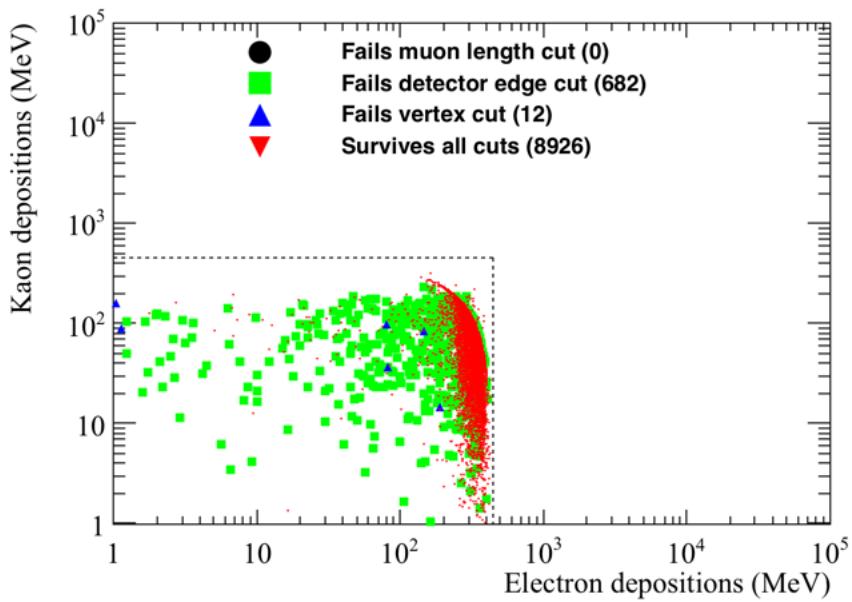
⁵ From Figure 6.18, it can be seen that the electron energy distribution in the cosmic back-
⁶ ground sample, is very different from the one seen in the simulated neutron decay sample.
⁷ The energies deposited by electrons in the nucleon decay sample are tightly concentrated
⁸ between ~ 200 and ~ 400 MeV, whilst in the cosmic background sample, the energy deposited
⁹ by the electron is almost always less than ~ 50 MeV. Many of the electrons in the cosmic
¹⁰ background sample deposit less than 1 MeV of energy in the detector, and so are not shown
¹¹ in Figure 6.18b. This is why Figure 6.18b is very sparse despite 606 events being plotted.
¹² Realistically, these electrons are unlikely to be reconstructed due to their extremely low
¹³ energies. From Figure 6.18a, it can be seen that some of the electrons produced in the nucleon
¹⁴ decay events also deposit very little energy in the detector, though these events generally fail
¹⁵ either the fiducial cut, or the vertex cut. An explanation as to why these events fail the cuts
¹⁶ was presented in Section 6.3.2, when considering Figure 6.14.

¹⁷

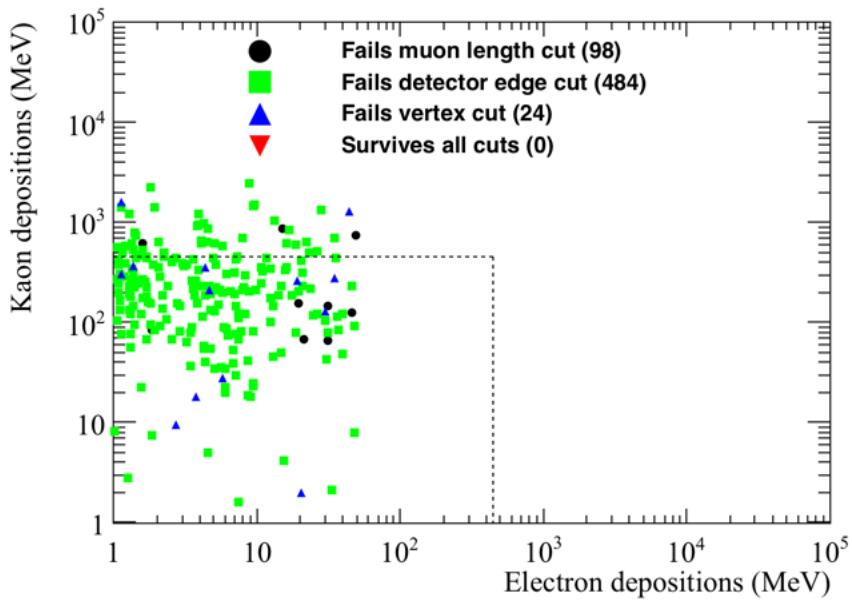
¹⁸ From Figure 6.19a, it can be seen that as the energy deposited near the kaon and electron
¹⁹ vertex increases, the sum of the energy deposited by the kaon and electron decreases. This is
²⁰ reasonable, because, when the spallation protons and neutrons have more energy, the kaon
²¹ and electron will have less energy. The decrease in the energy deposited by the kaon and the
²² electron, is roughly consistent with the increase in the amount of energy which is deposited
²³ near their shared vertex. This means that the sum of the three energies is generally around
²⁴ 450 MeV, as expected from the kinematic calculations shown in Section 6.3.3 where a two
²⁵ body decay was assumed. Though this is no longer valid when 3+ particles are produced,
²⁶ the majority of the rest mass energy of the neutron will be still be measured if the neutron
²⁷ energies are small. This is largely inconsistent with the simulated cosmic background events,
²⁸ where many events have no energy deposited near the shared vertex of the kaon and electron.
²⁹ However, the lack of energy deposited near the shared kaon and electron vertex cannot be
³⁰ used to discriminate against cosmogenic background events, as this is also observed in many
³¹ simulated nucleon decay events. What can be used to separate cosmic background events
³² from nucleon decay events though, is the sum of the three energy criteria, as it can be seen
³³ that this would rarely be around 450 MeV in the cosmic background sample.

³⁴

³⁵ Figure 6.20, shows that the sum of all energy deposits attributed to the kaon, against
³⁶ the sum of all energy depositions which are considered to be “other” energy depositions,

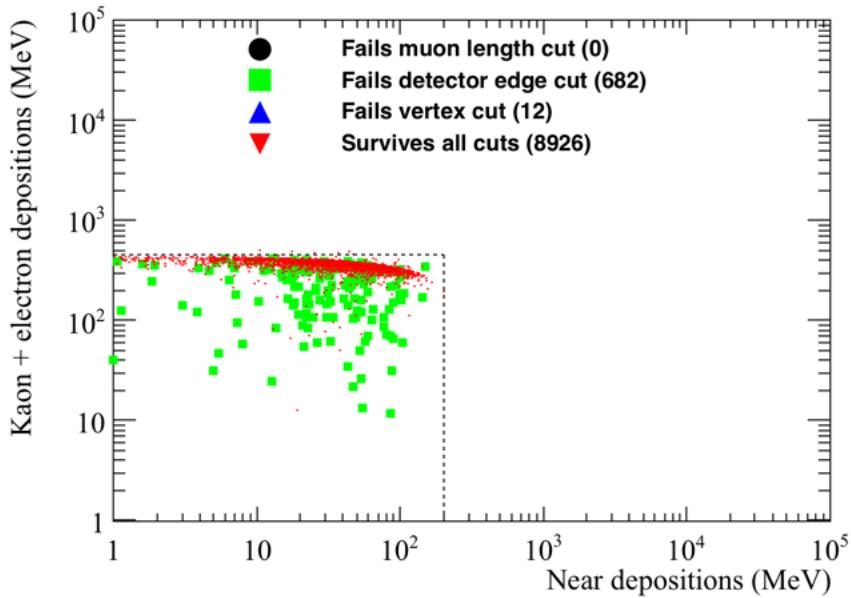


(a) Signal events.

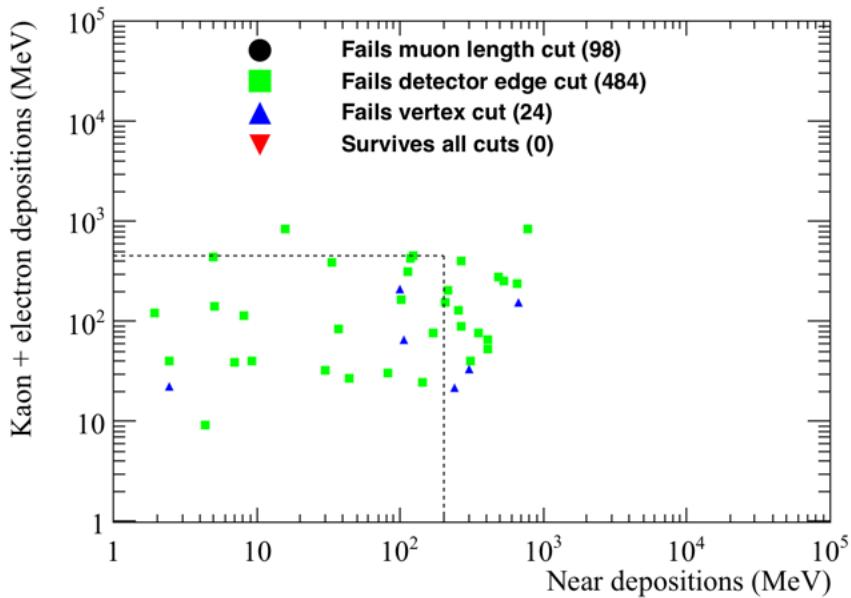


(b) Cosmic background events.

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.

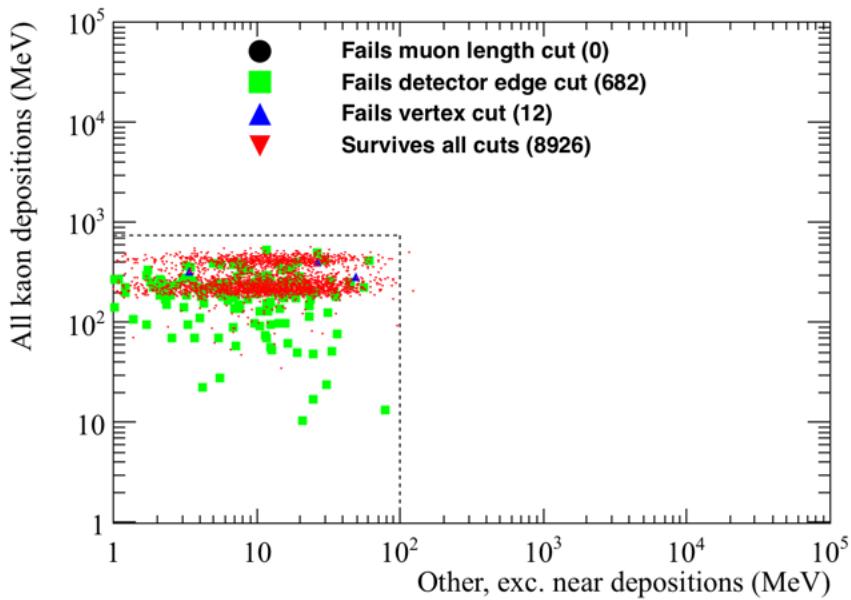


(a) Signal events.

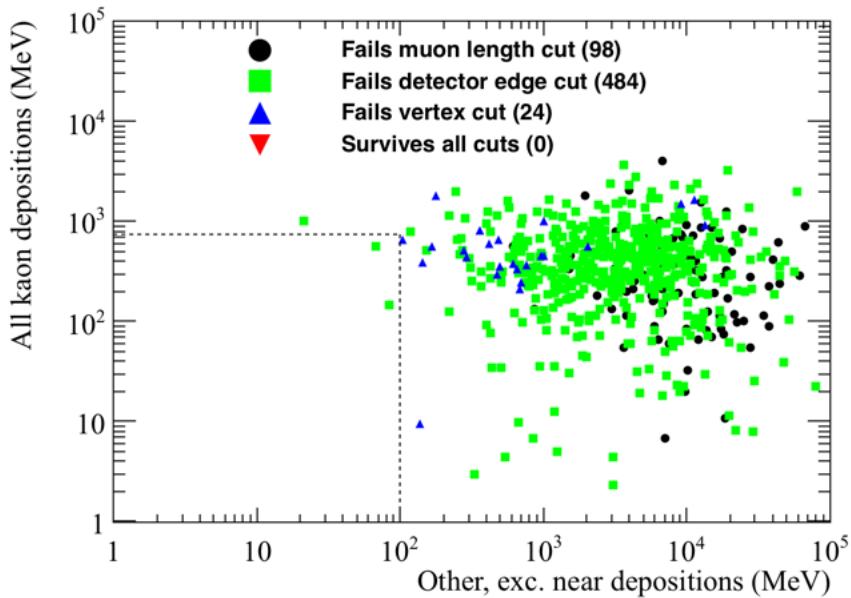


(b) Cosmic background events.

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.

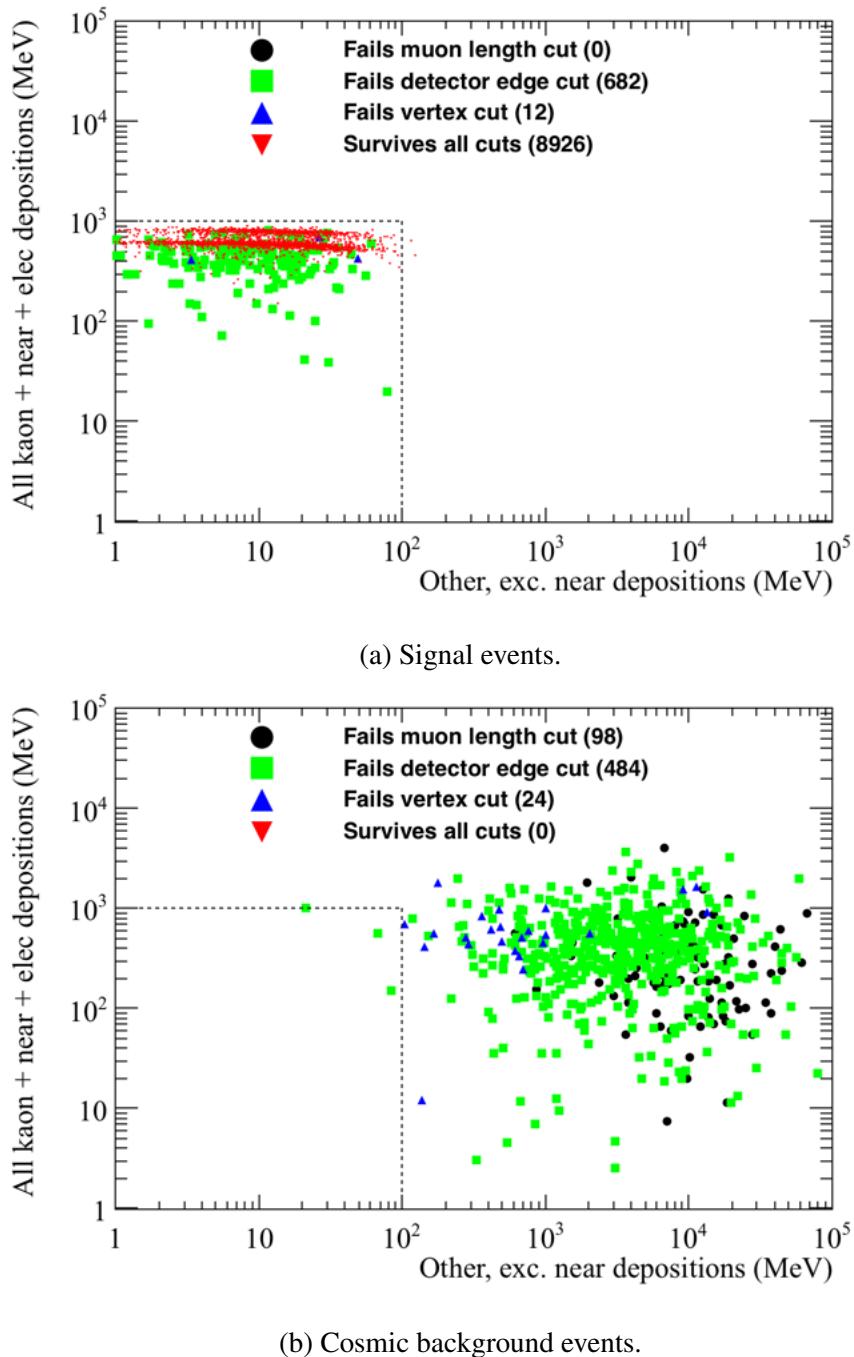


Fig. 6.21 The energy directly deposited by kaons, plus the energy deposited by the kaon decay products, plus the energy directly deposited by electrons, plus the energy deposited near the kaon and electron vertex versus the energy depositions which do not fit any of the other criteria, in the simulated 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.

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.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35

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

9

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

17

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

31

32 Therefore, the result of a preliminary study on background rejection for the $n \rightarrow K^+ + e^-$
33 decay channel, found no muon-induced background events which would survive the ap-
34 plication of cuts, in a sample of 2×10^9 muons representing 401.6 years of detector for
35 a single FD module. This corresponds to a limit of the background rate of less than 0.44
36 events/Mt/year 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 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 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 6.15 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 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 still 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 on 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 [175];

6.3 Nucleon decay channels in DUNE

218

- ¹ • A minimum of 10 MeV deposited in the detector volume.
 - ² • A maximum of 3,000 MeV deposited in the detector volume.
 - ³ • A maximum of 5 MeV deposited within 10 cm of the detector edge.
- ⁴ These criteria are designed to be soft enough that the full range of nucleon decay modes can
⁵ be studied, including di-nucleon decay modes and neutron-antineutron oscillations. This is
⁶ why the maximum deposited energy greatly exceeds the rest mass of a single nucleon. This
⁷ also accounts for the fact that during reconstruction some energy depositions may not be
⁸ reconstructed, so even though the total energy deposited in the detector is 3 GeV, the total
⁹ reconstructed energy may be less.

¹⁰

¹¹ Upon performing reconstruction, and hence energy and position smearing, it is likely that
¹² the energy cuts which were outlined in Section 6.3.1 will have to be broadened. This may
¹³ result in more background events being contained within the expected energy distributions for
¹⁴ signal events, however the effectiveness of the cuts to remove background events should be
¹⁵ largely unaffected. The reason for this, is that upon performing reconstruction, the separation
¹⁶ of the kaon and electron track will not decrease, and so the cut is expected to remain very
¹⁷ efficient. The only cut which may be less effective is the requirement that there only be
¹⁸ one kaon track, and only one electron shower in the event. This is because some of the low
¹⁹ energy electron showers may not be reconstructed. However, even when this cut was relaxed
²⁰ no background events survived the application of all cuts.

²¹

²² The reduced reconstruction efficiency will also affect the nucleon decay signals and so
²³ this, combined with the broader distributions of energies which will be reconstructed, will
²⁴ mean that the efficiency with which signal events can be identified will probably decrease.
²⁵ Studies have been performed to establish the event identification efficiencies when using
²⁶ reconstructed objects [177–180], though currently these have only been done with respect to
²⁷ the $p \rightarrow K^+ + \bar{\nu}_e$ decay channel. These studies have found that the current signal reconstruc-
²⁸ tion efficiency is much less than the 100% assumed here, at around 30.5% [178], though it is
²⁹ imagined that this will improve significantly before the DUNE detector is built.

³⁰

³¹ However, it is envisioned that there will still be no muon-induced background events
³² which survive the application of all cuts, though the impact of atmospheric neutrinos remains
³³ to be seen, and so the observation of a single event in the DUNE detector could provide a
³⁴ robust indication of nucleon decay.

³⁵

Chapter 7

Concluding remarks

This thesis has concerned work performed done in relation 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 [131].

Chapter 4 outlined the simulation efforts which were performed in preparation for the 35 ton Phase II run, culminating in a proposed 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, when the simulation was extended to contain cosmic rays, as would be present in the 35 ton dataset, the method was seen to be less effective. This was primarily down to the large number of very short reconstructed tracks which were associated with the 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 separations could be improved by using Monte Carlo truth information regarding the true stopping point of the simulated particles. 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

¹ be fully developed in order for this to be possible, as otherwise through going muons would
² saturate any sample of stopping particles.

³

⁴ A discussion of the features and issues surrounding the 35 ton dataset were presented
⁵ in Chapter 5, concluding in the proposal of a new method of interaction time determination
⁶ using the effects of longitudinal diffusion. Two metrics were proposed whereby this could be
⁷ done, though considering the *RMS/Charge* of hits in reconstructed tracks was found to be
⁸ the better identifier, where the *RMS* is the width of the reconstructed hit in ticks (500 ns),
⁹ and the *Charge* is the integrated area of the reconstructed hit (ADC). It was found that the
¹⁰ interaction time of through-going tracks could be determined to an accuracy of 171 μ s, where
¹¹ the distribution has a FWHM of 210 μ s, over a drift window of 5,200 μ s. When this is
¹² converted to a prediction of the *x* (drift) position of the particle, the accuracy is found to be
¹³ 18.5 cm, with a FWHM of 23.0 cm, over a drift length of 223 cm. This was then compared
¹⁴ to a simulated low-noise detector, where the accuracy of the interaction time (*x* position)
¹⁵ determination was found to be 3 μ s (0.4 cm) with a FWHM of 114 μ s (12.6 cm). The
¹⁶ effect of changing detector conditions in Monte Carlo was then observed with relation to; the
¹⁷ electronics noise level of the detector, the applied electric field, the electron lifetime, and the
¹⁸ magnitude of the longitudinal diffusion constant. As expected, lower noise levels, and larger
¹⁹ longitudinal diffusion constants were seen to increase the accuracy of the method. Lower
²⁰ electron lifetimes were seen to cause the method to be more effective, as this corresponds to
²¹ larger decreases in hit charge, giving the method a stronger handle on the *x* (drift) position of
²² the particle. The method was seen to be largely unaffected by the electric field which was
²³ applied to the detector.

²⁴

²⁵ Finally, Chapter 6 concerns simulations of background events in LArTPCs. The back-
²⁶ grounds to neutrino interactions in a surface detector built on work presented in [153], where
²⁷ a system of cuts designed to reject background events had been developed. The authors
²⁸ contribution concerned the addition of an accurate detector geometry, and an accurate surface
²⁹ profile to simulations. The accurate detector geometry was seen to significantly affect the
³⁰ predicted background rate, causing the hadronic flux to be reduced by an order of magnitude
³¹ due to the addition of extra shielding, whilst the muon-induced background was seen to
³² increase due to the detector having a larger surface area. The background rate was seen to
³³ be largely unaffected by the addition of the accurate surface profile. The combined annual
³⁴ background rate from muons, protons and neutrons, after all cuts were applied, was seen to
³⁵ be reduced from \approx 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.

1

2

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 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.

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

¹ References

- ² [1] G. Aad et al. (ATLAS Collaboration), Phys. Lett. **B716**, 1 (2012).
- ³ [2] S. Chatrchyan et al. (CMS Collaboration), Phys. Lett. **B716**, 30 (2012).
- ⁴ [3] B. P. Abbott et al. (LIGO Scientific Collaboration and Virgo Collaboration), Phys. Rev. Lett. **116**, 061102 (2016).
- ⁵
- ⁶ [4] Y. Fukuda et al. (Super-Kamiokande Collaboration), Phys. Rev. Lett. **81**, 1562 (1998).
- ⁷ [5] Q. R. Ahmad et al. (SNO Collaboration), Phys. Rev. Lett. **89**, 011301 (2002).
- ⁸ [6] W. Pauli, *Open letter to the Gauverein meeting in Tubingen* (1930).
- ⁹ [7] E. Fermi, Z. Phys. **88**, 161 (1934).
- ¹⁰ [8] O. Kofoed-Hansen, Phys. Rev. **71**, 451 (1947).
- ¹¹ [9] G. C. Hanna and B. Pontecorvo, Phys. Rev. **75**, 983 (1949).
- ¹² [10] L. M. Langer and R. J. D. Moffat, Phys. Rev. **88**, 689 (1952).
- ¹³ [11] C. L. Cowan et al., Science **124**, 103 (1956).
- ¹⁴ [12] G. Danby et al., Phys. Rev. Lett. **9**, 36 (1962).
- ¹⁵ [13] R. Davis et al., Phys. Rev. Lett. **20**, 1205 (1968).
- ¹⁶ [14] B. T. Cleveland et al., Astrophys. J. **496**, 505 (1998).
- ¹⁷ [15] B. Pontecorvo, Sov. Phys. JETP **6**, 429 (1957), [Zh. Eksp. Teor. Fiz. 33,549(1957)].
- ¹⁸ [16] K. S. Hirata et al., Phys. Rev. Lett. **63**, 16 (1989).
- ¹⁹ [17] M. L. Perl et al., Phys. Rev. Lett. **35**, 1489 (1975).
- ²⁰ [18] K. Kodama et al., Physics Letters B **504**, 218 (2001), ISSN 0370-2693.
- ²¹ [19] Q. R. Ahmad et al. (SNO), Phys. Rev. Lett. **87**, 071301 (2001).
- ²² [20] Z. Maki et al., Progress of Theoretical Physics **28**, 870 (1962).
- ²³ [21] S. Andringa et al. (SNO+ Collaboration), Adv. High Energy Phys. **2016**, 6194250 (2016).
- ²⁴

-
- [22] R. Arnold et al., The European Physical Journal C **70**, 927 (2010), ISSN 1434-6052. 1
- [23] D. Michael et al. (Minos Collaboration), Nuclear Instruments and Methods in Physics Research A **596**, 190 (2008). 2
3
- [24] G. Tzanankos et al. (MINOS+ Collaboration), Tech. Rep. (2011). 4
- [25] H. Nunokawa et al., Prog. Part. Nucl. Phys. **60**, 338 (2008). 5
- [26] C. Patrignani et al. (Particle Data Group), Chin. Phys. C**40**, 100001 (2016). 6
- [27] N. Cabibbo, Phys. Rev. Lett. **10**, 531 (1963). 7
- [28] M. Kobayashi and T. Maskawa, Prog. Theor. Phys. **49**, 652 (1973). 8
- [29] L. Wolfenstein, Phys. Rev. D **17**, 2369 (1978). 9
- [30] S. P. Mikheev and A. Yu. Smirnov, Sov. J. Nucl. Phys. **42**, 913 (1985), [Yad. Fiz.42,1441(1985)]. 10
11
- [31] G. Fogli et al., Physics Letters B **583**, 149 (2004). 12
- [32] Hewett, J. L. and others, Tech. Rep. ANL-HEP-TR-12-25, SLAC-R-991, FERMILAB-CONF-12-879-PPD (2012). 13
14
- [33] I. Esteban et al., JHEP **01**, 087 (2017). 15
- [34] P. A. R. Ade et al. (Planck), Astron. Astrophys. **594**, A13 (2016). 16
- [35] S. Mertens, Physics Procedia **61**, 267 (2015). 17
- [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). 18
19
20
- [37] C. Athanassopoulos et al. (LSND Collaboration), Phys. Rev. Lett. **81**, 1774 (1998). 21
- [38] A. Aguilar et al. (LSND Collaboration), Phys. Rev. D **64**, 112007 (2001). 22
- [39] A. A. Aguilar-Arevalo et al. (MiniBooNE Collaboration), Phys. Rev. Lett. **110**, 161801 (2013). 23
24
- [40] W. J. Marciano and G. Senjanović, Phys. Rev. D **25**, 3092 (1982). 25
- [41] H. Georgi and S. L. Glashow, Phys. Rev. Lett. **32**, 438 (1974). 26
- [42] H. Weyl, Zeitschrift für Physik **56**, 330 (1929). 27
- [43] G. Senjanovic, AIP Conf. Proc. **1200**, 131 (2010). 28
- [44] R. Foot et al., Zeitschrift für Physik C Particles and Fields **44**, 441 (1989). 29
- [45] I. Dorsner and P. Fileviez Perez, Nucl. Phys. **B723**, 53 (2005), hep-ph/0504276. 30

- ¹ [46] S. P. Martin (1997), [Adv. Ser. Direct. High Energy Phys.18,1(1998)], arXiv:hep-ph/9709356.
- ²
- ³ [47] D. J. Castano et al., Phys. Rev. **D49**, 4882 (1994), hep-ph/9308335.
- ⁴ [48] M. E. Gomez, PoS **trieste99**, 006 (1999), hep-ph/9908419.
- ⁵ [49] G. R. Farrar and P. Fayet, Physics Letters B **76**, 575 (1978), ISSN 0370-2693.
- ⁶ [50] F. Vissani, Phys. Rev. **D52**, 4245 (1995).
- ⁷ [51] M. Severson, Phys. Rev. **D92**, 095026 (2015).
- ⁸ [52] G. Battistoni et al., Physics Letters B **133**, 454 (1983).
- ⁹ [53] C. Berger et al., Physics Letters B **269**, 227 (1991).
- ¹⁰ [54] M. C. Goodman (Soudan-2), in *Proceedings of the 26th International Cosmic Ray Conference. August 17-25, 1999. Salt Lake City, Utah, USA* (1999), p. 364.
- ¹¹
- ¹² [55] W. Gajewski et al., Phys. Rev. **D42**, 2974 (1990).
- ¹³ [56] K. Abe et al. (Super-Kamiokande Collaboration), Phys. Rev. D **95**, 012004 (2017).
- ¹⁴ [57] K. Abe et al., *Letter of Intent: The Hyper-Kamiokande Experiment — Detector Design and Physics Potential* — (2011), arXiv:1109.3262.
- ¹⁵
- ¹⁶ [58] S. Ritz et al. (HEPAP Subcommittee), *Building for Discovery: Strategic Plan for U.S. Particle Physics in the Global Context* (2014), INSPIRE-1299183.
- ¹⁷
- ¹⁸ [59] LBNE Collaboration, Tech. Rep. LBNE-doc-5235 (2012), URL <http://lbne2-docdb.fnal.gov/cgi-bin/RetrieveFile?docid=5235&filename=CDR-intro-volume-2012nov20.pdf&version=9>.
- ¹⁹
- ²⁰
- ²¹ [60] LBNE Collaboration, Tech. Rep. LBNE-doc-4317 (2012), URL <http://lbne2-docdb.fnal.gov/cgi-bin/RetrieveFile?docid=4317&filename=CDR-beam-volume-101812.pdf&version=25>.
- ²²
- ²³
- ²⁴ [61] LBNE Collaboration, Tech. Rep. LBNE-doc-4724 (2012), URL <http://lbne2-docdb.fnal.gov/cgi-bin/RetrieveFile?docid=4724&filename=CDR-NDC-volume-final-101912.pdf&version=12>.
- ²⁵
- ²⁶
- ²⁷ [62] LBNE Collaboration, Tech. Rep. LBNE-doc-4892 (2012), URL <http://lbne2-docdb.fnal.gov/cgi-bin/RetrieveFile?docid=4892&filename=CDR-LArFD-volume-final-2012Oct22.pdf&version=12>.
- ²⁸
- ²⁹
- ³⁰ [63] LBNE Collaboration, Tech. Rep. LBNE-doc-4623 (2012), URL <http://lbne2-docdb.fnal.gov/cgi-bin/RetrieveFile?docid=4623&filename=CDR-NSCF-volume-final-2012Oct10.pdf&version=9>.
- ³¹
- ³²
- ³³ [64] LBNE Collaboration, Tech. Rep. LBNE-doc-5017 (2012), URL <http://lbne2-docdb.fnal.gov/cgi-bin/RetrieveFile?docid=5017&filename=CDR-FSCF-volume-final-2012Oct10.pdf&version=5>.
- ³⁴
- ³⁵

-
- [65] A. Stahl et al., Tech. Rep. CERN-SPSC-2012-021, SPSC-EOI-007 (2012). 1
- [66] 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. 2
3
4
- [67] 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. 5
6
7
- [68] S. D. Holmes et al., Proc. of IPAC 2015 pp. 3982–3985 (2015). 8
- [69] 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. 9
10
11
- [70] 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. 12
13
14
- [71] DUNE Collaboration, *Proposals and Design Reports*, [online] <https://web.fnal.gov/project/LBNF/SitePages/Proposals%20and%20Design%20Reports.aspx>, URL <https://web.fnal.gov/project/LBNF/SitePages/Proposals%20and%20Design%20Reports.aspx>. 15
16
17
18
- [72] E. Worcester, Tech. Rep. DUNE-doc-2329 (2017), URL <http://docs.dunescience.org:8080/cgi-bin>ShowDocument?docid=2329>. 19
20
- [73] M. Bass et al., Phys. Rev. **D91**, 052015 (2015). 21
- [74] M. V. Diwan, Frascati Phys. Ser. **35**, 89 (2004). 22
- [75] E. Worcester, Tech. Rep. DUNE-doc-2335 (2017), URL <http://docs.dunescience.org:8080/cgi-bin>ShowDocument?docid=2335>. 23
24
- [76] E. Worcester, Tech. Rep. DUNE-doc-2401 (2017), URL <http://docs.dunescience.org:8080/cgi-bin>ShowDocument?docid=2401>. 25
26
- [77] E. Worcester, Tech. Rep. DUNE-doc-2332 (2017), URL <http://docs.dunescience.org:8080/cgi-bin>ShowDocument?docid=2332>. 27
28
- [78] E. Worcester, Tech. Rep. DUNE-doc-2377 (2017), URL <http://docs.dunescience.org:8080/cgi-bin>ShowDocument?docid=2377>. 29
30
- [79] M. C. Gonzalez-Garcia et al., JHEP **11**, 052 (2014). 31
- [80] J. Beringer et al. (Particle Data Group), Phys. Rev. D **86**, 010001 (2012). 32
- [81] H. Nishino et al. (Super-Kamiokande Collaboration), Phys. Rev. D **85**, 112001 (2012). 33
- [82] A. Bueno et al., JHEP **04**, 041 (2007). 34
- [83] J. Klinger et al., Phys. Lett. **B746**, 44 (2015). 35

- ¹ [84] M. Antonello et al., *Adv. High Energy Phys.* **2013**, 260820 (2013).
- ² [85] L. Lin and M. Goodman, *Tech. Rep. DUNE-doc-679* (2016), URL <http://docs.dunescience.org:8080/cgi-bin>ShowDocument?docid=679>.
- ⁴ [86] C. Adams et al. (LBNE Collaboration), *The Long-Baseline Neutrino Experiment: Exploring Fundamental Symmetries of the Universe* (2013), arXiv: 1307.7335.
- ⁶ [87] A. Blake, *Tech. Rep. LBNE-doc-8836* (2014), URL http://lbne2-docdb.fnal.gov:8080/cgi-bin/RetrieveFile?docid=8836&filename=AtmosphericNeutrinos_20140402.pdf&version=1.
- ⁹ [88] K. Scholberg, *Annual Review of Nuclear and Particle Science* **62**, 81 (2012).
- ¹⁰ [89] R. Laha and J. F. Beacom, *Phys. Rev.* **D89**, 063007 (2014).
- ¹¹ [90] R. M. Bionta et al., *Phys. Rev. Lett.* **58**, 1494 (1987).
- ¹² [91] K. Hirata et al., *Phys. Rev. Lett.* **58**, 1490 (1987).
- ¹³ [92] M. G. Aartsen et al. (IceCube), *Phys. Rev. Lett.* **110**, 131302 (2013).
- ¹⁴ [93] K. Choi et al. (Super-Kamiokande), *Phys. Rev. Lett.* **114**, 141301 (2015).
- ¹⁵ [94] J. F. Beacom, *Ann. Rev. Nucl. Part. Sci.* **60**, 439 (2010).
- ¹⁶ [95] O. L. Caballero et al., *Astrophys. J.* **745**, 170 (2012).
- ¹⁷ [96] O. L. Caballero et al., *Phys. Rev.* **D80**, 123004 (2009).
- ¹⁸ [97] M. Thomson, *Tech. Rep. LBNE-doc-11100* (2015), URL <http://lbne2-docdb.fnal.gov:8080/cgi-bin/RetrieveFile?docid=11100&filename=P03-Thomson-DUNE-CollaborationStrategy.pdf&asof=2015-6-2>.
- ²¹ [98] M. Adamowski et al., *JINST* **9**, P07005 (2014).
- ²² [99] D. Montanari et al., *Phys. Procedia* **67**, 308 (2015).
- ²³ [100] A. Hahn et al. (LBNE), in *Proceedings, 21st Symposium on Room-Temperature Semiconductor X-ray and Gamma-ray Detectors (RTSD 2014): Seattle, WA, USA, November 8-15, 2014* (2016), p. 7431158.
- ²⁶ [101] E. D. Church, *LArSoft: A Software Package for Liquid Argon Time Projection Drift Chambers* (2013), arXiv:1311.6774.
- ²⁸ [102] The LArSoft collaboration, *Software for LArTPCs* (2017), URL <http://larsoft.org>.
- ²⁹ [103] R. Brun and F. Rademakers, *Nucl. Instrum. Meth.* **A389**, 81 (1997).
- ³⁰ [104] S. Agostinelli et al., *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **506**, 250 (2003).
- ³² [105] C. Andreopoulos et al., *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **614**, 87 (2010).

-
- [106] D. Casper, Nucl. Phys. Proc. Suppl. **112**, 161 (2002). 1
- [107] C. Hagman et al., Tech. Rep. UCRL-TM-229453 (2012), URL http://nuclear.llnl.gov/simulation/doc_cry_v1.7/cry.pdf. 2
3
- [108] C. Hagmann et al., Tech. Rep. UCRL-TM-229452 (2012), URL http://nuclear.llnl.gov/simulation/doc_cry_v1.7/cry_physics.pdf. 4
5
- [109] D. Heck et al., Tech. Rep. (1998). 6
- [110] V. Kudryavtsev, Computer Physics Communications **180**, 339 (2009). 7
- [111] V. A. Kudryavtsev, N. J. C. Spooner, and J. E. McMillan, Nucl. Instrum. Meth. **A505**, 688 (2003). 8
9
- [112] P. Antonioli et al., Astropart. Phys. **7**, 357 (1997). 10
- [113] V. A. Kudryavtsev et al., Phys. Lett. **B471**, 251 (1999). 11
- [114] T. Alion, *LBNE Geometries* (2014), URL https://cdcvns.fnal.gov/redmine/projects/lbnecode/wiki/LBNE_Geometries. 12
13
- [115] J. Asaadi et al., Tech. Rep. DUNE-doc-1689 (2016), URL <http://docs.dunescience.org:8080/cgi-bin>ShowDocument?docid=1689>. 14
15
- [116] C. Harris and M. Stephens, in *In Proc. of Fourth Alvey Vision Conference* (1988), pp. 147–151. 16
17
- [117] P. V. C. Hough, in *Proceedings, 2nd International Conference on High-Energy Accelerators and Instrumentation, HEACC 1959: CERN, Geneva, Switzerland, September 14-19, 1959* (1959), vol. C590914, pp. 554–558. 18
19
20
- [118] J. S. Marshall and M. A. Thomson, Eur. Phys. J. **C75**, 439 (2015). 21
- [119] M. Wallbank, Tech. Rep. DUNE-doc-1369 (2016), URL <http://docs.dunescience.org:8080/cgi-bin>ShowDocument?docid=1369>. 22
23
- [120] R. Sulej and D. Stefan, *Reconstruction: Projection Matching*, Reconstruction working group meeting (2016), URL <https://indico.fnal.gov/getFile.py/access?contribId=62&sessionId=40&resId=1&materialId=slides&confId=10276>. 24
25
26
- [121] J. B. Birks, Proc. Phys. Soc. **A64**, 874 (1951). 27
- [122] R. Acciarri et al. (ArgoNeuT Collaboration), JINST **8**, P08005 (2013). 28
- [123] J. Thomas and D. A. Imel, Phys. Rev. A **36**, 614 (1987). 29
- [124] A. Blatter et al., Journal of Instrumentation **9**, P04006 (2014). 30
- [125] S. C. Delaquis et al., JINST **8**, T12001 (2013). 31
- [126] K. Mavrokordidis et al., JINST **9**, P02006 (2014). 32
- [127] E. Erdal et al., JINST **10**, P11002 (2015). 33

- ¹ [128] N. McConkey et al., JINST **12**, P03014 (2017).
- ² [129] G. Ghibaudo and F. Balestra, Microelectronics Reliability **37**, 1353 (1997).
- ³ [130] H. Zhao and X. Liu, Cryogenics **55**, 79 (2013).
- ⁴ [131] 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.
- ⁷ [132] H. Bethe, Annalen der Physik **397**, 325 (1930).
- ⁸ [133] F. Bloch, Annalen der Physik **408**, 285 (1933).
- ⁹ [134] X. Qian, *A Four-Plane proposal for Single-Phase LArTPC*, Far Detector Sim/Reco working group meeting (2016), URL <https://indico.fnal.gov/getFile.py/access?contribId=78&sessionId=20&resId=0&materialId=slides&confId=10613>.
- ¹² [135] J. Insler, *Module to unstick ADC codes*, 35 ton working group meeting (2016), URL <https://indico.fnal.gov/getFile.py/access?contribId=2&resId=0&materialId=slides&confId=11627>.
- ¹⁵ [136] MicroBooNE Collaboration, Tech. Rep. MICROBOONE-NOTE-1016-PUB (2016), URL <http://www-microboone.fnal.gov/publications/publicnotes/MICROBOONE-NOTE-1016-PUB.pdf>.
- ¹⁸ [137] J. Cooley and J. Turkey, Math. Comp. **19**, 297 (1965).
- ¹⁹ [138] N. Wiener, *Extrapolation, Interpolation, and Smoothing of Stationary Time Series* ([Cambridge]: Technology Press of the Massachusetts Institute of Technology., 1942).
- ²¹ [139] M. Johnson et al., Tech. Rep. DUNE-doc-1704 (2016), URL <http://docs.dunescience.org:8080/cgi-bin>ShowDocument?docid=1704>.
- ²³ [140] L. Bagby et al., Tech. Rep. DUNE-doc-1834 (2016), URL <http://docs.dunescience.org:8080/cgi-bin>ShowDocument?docid=1834>.
- ²⁵ [141] T. Yang, *Update on reconstruction of 35t data*, 35 ton working group meeting (2016), URL <https://indico.fnal.gov/getFile.py/access?contribId=1&resId=0&materialId=slides&confId=12349>.
- ²⁸ [142] P. Torr and A. Zisserman, Computer Vision and Image Understanding **78**, 138 (2000).
- ²⁹ [143] M. Thiesse, *Hit reconstruction and purity*, 35 ton working group meeting (2016), URL <https://indico.fnal.gov/getFile.py/access?contribId=1&resId=0&materialId=slides&confId=12859>.
- ³² [144] C. Anderson et al., JINST **7**, P10019 (2012).
- ³³ [145] T. Yang, *Update on 35 ton analysis*, 35 ton working group meeting (2016), URL <https://indico.fnal.gov/getFile.py/access?contribId=1&resId=0&materialId=slides&confId=12396>.

-
- [146] S. Derenzo et al., Nuclear Instruments and Methods **122**, 319 (1974). 1
- [147] 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. 2
3
4
- [148] E. Shibamura et al., Phys. Rev. A **20**, 2547 (1979). 5
- [149] V. M. Atrazhev and I. V. Timoshkin, IEEE Trans. Dielectrics and Electrical Insulation **5**, 450 (1998). 6
7
- [150] Y. Li et al., Nucl. Instrum. Meth. **A816**, 160 (2016). 8
- [151] D. Brailsford, *Looking at electron diffusion in 35t data and MC*, 35 ton working group meeting (2016), URL <https://indico.fnal.gov/getFile.py/access?contribId=43&sessionId=19&resId=0&materialId=slides&confId=10613>. 9
10
11
- [152] C. Bromberg et al., Journal of Instrumentation **10**, P07015 (2015). 12
- [153] M. Richardson, Ph.D. thesis, The University of Sheffield (2016). 13
- [154] N. S. Bowden, Tech. Rep. LBNE-doc-6493 (2012), URL <http://lbne2-docdb.fnal.gov:8080/cgi-bin/RetrieveFile?docid=6493&filename=LBNENearSurfaceReviewrReports.pdf&version=2>. 14
15
16
- [155] D. Wright, *Shielding Physics List Description* (2012), URL http://www.slac.stanford.edu/comp/physics/geant4/slac_physics_lists/shielding/physlistdoc.html. 17
18
- [156] B. Baller, Tech. Rep. LBNE-doc-3383 (2015), URL http://lbne2-docdb.fnal.gov:8080/cgi-bin/RetrieveFile?docid=3383&filename=ParametersAPAsOutsideAug11_2015.pdf&version=59. 19
20
21
- [157] D. Cherdack, Tech. Rep. LBNE-doc-7806 (2013), URL http://lbne2-docdb.fnal.gov:8080/cgi-bin/RetrieveFile?docid=7806&filename=SoftwarePreMeeting_Talk_DDC.pdf&version=1. 22
23
24
- [158] D. Barker et al., *Muon-induced background for beam neutrinos at the surface*, IBNE Reconfiguration Report (2012), URL http://www.fnal.gov/directorate/lbne_reconfiguration/files/LBNE-Reconfiguration-Surface-Detector-August2012.pdf. 25
26
27
- [159] D. Mei et al., Tech. Rep. LBNE-doc-6621 (2012), URL <http://lbne2-docdb.fnal.gov:8080/cgi-bin/RetrieveFile?docid=6621&filename=CosmogenicsSimulationOverview.pdf&version=1>. 28
29
30
- [160] D. Coelho et al., Tech. Rep. LBNE-doc-8458 (2012), URL http://lbne2-docdb.fnal.gov:8080/cgi-bin/RetrieveFile?docid=8458&filename=DC_LBNE_FEB1.pdf&version=1. 31
32
33
- [161] A. Hitachi et al., Phys. Rev. B **27**, 5279 (1983). 34
- [162] T. K. Gaisser, *Cosmic rays and particle physics* (Cambridge, UK: Univ. Pr. (1990) 279 p, 1990). 35
36

-
- ¹ [163] M. Aglietta et al. (LVD Collaboration), Phys. Rev. D **58**, 092005 (1998).
- ² [164] A. Dar, Phys. Rev. Lett. **51**, 227 (1983).
- ³ [165] T. Blackwell et al., Tech. Rep. LBNE-doc-7517 (2013), URL <http://lbne2-docdb.fnal.gov:8080/cgi-bin/RetrieveFile?docid=7517&filename=cosmogenics-vk1-130712-v3.pdf&version=3>.
- ⁶ [166] G. J. Feldman and R. D. Cousins, Phys. Rev. D **57**, 3873 (1998).
- ⁷ [167] D. M. Mei et al., Astropart. Phys. **34**, 33 (2010).
- ⁸ [168] F. E. Gray et al., Nucl. Instrum. Meth. **A638**, 63 (2011).
- ⁹ [169] J. Heise, AIP Conf. Proc. **1604**, 331 (2014).
- ¹⁰ [170] M. L. Cherry et al., Phys. Rev. D **27**, 1444 (1983).
- ¹¹ [171] N. Abgrall et al. (MAJORANA) (2016), arXiv:1602.07742.
- ¹² [172] E. Church, Tech. Rep. DUNE-doc-1644 (2016), URL <http://docs.dunescience.org:8080/cgi-bin>ShowDocument?docid=1644>.
- ¹⁴ [173] de Jong, J. K., Tech. Rep. LBNE-doc-3144 (2012), URL <http://lbne2-docdb.fnal.gov/cgi-bin>ShowDocument?docid=3144>.
- ¹⁶ [174] V. A. Kudryavtsev et al., Tech. Rep. LBNE-doc-9673 (2014), URL <http://lbne2-docdb.fnal.gov/cgi-bin>ShowDocument?docid=9673>.
- ¹⁸ [175] M. Robinson, *Cosmogenic background to nucelon decay*, Cosmogenics working group meeting (2017), URL <https://indico.fnal.gov/getFile.py/access?contribId=104&sessionId=14&resId=0&materialId=slides&confId=10641>.
- ²¹ [176] D. Stefan and A. M. Ankowski, Acta Phys. Polon. **B40**, 671 (2009).
- ²² [177] K. Wood, *Reconstruction Efficiencies of Scintillation Light from PDK events*, Cosmogenics working group meeting (2016), URL <https://indico.fnal.gov/getFile.py/access?contribId=98&sessionId=25&resId=0&materialId=slides&confId=10613>.
- ²⁵ [178] A. Higuera, *Proton Decay at DUNE $p \rightarrow K^+ + \bar{\nu}_e$* , cosmogenics working group meeting (2017), URL <https://indico.fnal.gov/getFile.py/access?contribId=1&resId=0&materialId=slides&confId=13789>.
- ²⁸ [179] T. Yang, *Updates on reconstruction*, Cosmogenics working group meeting (2017), URL <https://indico.fnal.gov/getFile.py/access?contribId=2&resId=0&materialId=slides&confId=13789>.
- ³¹ [180] A. Higuera, *Proton Decay at DUNE $p \rightarrow K^+ + \bar{\nu}_e$ and ${}^{39}\text{Ar}$* , Cosmogenics working group meeting (2017), URL <https://indico.fnal.gov/getFile.py/access?contribId=3&resId=0&materialId=slides&confId=13947>.

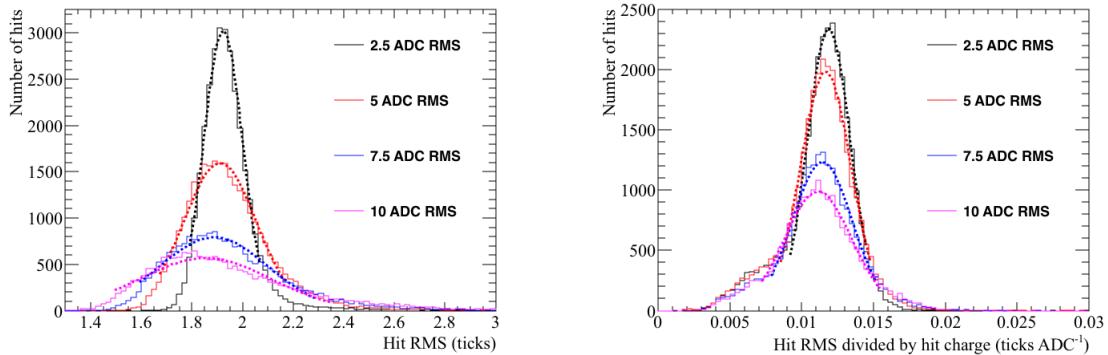
Appendix A

Supporting figures to Monte Carlo studies concerning determining interaction times using the effects of diffusion

Figure A.1 shows how the distributions of the hit RMS and hit $RMS/Charge$ change as the electronics noise increases, for hits between 20 and 30 cm from the APAs. Figure A.2, shows how the most probable values of hit RMS changes with drift distance for tracks associated with counter differences of 4, for different values of the electronics noise. Figure A.3, shows how the most probable value of hit RMS next to the APAs changes with increasing counter difference. Figure A.4, shows the normalised hit charge distributions for increasing noise levels, and the cut which is applied to remove the tails of the distribution.

Figure A.5, shows how the distributions of the hit RMS and hit $RMS/Charge$ change as the electron lifetime increases, for hits between 20 and 30 cm from the APAs. Figure A.6, shows how the most probable values of hit RMS change with drift distance for tracks associated with counter differences of 4, for different values of the electron lifetime. Figure A.7, shows how the most probable value of hit RMS next to the APAs changes with increasing counter difference. Figure A.8, shows the normalised hit charge distributions for increasing electron lifetimes, and the cut which is applied to remove the tails of the distribution.

Figure A.9, shows how the distributions of the hit RMS and hit $RMS/Charge$ change as the electric field increases, for hits between 20 and 30 cm from the APAs. Figure A.10, shows



(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.

1 how the most probable values of hit *RMS* changes with drift distance for tracks associated
 2 with counter differences of 4, for different values of the electric field. Figure A.11, shows
 3 how the most probable value of hit *RMS* next to the APAs changes with increasing counter
 4 difference. Figure A.12, shows the normalised hit charge distributions for increasing electric
 5 fields, and the cut which is applied to remove the tails of the distribution.

6
 7 Figure A.13, shows how the distributions of the hit *RMS* and hit *RMS/Charge* change as
 8 the constant of longitudinal diffusion increases, for hits between 20 and 30 cm from the APAs.
 9 Figure A.14, shows how the most probable values of hit *RMS* changes with drift distance
 10 for tracks associated with counter differences of 4, for different values of the constant of
 11 longitudinal diffusion. Figure A.15, shows how the most probable value of hit *RMS* next to
 12 the APAs changes with increasing counter difference. Figure A.16, shows the normalised hit
 13 charge distributions for increasing values of the constant of longitudinal diffusion, and the
 14 cut which is applied to remove the tails of the distribution.

15

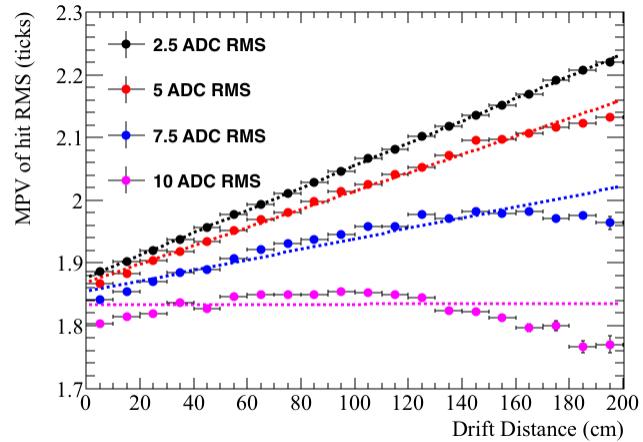


Fig. A.2 The most probable values of hit *RMS* as a function of drift distance, for tracks associated with a coincidence that had a counter difference of 4, for different values of the electronics noise.

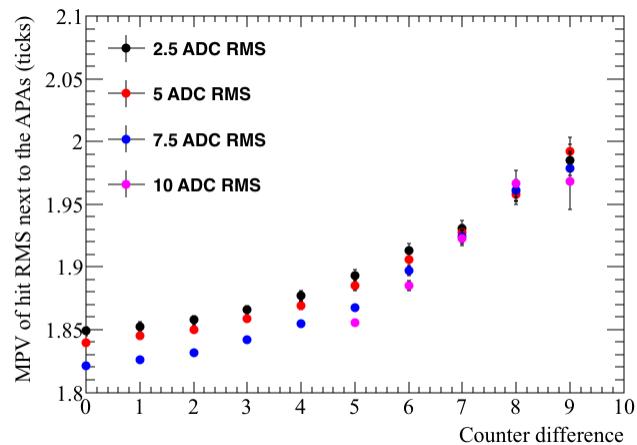


Fig. A.3 The most probable values of hit *RMS* within 10 cm of the APAs, as a function of the counter difference of the coincidence that the track was associated with, for different values of the electronics noise.

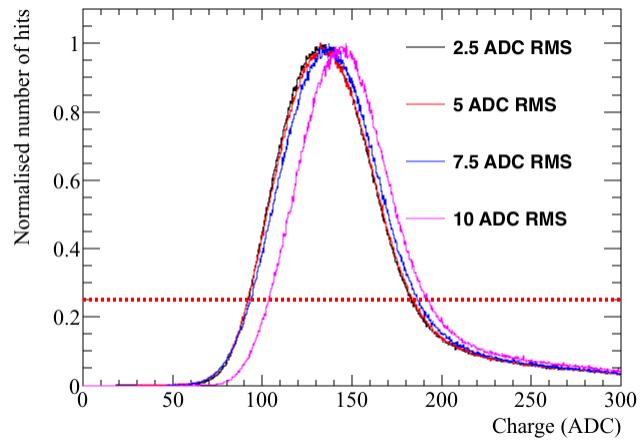
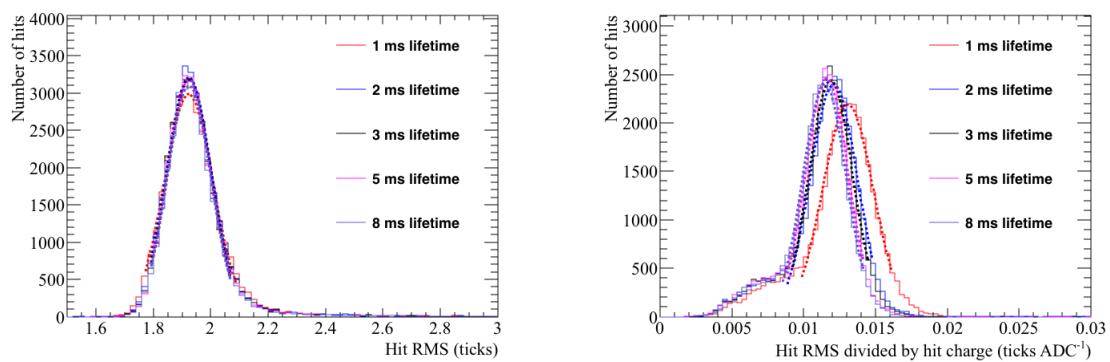


Fig. A.4 The normalised hit charge distribution for different values of the electronics noise. The hit charge is shown in units of ADC, and is normalised so that the most common hit charge has a value of 1. A cut on the normalised number of hits being greater than 0.25 is shown, the aim of this cut is 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.

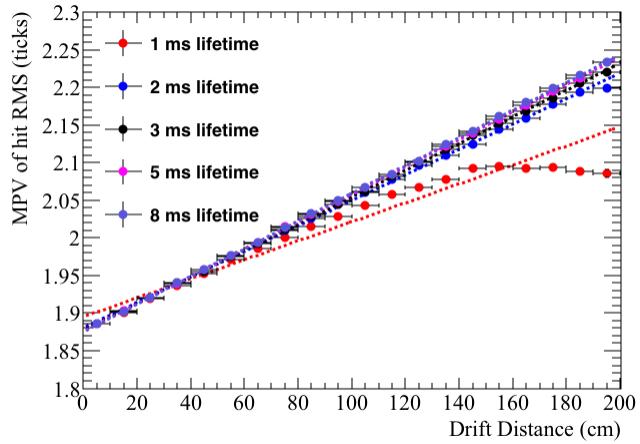


Fig. A.6 The most probable values of hit *RMS* as a function of drift distance, for tracks associated with a coincidence that had a counter difference of 4, for different values of the electron lifetime.

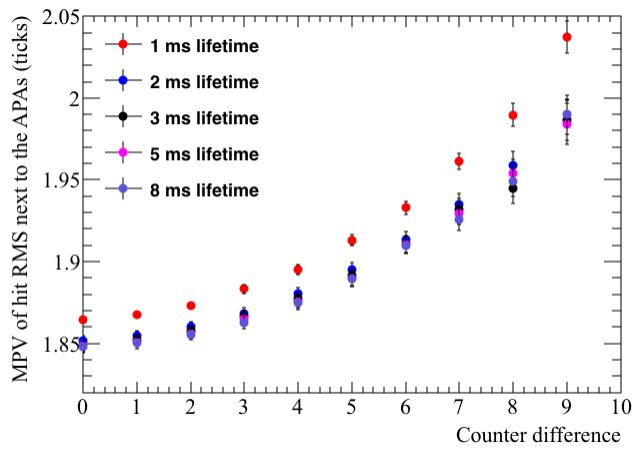


Fig. A.7 The most probable values of hit *RMS* within 10 cm of the APAs, as a function of the counter difference of the coincidence that the track was associated with, for different values of the electron lifetime.

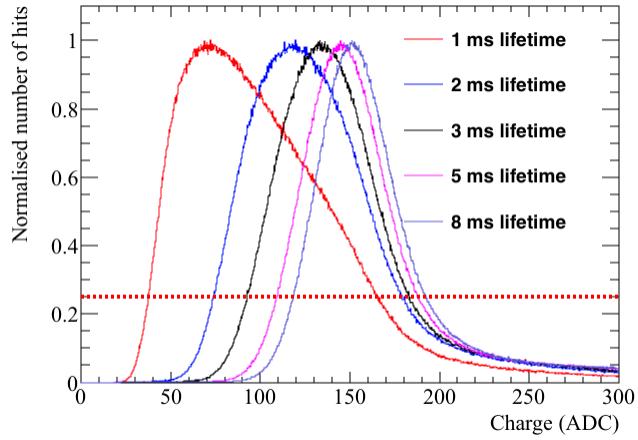
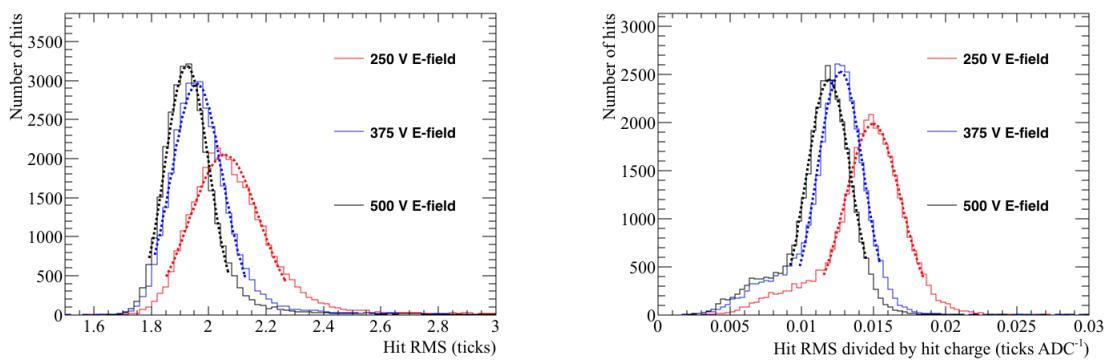


Fig. A.8 The normalised hit charge distribution for different values of the electron lifetime. The hit charge is shown in units of ADC, and is normalised so that the most common hit charge has a value of 1. A cut on the normalised number of hits being greater than 0.25 is shown, the aim of this cut is 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.

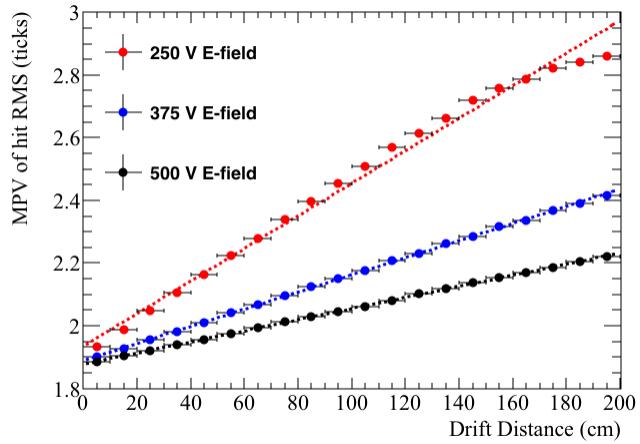


Fig. A.10 The most probable values of hit *RMS* as a function of drift distance, for tracks associated with a coincidence that had a counter difference of 4, for different values of the electric field.

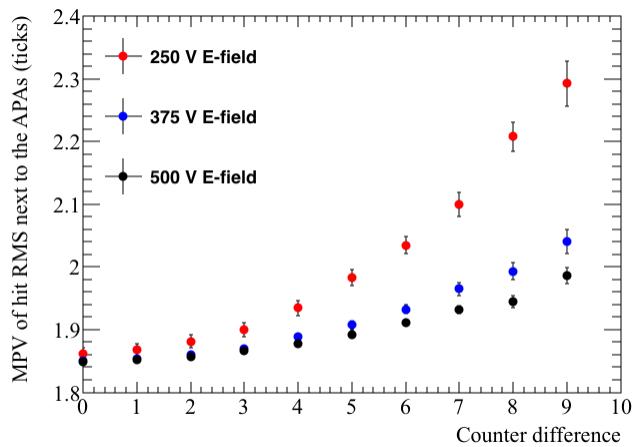


Fig. A.11 The most probable values of hit *RMS* within 10 cm of the APAs, as a function of the counter difference of the coincidence that the track was associated with, for different values of the electric field.

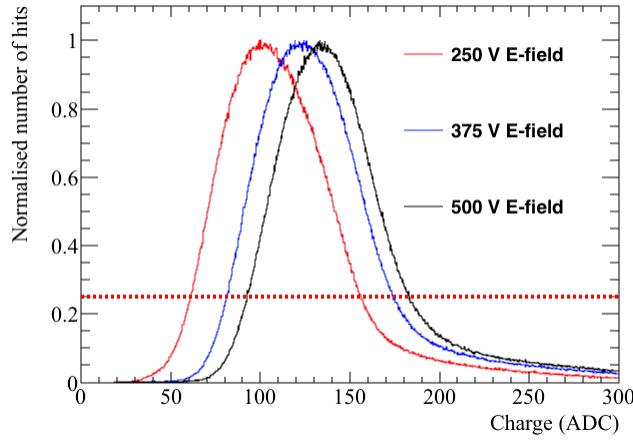
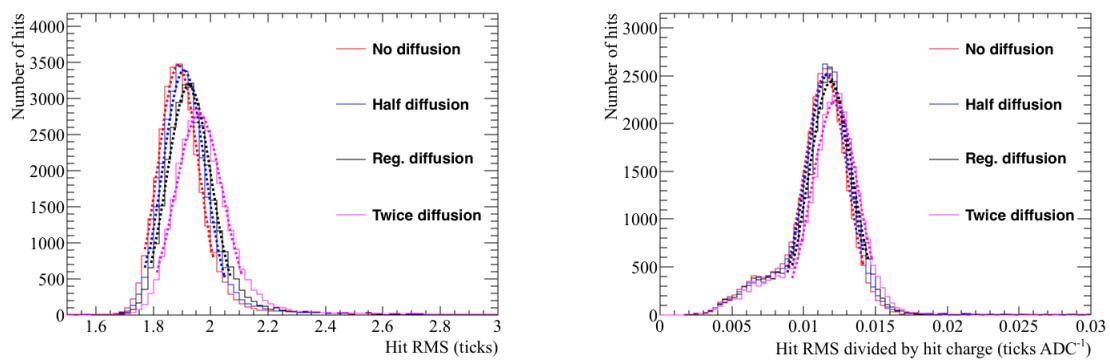


Fig. A.12 The normalised hit charge distribution for different values of the electric field. The hit charge is shown in units of ADC, and is normalised so that the most common hit charge has a value of 1. A cut on the normalised number of hits being greater than 0.25 is shown, the aim of this cut is 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 probable 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.

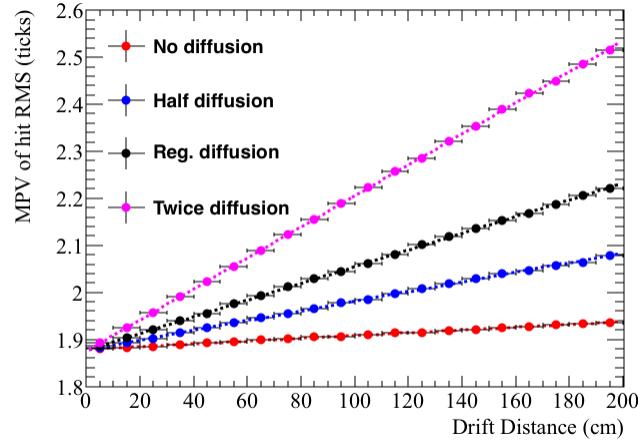


Fig. A.14 The most probable values of hit *RMS* as a function of drift distance, for tracks associated with a coincidence that had a counter difference of 4, for different values of the constant of longitudinal diffusion.

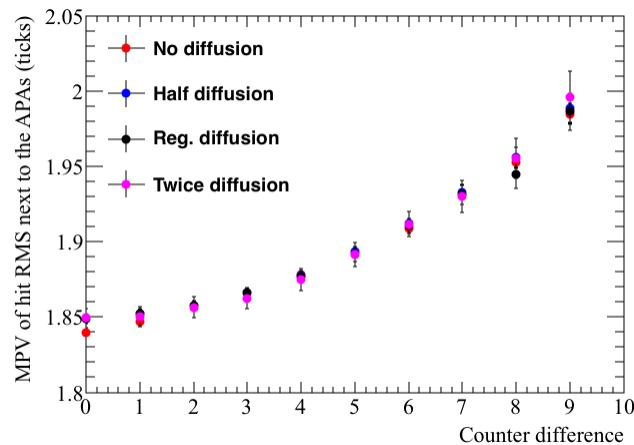


Fig. A.15 The most probable values of hit *RMS* within 10 cm of the APAs, as a function of the counter difference of the coincidence that the track was associated with, for different values of the constant of longitudinal diffusion.

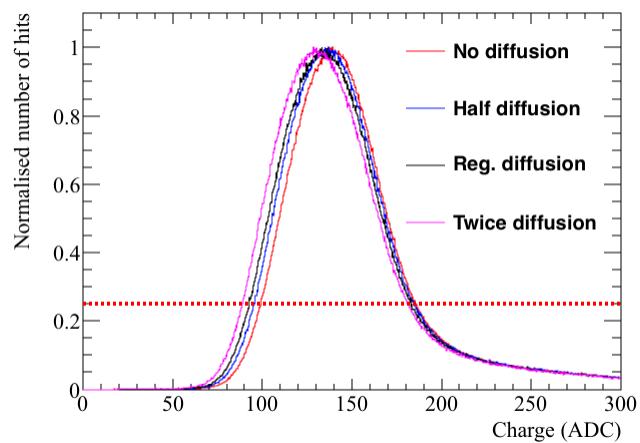


Fig. A.16 The normalised hit charge distribution for different values of the constant of longitudinal diffusion. The hit charge is shown in units of ADC, and is normalised so that the most common hit charge has a value of 1. A cut on the normalised number of hits being greater than 0.25 is shown, the aim of this cut is to remove the tails of the hit charge distribution.