

Background Mitigation in Dual Phase Xenon Time Projection Chambers



The
University
Of
Sheffield.

Peter Lyle Rossiter

Supervisor: Vitaly Kudryavtsev

Department of Physics & Astronomy
University of Sheffield

This dissertation is submitted for the degree of
Doctor of Philosophy

May 2021

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation 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 dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements.

Peter Lyle Rossiter
May 2021

Acknowledgements

There have been many people who have accompanied me along this PhD journey, and many more who have been exceptionally helpful and supportive during specific stretches. Without a doubt, this work would not be what it is if not for each and every one of them, and I am immensely grateful to more people than I could possibly mention. But if I were to try and thank them all I would certainly begin with my supervisor, Vitaly Kudryavtsev. He not only gave me this fantastic opportunity, but his patience, professionalism, and guidance ensured I got the most from the experience. This extends to our group here in Sheffield of Elena, Andrew and, before he accepted a post-doc at Penn State, David. At various times each of you has shared your knowledge, experience, and know-how with me; I can't imagine what my thesis would have looked like in the end without the input of each of you.

A big thank you also needs to be extended to everyone in the LZ and LUX collaborations for their tireless work, and their commitment to a welcoming culture of mutual respect. In particular, I'd like to thank Amy and Ben for insights they've shared at critical junctures in my research with LZ; and Claudio for helping me to find my feet and a direction within the LUX collaboration.

Given the amount of time I spent in the Bay Area I would like to express my gratitude towards Kevin Lesko for facilitating a place for me to work with his team at LBNL; and also to Evan, Kelsey, Quentin, and everyone else there for their help, support, and warm welcomes. I would also like to thank Greg who was a pleasure to work with on the LUX gamma-X analysis; your speedy simulations certainly saved me a lot of time towards the end of my PhD.

To my colleagues in Sheffield Tom, Dom, Warren, Scarfy, Viktor, and Jordan; thank you all for contributing towards the sense of community and solidarity I felt, particularly while writing my thesis during lockdown. Also for your invaluable proofreading!

Thank you too to Joe, Maddie, Pedro, Damien, Claire, Bernie, and Mr. Steam Roller; your various stylistic, writing, and formatting tips were just what I needed during the final

stages of thesis drafting. My use of semicolons may not be perfect yet; but, it is much better now. Thanks also to Christine who for her support, encouragement, and unique perspective.

Finally, I'd like to finish by thanking my family for their love and faith in me over the years. In particular, I want to thank my Dad for his tons, and tonnes, of proofreading; and my Mum, whose capacity to face life's challenges with courage and good humour is a constant source of inspiration to me.

Abstract

The focus of this thesis is on the identification and removal of background events in the dual-phase xenon TPC experiments of LUX and LZ. In particular this effort focuses on a few hard to classify background types which appear in the traditional WIMP search and in the EFT search. The first half of the original research presented in this thesis pertains to backgrounds expected to be present in LZ. This includes multiple scatter events which only deposit a proportion of their energy in the LXe, and multiple scatter events which only produce a single ionisation signal. The thesis then goes on to examine how backgrounds such as these, and other backgrounds occurring close to the edge of the LXe can be reduced with an improved modelling of ionisation electron paths in the LXe, and describes how this modelling was used by the LZ collaboration. The second half of this thesis focuses exclusively on these multiple scatter events which only produce a single ionisation signal, but in the context of LUX data. It begins by examining how previous studies have identified and removed them; then presents a novel method for the identification and removal of these events using a machine learning algorithm called a BDT. Finally the thesis concludes after using data and simulations to compare this new BDT cut against previously developed cuts for this class of event, and shows this new BDT cut to be the best option for the ongoing EFT analysis of LUX.

Table of contents

List of figures	xvii
List of tables	xxv
Nomenclature	xxvii
1 Introduction	1
2 Dark matter	3
2.1 Introduction	3
2.2 Evidence for dark matter	3
2.2.1 Early dark matter history	3
2.2.2 Galactic rotation curves	4
2.2.3 Gravitational lensing	5
2.2.4 Cosmic microwave background radiation	7
2.2.5 Structure formation	9
2.3 Dark matter landscape	10
2.3.1 Dark matter properties	10
2.3.2 Standard model dark matter	10
2.3.3 WIMPs	11
2.3.4 Axions	13
2.4 Direct detection of dark matter	14
2.4.1 Concept	14
2.4.2 Interaction rate	15
2.4.3 WIMP halo	15
2.4.4 Nuclear scattering cross-section	16
2.4.5 Nuclear form factor	17
2.4.6 Detection signals and technology	18
2.4.7 Status of field	21

2.5	Other search methods	23
2.5.1	Collider searches	23
2.5.2	Indirect detection	24
2.6	Effective field theory of dark matter detection	25
2.7	Summary	28
3	The LUX and LZ experiments	31
3.1	Dual-phase xenon time projection chambers	31
3.1.1	Concept	31
3.1.2	Ionisation and scintillation signals	32
3.1.3	$S1$ production	32
3.1.4	$S2$ production	35
3.1.5	Energy reconstruction	36
3.1.6	Electron recoil event discrimination	36
3.1.7	Xenon self-shielding	37
3.2	The LUX experiment	39
3.2.1	Overview	39
3.2.2	Timeline and results	40
3.2.3	Electric field	40
3.3	The LZ experiment	41
3.3.1	Overview	41
3.3.2	Outer detector	42
3.3.3	LXe skin region	43
3.3.4	LZap	43
3.3.5	Mock data challenges	44
3.4	Backgrounds	45
3.4.1	Cosmogenic backgrounds	45
3.4.2	Intrinsic xenon backgrounds	46
3.4.3	Detector material backgrounds	47
3.5	Background modelling	48
3.5.1	Overview	48
3.5.2	LUXSim and BACCARAT	48
3.5.3	NEST	49
3.5.4	LLAMA	49
3.6	Calibration for position reconstruction	50
3.7	Data analysis cuts	51
3.7.1	Region of Interest	53

3.7.2	Multiple scatter cut	53
3.7.3	Fiducial volume cut	53
3.7.4	LXe skin cut	54
3.7.5	Outer detector cut	54
4	^{131m}Xe simulation studies for LZ	55
4.1	^{131m}Xe background production mechanisms	55
4.2	Metastable ^{131}Xe	56
4.3	^{131m}Xe emissions	56
4.4	^{131m}Xe Monte Carlo Generator	58
4.4.1	Primary particle energy validation	58
4.4.2	Primary particle spatial distribution	61
4.5	^{131m}Xe simulation	61
4.6	Partial Energy Deposition Analysis	62
4.7	Multiple Scatter Single Ionisation Analysis	66
4.8	Conclusion	70
5	The Effect of Electric Field Non-Uniformity on Position Reconstruction in LZ	73
5.1	Rationale	73
5.2	Electric field in LZ	74
5.3	Electron trajectory in LZ	75
5.4	The Electron Drifter Module	77
5.5	ED module analyses	80
5.5.1	LZ grids analysis	80
5.5.2	Diffusion analysis	82
5.6	MDC2 application	85
5.6.1	Rationale	85
5.6.2	Electron drift in BACCARAT	85
5.6.3	LZap XYZ Position Corrector Module	86
5.6.4	Validation	89
5.7	Conclusion	91
6	Gamma-X event simulation for LUX	93
6.1	Rationale	93
6.2	Gamma-X events	94
6.2.1	Description	94
6.2.2	Properties	98

6.2.3	Sources	98
6.3	History of gamma-X cuts in LUX	101
6.3.1	Background	101
6.3.2	WIMP search gamma-X cut	101
6.3.3	Double electron capture gamma-X cut	102
6.3.4	EFT search leakage cut	104
6.3.5	Similarities, differences, and expected performance	105
6.4	Run04 EFT data analysis cuts	106
6.5	Gamma-X event simulation	109
6.5.1	Simulation process	109
6.5.2	Simulation tuning	111
6.5.3	Simulation chain validation	112
6.5.4	Source selection for gamma-X simulations	115
6.5.5	Gamma-X simulations from the bottom PMT array	117
6.5.6	Comparison of multiple scatter simulations with data	119
6.6	Gamma-X feature search	121
6.7	Conclusion	126
7	Identification of gamma-X events with a BDT in LUX	129
7.1	Aim	129
7.2	The XGBoost Boosted Decision Tree	129
7.2.1	Overview	129
7.2.2	BDT building blocks	130
7.2.3	Growing a decision tree	132
7.2.4	Decision tree pruning	135
7.2.5	Boosting	136
7.2.6	Decision tree stopping conditions	137
7.3	BDT training event selection	138
7.3.1	Rationale	138
7.3.2	Toy Model simulations	139
7.3.3	Overtraining mitigation	142
7.3.4	BDT performance	144
7.3.5	BDT stability	145
7.4	Hyperparameter selection	147
7.4.1	Purpose of hyperparameter optimisation	147
7.4.2	Coarse N and D simulations grid search	148
7.4.3	Fine N and D simulations grid search	149

7.4.4	<i>N</i> and <i>D</i> selection	150
7.4.5	Learning rate	151
7.4.6	Other hyperparameters and final hyperparameter selection	152
7.5	BDT cut development with tritium and ^{14}C data	153
7.6	Conclusion	156
8	Gamma-X cut performance on LUX data	159
8.1	Cut benchmarking	159
8.2	Gamma-X cut performance on Run04 EFT data	164
8.3	Conclusion	167
9	Concluding remarks	169
	References	171

List of figures

2.1	Galactic rotation curve for NGC 6503. Dark matter halo (dash-dotted line), optical disk (dashed line), and gas (dotted line) contributions used to match the data are indicated.	5
2.2	A gravitational lensing of a distant blue galaxy surrounding the red galaxy LRG 3-757. Light from the distant galaxy is distorted as it passes through the strong gravitational field, pulling the image of the distant blue galaxy into a ring-like shape.	6
2.3	Dark matter in the Bullet Cluster. This figure has been artificially coloured to highlight the separation of the interstellar gas (in pink) and the gravitational centre of mass (blue).	7
2.4	The all-sky map of the CMB radiation temperature variations as detected by the Planck satellite. The differences in temperature are highlighted on this map, with the coldest regions shown in blue, and the hottest regions in red. The average temperature is 2.7 K, and the range of temperature variation is 0.0006 K.	8
2.5	The CMB power spectrum shows the relationship between the angular size of CMB anisotropies (lower horizontal axis) and their temperature fluctuation. Multipole moment ℓ is shown in the upper horizontal axis. Note that a scale change on the horizontal axis from logarithmic to linear exists at $\ell=500$. Data (red dots) is measured by the Planck collaboration. The green line shows the prediction from the Λ CDM model.	9
2.6	Predicted WIMP event rate based on XENON1T nucleon cross-section limits, for various target materials and WIMP masses. (a) shows event rate in ^{131}Xe for three WIMP masses: 20 GeV, 35 GeV, and 50 GeV. (b) shows event rate for a 100 GeV WIMP in three target materials: ^{40}Ar , ^{72}Ge , and ^{131}Xe	18
2.7	Nuclear form factor for ^{40}Ar , ^{72}Ge , and ^{131}Xe	19

2.8	Limits on spin-independent WIMP-nucleon cross-section as a function of WIMP mass.	22
2.9	WIMP-nucleon recoil spectra in xenon for a representative sample of EFT operators, weighted by the natural abundance of isotopes.	29
3.1	Operating principle of a dual-phase Xe TPC.	33
3.2	Discrimination parameter $\log_{10}(S2/S1)$ as a function of $S1$ obtained with LUX calibration data.	38
3.3	The LUX TPC. In this figure, the locations of PMT arrays, field rings, shielding, and the grids, are explicitly indicated. The Inner Cryostat Vessel (ICV) lid is included in the illustration on the right.	39
3.4	LUX wall position based on ^{83m}Kr data. The plot on the left shows Run03 data (before grid conditioning). The plot on the right shows various wall positions through 2014-2015 (after grid conditioning).	41
3.5	Overview of the LZ experiment showing the main components. Key differences to LUX are the size of the TPC, the LXe skin, and a liquid scintillator veto placed inside the water tank and the cathode high voltage connection.	42
3.6	A 1.5 keV ER event taken from a three month above ground commissioning of LUX. The image shows a five-fold PMT hit coincidence for $S1$, and the corresponding, and larger, $S2$ delayed by 20 μs	44
3.7	The red-black heat map shows where simulated mock data taken from LZ's MDC2 are proportionally located in $S2$ vs. $S1$ space. The blue line represents the reconstructed energy of the calibration source ^{131m}Xe ($E = 163.93$ keV), using Eq. 3.11 with: $g_1 = 0.128$, $g_2 = 92.26$ (a higher value of g_2 compared to past experiments [1]). Also visible in this plot are a suite of other background events with low $S1$ (as seen on the lower left corner).	52
4.1	A decay mode for ^{131m}Xe	59
4.2	Spectral output ^{131m}Xe . In red are the photon peaks (excluding the 163.93 keV peak), and in blue are the electron peaks. Intensity is given on a logarithmic scale.	60
4.3	Generator validation plot.	60

4.4	Simulated primary particle initial positions for 10^5 events. Particles are randomly distributed throughout all the LXe volumes which the generator was called in. Volume encircled in red represents the sensitive LXe target. In green is the skin. The cathode position is represented with the yellow line at $Z = 0$. A small contribution can also be seen coming from the LXe in the high voltage feed through.	62
4.5	Photon mean free path in LXe of density 2.88 g/cm^3	63
4.6	Position of all simulated $^{131\text{m}}\text{Xe}$ partial energy depositions. All events are removed by the fiducial cut. The central plot shows all events in R - Z space. The plot in the horizontal margin shows a close up of events near the wall. The plot in the vertical margin shows a close up of events near the phase boundary.	66
4.7	Position of MSSSI backgrounds originating from $^{131\text{m}}\text{Xe}$ events. Figure (a) shows the R - Z distribution of MSSSI events after MS cuts. Figure (b) shows the distribution of events after the MS cut and radial fiducial cut.	68
4.8	$^{131\text{m}}\text{Xe}$ MSSSI events in S_2 vs. S_1 space. In blue are all simulated MSSSI events, with the final background selection for the EFT search shown in black. The green region and magenta regions represent the EFT and standard WIMP search energy ranges respectively. The lack of black events in the magenta region indicates no known $^{131\text{m}}\text{Xe}$ induced MSSSI background for the standard WIMP search.	69
5.1	Electric field strength in LZ fiducial region.	75
5.2	Electron drift velocity in LXe.	76
5.3	Diagram of an electron drifting in an electric field for time Δt . Electron begins at the coordinates (R_i, Z_i) , and finish at (R_f, Z_f) . This diagram represents a single step in the ED module algorithm described in this section.	79
5.4	Electron trajectory in LZ field as indicated by a series of drift lines.	80

- 5.5 LZ's radial electric field (blue-yellow heat map) and field lines (white lines) near the grids and cathode. Figure (a) shows the region around the grids at the top of the TPC. The lower plot in (a) shows a heat map of the electric field strength, with yellow/green hot spots representing the high voltage of the grid wires. Field lines are overlaid on this heat map, with each field line represented in white. These field lines were generated with the ED module originating at $Z = 145$ cm with a radial spacing of 0.02 cm and terminate at $Z = 146.1$ cm. These field lines can be seen to converge in on one another as they pass through the grid. The hit pattern of electrons generated at $Z = 145$ cm with a tighter radial spacing of 0.001 cm can be seen in the histogram directly above. Figure (b) shows a similar analysis conducted around the cathode. This plot however uses a radial field line density of 0.01 cm, and the upper histogram shows the radial position at $Z = 2$ cm. 82
- 5.6 Comparison between the extracted position of electrons with diffusion turned on and off. The blue histogram in each plot represents the extracted position of electrons with diffusion turned on. The dashed red line is the Gaussian obtained by calculating the mean and standard deviation of the blue distribution. The dashed black line is the parameterised calculation. 83
- 5.7 TPC map showing the effect of the electric field in LZ. The heat map shows the difference in $R_{EXTRACTED}$ and R_{ORIGIN} for each point in the LXe. The dashed aqua lines show the trajectory of selected electrons to the phase boundary. The fiducial volume is displayed, outlined in dashed white lines. 86
- 5.8 Position of all simulated ^{131m}Xe PED events from Chapter 4 offset by the MDC2 drift map used by BACCARAT. Central plot shows all events in R - Z space which have clearly experienced an inward shift which scales with depth. The two histograms in the margins shows the same events in a single dimension. 87
- 5.9 The difference between reconstructed and true positions of events as a function of their position in the TPC. (a) shows the difference between $R_{RECONSTRUCTED}$ and R_{ORIGIN} , and (b) shows the same for Z . The dashed green line in both figures represents the fiducial region. 89
- 5.10 The difference between reconstructed and true positions for all events. (a) represents this difference in R , and (b) shows it for Z 90

5.11	All non-fiducial events which are reconstructed inside the fiducial region. In both figures, the dashed green line represents the fiducial region, and the magenta dots the event positions. (a) shows the true position of these events, and (b) shows their reconstructed positions.	91
6.1	A schematic of a gamma-X event. This example illustrates a gamma-ray originating from the bottom PMT array (gamma-ray displayed as a bold black arrow, and the PMT array as grey blocks). The first scatter (represented in yellow) is below the cathode (therefore has its ionisation signal lost), and the second is in the fiducial region. For positional reference the height above the cathode is listed for various detector components in the LUX TPC.	95
6.2	An example of a suspected gamma-X event in the Run04 data, showing the phd for $S1$ and $S2$ on the Y axes, and time in Samples on the X (1 Sample = 10 ns). The overall event summary can be seen in the grey box, located in the lower centre of the figure. The upper left box in blue is a close up of the $S1$ signal. The red box on the upper right is a close up of the $S2$ signal. The tight clustering of the $S1$ signal in the bottom PMT array and long drift time are both typical features of a gamma-X event.	97
6.3	Measured gamma-ray spectrum in LUX during Run03	99
6.4	Definition of the CS_{S1} gamma-X cut used by LUX Run04 WIMP search	103
6.5	Gamma-X cut developed for the LUX double electron capture analysis. The heat map shows ^{14}C calibration data in $(S1, MPA_{S1})$ space. The red line overlaying this data is the cut in question, which was developed with a linear interpolation of 99% data acceptance bins in $S1$. Bins are defined by a width of 50 phd. Events above the red line are classified as gamma-X-like by this cut and removed.	104
107	figure.caption.52	
6.7	Simulation chain used for producing gamma-X events.	110
6.8	Tuning of raw simulations based on comparison with Run04 $^{83\text{m}}\text{Kr}$ data. For each variable listed on the vertical axis, data and simulations were compared in 20 different Z bins. The mean of each simulated bin is shown in turquoise, along with a 1σ band. Plots on the left show the simulated bands before tuning overlaid on data. Plots on the right show simulated bands after tuning.	113

6.9	A comparison of simulated and real $^{83\text{m}}\text{Kr}$ data for LUX Run04. To ease visual comparison all plots display one dataset as a heatmap, with the other represented as a contour plot overlaid on top. Regardless of whether data or simulations are represented by the heatmap - simulations are always shown in blue, and data in red. The top left plot displays $S1$ vs CS_{S1} , the top right displays $S1$ vs $\log(\text{MPAF})$, and the middle right plot displays CS_{S1} vs $\log(\text{MPAF})$. The remaining three plots show the same datasets, but with the axes flipped, and data as a contour plot laid over the simulations.	114
6.10	Simulated and real nearGX events. (a) the number of MS events which occur within 3.8 cm of the cathode, (b) average $S2$ separation in Z as a function of the height above the cathode for data and simulations.	121
6.11	Potential parameters for a gamma-X cut. This figure examines the parameter spaces of $S1$, $S2$, and Z by plotting three distributions side by side: $^{83\text{m}}\text{Kr}$ data as a red-black heat map, tritium as a blue-black heat map, and the simulated gamma-X events as a yellow scatter plot.	123
6.12	Potential parameters for a gamma-X cut. This figure examines some of the parameter spaces of $S1$, $S2$, and Z , TB ratio, CS_{S1} and MPAF_{S1} by plotting three distributions side by side: $^{83\text{m}}\text{Kr}$ data as a red-black heat map, tritium as a blue-black heat map, and the simulated gamma-X events as a yellow scatter plot.	124
6.13	Potential parameters for a gamma-X cut. This figure examines some of the parameter spaces of TB ratio, CS_{S1} and MPAF_{S1} by plotting three distributions side by side: $^{83\text{m}}\text{Kr}$ data as a red-black heat map, tritium as a blue-black heat map, and the simulated gamma-X events as a yellow scatter plot. . . .	125
7.1	Root tree from the BDT used for gamma-X classification. The tree has been simplified by removing most internal nodes (as indicated by the suspension periods). The names variables A - F are not relevant to the structure of this tree but are provided in Section 7.2.3.	131
7.2	Ordering and splitting of training data in the root node. All events (SS and gamma-X) are ordered based on a particular variable; Z in this case. The node then chooses a splitting value which maximises the <i>Gain</i> given by Eq. 7.3.	133
7.3	Toy Model validation	141

- 7.4 The Classifier Output Distribution used to assess BDT performance. Training events are given by the histograms, test events by the dots. Red indicates the event is gamma-X, blue is for SS. The dotted black line indicates the classification threshold. All events below this line are categorised as gamma-X. The BDT is trained to give events which appear more single scatter-like a score closer to 1, and events more gamma-X-like a score closer to 0. 144
- 7.5 BDT performance on a varying number of simulated events. The top figure shows how the average of 10 BDTs classification efficiencies varied as the number of simulated events increased. The error-bar represents 1σ on the variation of the mean represented in each bin. The lower figure shows the standard deviation of the upper plot. The number of simulations increased until this σ was below 0.001. 147
- 7.6 Figure (a) shows the variation of the classification score with the number of trees and the tree depth. Figure (b) shows how these two hyperparameters affect the p value measuring overtraining. This grid search was performed over a coarse range of N ranging from 50 - 200, at steps of 50, and D ranging from 5 - 20 with steps of 5. 149
- 7.7 (a) shows the variation of the classification score with the number of trees and tree depth. (b) shows how these two hyperparameters affect the p value measuring overtraining. This grid search was performed for all (N, D) combinations in the N range of 75 to 225, and the D range of 3 to 8. For visual clarity of the trend not all values for N are displayed. 150
- 7.8 This figure shows the effect of changing ϵ on the classification score (upper plot), and the p value measuring overtraining (lower plot). 151
- 7.9 The effect of altering the hyper parameters I_0 (left), γ (centre), and λ (right) on classification score (upper), and the p value measuring overtraining (lower). 152
- 7.10 The BDT score for a combination of tritium and ^{14}C calibration data. This plot was used to inform the position of the BDT cut which is displayed as a dashed black line. A cut value of 0.36 was chosen as this was the threshold which gave a 95% acceptance. 154
- 7.11 Efficiency curves for the BDT cut on a selection of tritium and ^{14}C calibration data. The upper plot in each figure shows all events in blue and removed events in red. The overall efficiency is shown below in black: (a) $S1$, (b) $S2$, (c) ϕ (azimuthal angle), and (d) drift time. 155

- 8.1 The rejection power of various cuts on 1S1 2S2 MS events near the cathode. In (a) the number of events removed by four different cuts is displayed for as the maximum height above the cathode required for nearGX selection is varied. The fraction of all events removed for each cut is displayed in (b). Cuts represented in this figure are: the BDT gamma-X cut (red dashed line), the CS_{S1} gamma-X cut (blue dashed line), the $MPAF_{S1}$ gamma-X cut (green dashed line), and the MPA_{S1} leakage cut (grey dashed line). The total number of events for each height above cathode selection criteria is also shown in (a) (thick black line). Note that the MPA_{S1} leakage cut is included in the data selection of each of the other cuts displayed. 160
- 8.2 Gamma-X cut benchmarking plot. In all sub-figures, blue-black heat maps represent tritium data, red-black represents ^{83m}Kr , and yellow scatter points represent simulated gamma-X events. The cut in (a) is the CS_{S1} gamma-X cut (solid green line) and removed all events below the curve. The cut in (b) is the $MPAF_{S1}$ gamma-X cut (dashed green line) and removes all events above the curve. In (c) tritium and ^{83m}Kr data are combined into one scatter plot (black) and shows events removed by the MPA_{S1} leakage cut in green. In (d) the BDT score used to inform the BDT cut is shown for the same tritium (blue) and ^{83m}Kr (red) data, and simulated gamma-X events (yellow). Events below the dashed black line are removed by the BDT gamma-X cut, so long as they are in the range: $Z < 20$ cm, and $S1 > 11$ phd. 162
- 8.3 BDT cut effect on Run04 EFT search data circa February 2020. In (a) data is displayed in (R^2, Z) space, with removed events in red. In (b) the same data is displayed in $(S1, S2)$ space, with the 1σ NR band displayed in magenta, and the 1σ ER band in blue. For comparison, the events removed by the $MPAF_{S1}$ cut are also displayed with green crosses. 165
- 8.4 BDT cut effect on salt and Run04 EFT search data. In (a) the BDT score of all Run04 EFT search data is displayed in grey, along with salt events in green, and the BDT cut as a dashed red line (events below the line are removed). In (b) are salt events (green) in $(S1, S2)$ space. Blue crosses represent events which would be removed by the CS_{S1} gamma-X cut. Red crosses show events with a BDT score < 0.36 . Note that these red events are not removed by the BDT cut on account of their low $S1$ value. 166

List of tables

3.1	Summary of backgrounds in LZ, showing the number of counts expected in 1,000 live days in an indicative 5.6-tonne fiducial mass in the region of interest with all cuts applied. A comprehensive set of numbers can be found in [2].	46
4.1	Emission spectra for $^{131\text{m}}\text{Xe}$ as it decays to ^{131}Xe . The various decay modes are formed by combinations of conversion electrons, Auger electrons, and X-rays. The probability of each mode along with the electron and photon energies are given here.	57
4.2	Result from PED background analysis with various data analysis cuts applied. The cut definition can be found in Section 3.7. The first column indicates which cuts have been applied for each of the WIMP and EFT searches. The second column shows how many PED events remain after each cut. The final column normalises these values to a 1000 day run assuming an average 0.02 Hz decay rate. WIMP search ROI is 1.5 keV to 6.0 keV. The EFT ROI extends up to 70 keV.	65
4.3	Result from MSSI background analysis with various data analysis cuts applied.	67
6.1	Run04 ER background model analysed for potential gamma-X events. . . .	118
6.2	Gamma-X simulation results. Four background sources from the bottom PMT array (left column) were simulated. The decay rate per PMT is shown in the second column, the total number of events in the third, the number of days this represents in the fourth, and the final column gives the total number of gamma-X events produced.	119
6.3	Confusion matrix for features with the potential for gamma-X identification.	126
7.1	Hyperparameter selection for the BDT	153

- 8.1 The fraction of rejected events of each cut on various datasets. Near gamma-X events are the same as from Figure 8.1 where the maximum height above cathode is set to 5 cm. Each cut was measured on datasets after all data analysis cuts from Section 6.4 were applied, except for the leakage cut, and the lower fiducial cut (the ROI cut was also abandoned on the simulated gamma-X events). 163
- 8.2 Various gamma-X cut performances on Run04 EFT data, and salt events. . . 166

Nomenclature

Acronyms / Abbreviations

CS_{S1} The cluster size of the S1-photo-hit-pattern in the bottom PMT array

$MPAF_{S1}$ The ratio between MPA_{S1} and S1 (a.k.a. Max Peak Area Fraction)

MPA_{S1} The amount of S1 light in the brightest PMT of the bottom PMT array (a.k.a. S1 Max Peak Area)

ADC Analog to Digital Converter

ArDM Argon Dark Matter - a single-phase liquid argon dark matter experiment

BACCARAT Basically a Component-Centric Analog Response to AnyThing

BAO Baryon Acoustic Oscillation

BDT Boosted Decision Tree

BooNE Booster Neutrino Experiment

BSM Beyond the Standard Model

CALET CALorimetric Electron Telescope

CDM Cold Dark Matter

CDMS Cryogenic Dark Matter Search

CL Confidence Level

CLEAN Cryogenic Low Energy Astrophysics with Noble liquids

CMB Cosmic Microwave Background

COD Classifier Output Distribution

COMSOL COMputer SOLution

CP Charge Parity

CRESST Cryogenic Rare Event Search with Superconducting Thermometers

DAMA DArk MATter

DAMPE Dark Matter Particle Explorer

DARWIN DArk matter Wimp search wIth Noble liquids

DD Deuterium-Deuterium fusion source

DEAP Dark matter Experiment using Argon Pulse-shape discrimination

DER Detector Electronic Response

DM Dark Matter

ED Module Electron Drifting Module

EDELWEISS Expérience pour DEtecter Les WIMPs En Site Souterrain

EFT Effective Field Theory

ER Electron Recoil

EXO Enriched Xenon Observatory

FFR Forward Field Region

IC Internal Conversion

ICV Inner Cryostat Vessel

KS test Kolmogorov-Smirnov test

LAr Liquid Argon

LHC Large Hadron Collider

LLAMA LUX Legacy Analysis Monte Carlo Application

LUX Large Underground Xenon

LUXSim LUX Simulation Package

LXe Liquid Xenon

LZap LZ Analysis Package

LZ LUX-ZEPLIN

MC Monte Carlo

MDC Mock Data Challenge

MFP Mean Free Path

ML Machine Learning

MOND MODified Newtonian Dynamics

MS Multiple Scatter

MSSI Multiple Scatter Single Ionisation

nearGX A 1S1 2S2 multiple scatter near the cathode (a.k.a. near miss gamma-X)

NEST Noble Element Simulation Technique

NR Nuclear Recoil

OCV Outer Cryostat Vessel

OD Outer Detector

PandaX Particle and Astrophysical Xenon Detector

PDE Partial Differential Equation

PED Partial Energy Deposition

PMT Photo-Multiplier Tube

PTFE Polytetrafluoroethylene

QCD Quantum Chromo-Dynamics

RFR Reverse Field Region

ROI Region of Interest

RQ Required Quantity

SD Spin Dependent

SI Spin Independent

SM Standard Model

SNOLab Sudbury Neutrino Observatory Lab

SS Single Scatter

SURF Sanford Underground Research Facility

TES Transient Edge Sensor

TPC Time Projection Chamber

VUV Vacuum Ultra-Violet

WIMP Weakly Interacting Massive Particle

XGBoost **eX**treme **G**radient **B**oosting

ZEPLIN ZonEd Proportional scintillation in LIquid Noble gases

Chapter 1

Introduction

One of modern physics most mystifying questions surrounds the nature of dark matter - an elusive substance that appears to make up 84% of the matter in our universe. Having only been observed via gravitational effects to date, experimental attempts to directly detect this dark matter represent a cornerstone approach in our effort to understand this substance. Maturation in the technology underpinning direct detection experiments has seen the sensitivity of these low background detectors being driven to unprecedented levels. Furthermore, with dark matter continuing to evade detection, analyses have begun looking for signals in previously uncharted energy ranges. In doing so the need to be able to identify and mitigate rarer and rarer backgrounds has grown in importance to these experiments.

This thesis is primarily concerned with original analysis pertaining to some methods for identifying and handling backgrounds in the Large Underground Xenon (LUX) experiment, and its successor LUX-ZEPLIN (LZ). Both experiments are examples of dual-phase xenon time projection chambers (TPCs), which represent the most sensitive type of device humanity currently has for the direct detection of dark matter.

The thesis will begin by giving an overview of the current status of dark matter, including the evidence for it, leading theories used to describe it, and search attempts; followed by a description of the LUX and LZ experiments. This is followed by a study of potential backgrounds that might be produced by $^{131\text{m}}\text{Xe}$, a calibration source used by LZ; and how event position reconstruction handles non-uniformities in LZ. Finally, this thesis will end with a description of an analysis into a rare background pathology seen in LUX known as a gamma-X event, which can allow a multiple scatter background to be detected as a positive dark matter signal. This will include a description of how gamma-X events were simulated,

the use of an observed machine learning algorithm to develop a cut on these events, and the performance of this cut on LUX data.

Chapter 2

Dark matter

2.1 Introduction

An astonishing fact uncovered from the advent of modern cosmology is that over 95% of the universe remains a mystery to us. For nearly 40 years dark matter, a large proportion of this dark universe has captured the attention and imagination of physicists. Appearing to have no electromagnetic interactions with light, this mysterious substance has only been observed via its gravitational effects. Despite this, its influence is an essential ingredient in the formation of galaxies and galaxy clusters. The story of dark matter begins in the 1930s. At that time the capability for astronomers to observe its gravitational effects were limited by the instruments of the day. But as time went on, and technology improved, so too did our awareness of dark matter. Nowadays, the unravelling of this dark matter mystery stands as one of the great challenges facing the physics community. This chapter will focus on the current scientific status of dark matter, beginning with its discovery, covering the leading theories devised to explain it, and finishing with the ongoing efforts to detect it.

2.2 Evidence for dark matter

2.2.1 Early dark matter history

For nearly a century, physicists have been pondering the source of an invisible form of gravity which seems to dominate our universe. The leading explanation for this phenomenon is known as dark matter (DM), an imperceivable mass resting in spherical halos around galaxies. The existence of ethereal forms of matter is an old idea, arguably tracing its origin back to the ancient Greeks [3]. But it was not until 1933, when Zwicky invoked DM as an explanation for the anomalous galactic motion of the Coma Cluster, that the modern, scientific notion of

DM was born [4].

Zwicky’s work focused on Hubble and Humason’s 1931 redshift data of various galaxy clusters [5]. In analysing this data, Zwicky noticed an unusually high galactic velocity dispersion in the Coma Cluster. This observation inspired Zwicky to apply the virial theorem in order to estimate the cluster’s mass based on gravitational effects (M_G). When comparing this to the alternative mass estimate based on the cluster’s luminosity (M_L) he found a disagreement of $\frac{M_G}{M_L} \approx 100$. This showed that rotational velocities of galaxies in the Coma Cluster were too high to be explained purely by the luminous matter of its galaxies. One explanation for this anomaly which Zwicky proposed was the presence of a non-luminous form of matter permeating the cluster, which he called “Dunkle Materie”. Notably, Zwicky’s use of the term “dark matter” was in continuity with contemporary astronomers of the 1930s such as Kapteyn, Oort, Jeans, and others who had been studying the dynamics of stars in the local Milky Way. Nevertheless, the concept was never taken very seriously at the time. After this brief period of consideration given to DM in the 1930s, it fell from scientific focus for several decades.

2.2.2 Galactic rotation curves

Then in the 1970s, observations by Rubin and Ford [6] of the rotational velocity of spiral galaxies suddenly brought DM back into scientific focus. Their observations using 21 cm hydrogen lines indicated that the rotational velocity of galaxies remained constant as orbital radius increased. This relationship between orbital radius and velocity is exemplified in Figure 2.1, and stands in contrast to predictions from Newtonian dynamics, which states that orbital velocity beyond the galactic bulge decreases according to¹:

$$v(r) = \sqrt{\frac{GM(r)}{r}} \quad (2.1)$$

where G is Newton’s gravitational constant, r is orbital radius, v is orbital velocity, $M(r) = 4\pi \int r^2 \rho(r) dr$, and $\rho(r)$ is the mass density profile of the galaxy.

Figure 2.1 illustrates the flatness of the observed rotation curves at large radii. This feature can be reconciled with Eq. 2.1 by invoking the existence of a DM halo with $\rho(r) \propto \frac{1}{r^2}$. This increase in mass explains how objects with longer orbits can move at the observed velocities without violating Newton’s laws.

¹This spherically symmetric formula is only an approximation for spiral galaxies. Nevertheless, the argument holds.

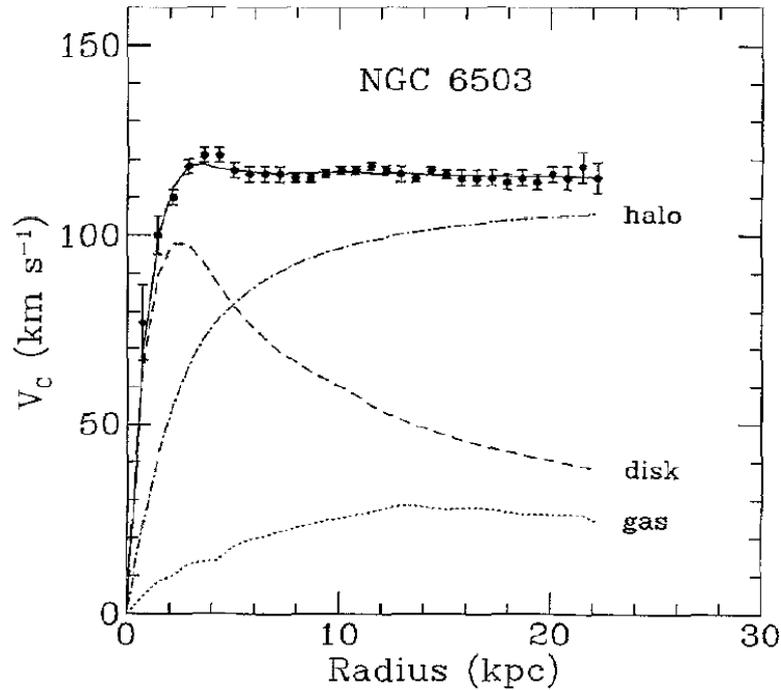


Figure 2.1 Galactic rotation curve for NGC 6503. Dark matter halo (dash-dotted line), optical disk (dashed line), and gas (dotted line) contributions used to match the data are indicated [7].

However, DM is not the only explanation for these anomalous rotation curves. An alternative proposal by the name of MOND (Modified Newtonian Dynamics) attempts to explain them without invoking the existence of any “missing” matter. Instead, this class of theory modifies Newtonian physics at low accelerations. Originally proposed by Milgrom in 1983 [8], MOND is capable of explaining some of these galactic rotation curves. However, MOND has not yet been able to explain temperature profiles [9] and stability [10] of the galactic disc without still invoking some form of DM. MOND additionally suffers from an inability to explain other observations, described in the next few sections, which contribute to the body of DM evidence. Although MOND theories and experiments remain under development [11], scientific consensus remains strongly in favour of DM as the preferred explanation for these phenomena.

2.2.3 Gravitational lensing

Gravitational lensing is a phenomenon that occurs when a large distribution of matter (such as a galaxy cluster) is positioned between a distant light source and an observer. The effect results in light from the distant source being gravitationally distorted, bending the trajectory of

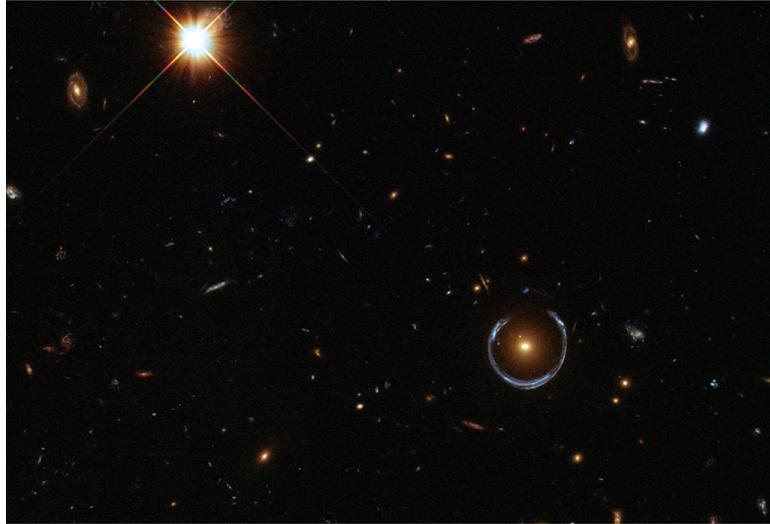


Figure 2.2 A gravitational lensing of a distant blue galaxy surrounding the red galaxy LRG 3-757. Light from the distant galaxy is distorted as it passes through the strong gravitational field, pulling the image of the distant blue galaxy into a ring-like shape [12].

the light. Under the right circumstances, the result on the source's image from the perspective of an observer is similar to that of an optical lens. When a distant source is gravitationally lensed by a galaxy or galaxy cluster for observers on Earth, it is common for astronomers to utilise this effect to peer deeper into the cosmos than would otherwise be allowed. As seen in Figure 2.2, the resulting image is a distorted version of the target galaxy, where the amount of distortion is directly related to the mass of the lensing structure. This effect allows for the measurement of the mass profile of the lensing object [13]. Over the past 20 years, several galaxy clusters have been observed to possess a mass density profile which varies from the baryonic mass distribution (taken from X-ray data). This provides convincing evidence for a significant DM component to these galaxy clusters.

A classic example is the Bullet Cluster which is in fact two galaxy clusters that have collided and subsequently passed through one another. To highlight the difference between the gravitational centre of mass and the baryonic mass density, the image in Figure 2.3 shows the Bullet Cluster which has been artificially coloured, with X-ray emission data in pink and gravitational lensing data in blue. The gravitational lensing measurements show two clear "cores" of mass, thought to be DM. These DM cores have passed through one another with minimal interaction. The baryonic matter on the other hand is observed to be clustering together in the centre. This is because the pink region predominantly consists of gas, which experiences scattering as the clusters collide. The result of these interactions is

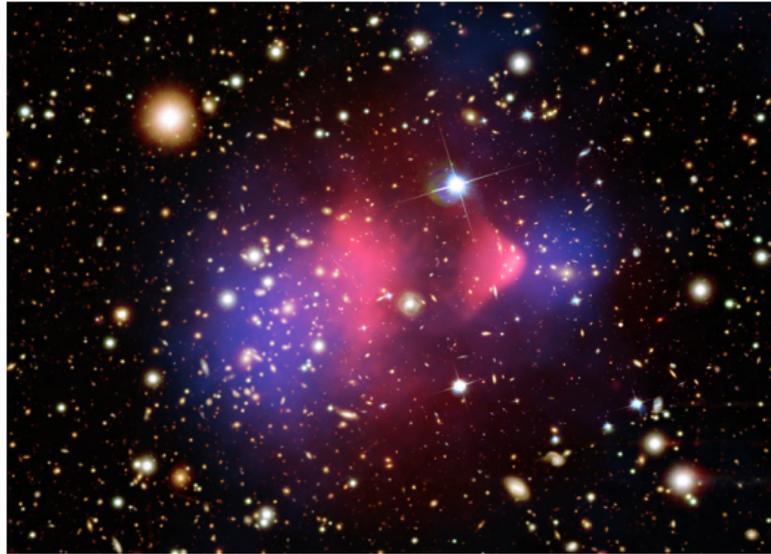


Figure 2.3 Dark matter in the Bullet Cluster. This figure has been artificially coloured to highlight the separation of the interstellar gas (in pink) and the gravitational centre of mass (blue) [15, 16].

a “drag” experienced by the baryonic matter, which separates it from the DM². The Bullet Cluster itself is specifically significant to DM searches because of the difficulty MOND has explaining it [14].

2.2.4 Cosmic microwave background radiation

The cosmic microwave background (CMB) radiation is the heat-glow remnant released when the hot-dense-early universe first became transparent to photons. The photons in the CMB last scattered when the universe was 380,000 years old, a moment known as recombination. This was when atoms first formed due to the cooling of the primordial plasma, which released photons from their coupled state within this plasma. The CMB has an average temperature of 2.7 K. The best map we have of the CMB currently comes from the Planck satellite, which collected data between 2009 and 2014 [17]. The CMB map illustrated in Figure 2.4 shows the tiny (1 part in 10,000) deviations from the average temperature. These variations (a.k.a. anisotropies) result from matter density fluctuations present immediately after the Big Bang. These density variations evolved into colossal sound waves, known as the baryon acoustic oscillations (BAO). BAOs reverberated through the universe during the first few hundred thousand years of its life, before being frozen in place by recombination.

²Notably it is only the gas of the interstellar medium which makes up the bulk of the baryonic matter which experiences this drag. The planets and stars which have not collided with one another remain in the blue DM region.

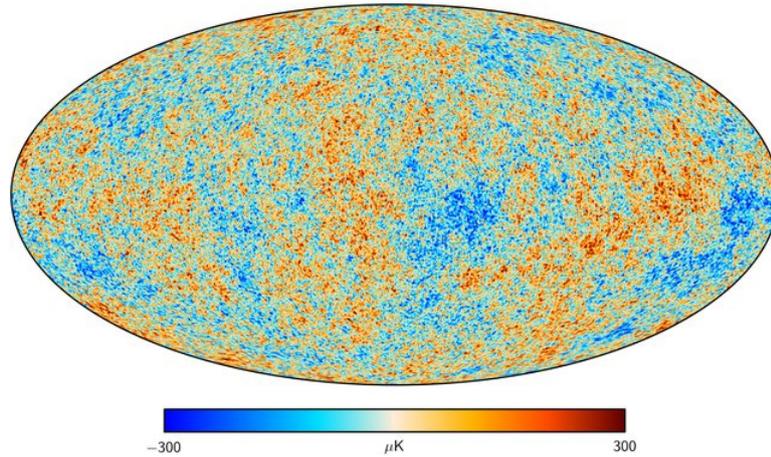


Figure 2.4 The all-sky map of the CMB radiation temperature variations as detected by the Planck satellite. The differences in temperature are highlighted on this map, with the coldest regions shown in blue, and the hottest regions in red. The average temperature is 2.7 K, and the range of temperature variation is 0.0006 K [17].

When rearranging the CMB temperature anisotropies by their angular size, one gets what is known as the CMB power spectrum. This spectrum reveals critical information about various cosmological parameters. These parameters include the flatness of the universe and the relative abundances of baryonic matter, DM, and dark energy. The CMB power spectrum (shown in Figure 2.5) consists of several obvious peaks. Each peak represents anisotropies caused by a BAO captured in a position of maximum compression or rarefaction at the point of recombination.

The position of the first peak indicates the size of anisotropies which resulted from BAOs that did not have time for a rarefaction before recombination. The size of these fluctuations matches predictions in a flat universe, indicating that our universe has a flat geometry. The first peak can also be used to determine the combined total matter and energy density of the universe. The location and height of the second peak can be used to determine the abundance of baryonic matter in the universe. The remaining peaks give information about the universe's DM density [18].

To quantify this density it is standard in cosmology to use the dimensionless ratios known as density parameters Ω_x , where x can be any one of several species³:

$$\Omega_x = \frac{8\pi G\rho_x}{3H^2} = \frac{\rho_x}{\rho_c} \quad (2.2)$$

³i.e. dark matter, electrons, photons e.t.c.

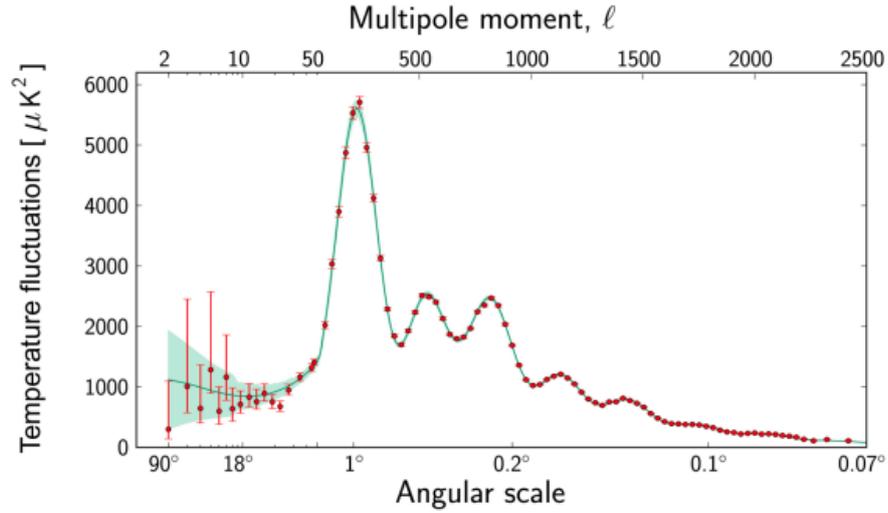


Figure 2.5 The CMB power spectrum shows the relationship between the angular size of CMB anisotropies (lower horizontal axis) and their temperature fluctuation. Multipole moment ℓ is shown in the upper horizontal axis. Note that a scale change on the horizontal axis from logarithmic to linear exists at $\ell=500$. Data (red dots) is measured by the Planck collaboration. The green line shows the prediction from the Λ CDM model [17].

Where G is the Newton's gravitational constant ($6.67 \times 10^{-11} \text{ m}^3\text{kg}^{-1}\text{s}^{-2}$), H is the Hubble parameter, ρ_x is the species density, and ρ_c is the critical density for which the universe is flat. The latest measurements of these parameters indicate the baryon density, $\Omega_b = 0.0484 \pm 0.00073$, and the DM density, $\Omega_{DM} = 0.2605 \pm 0.00624$ [17].

2.2.5 Structure formation

One of the best descriptions of the CMB power spectrum comes from the Lambda Cold Dark Matter Model of Cosmology (Λ CDM). It has been so effective that it has emerged as the standard of modern cosmology. As a consequence of the role of DM in Big Bang Nucleosynthesis, Λ CDM requires that the DM particle not be relativistic or baryonic (i.e. hence the name cold dark matter) [17].

The Λ CDM model is supported by observations of large-scale structures. Galaxies and galaxy clusters have formed from small fluctuations in the primordial density. Alternative models that invoke hot DM cannot explain this structure formation. This is because hot DM particles would escape the gravitational potential of the overdensity sites. This would dampen the BAOs, and preventing the universe from attaining the mass scale of observed large super-clusters [19]. A similar argument exists for why DM particles cannot be

baryonic, since electromagnetic interactions would prevent them from clustering in gravitational wells prior to decoupling.

2.3 Dark matter landscape

2.3.1 Dark matter properties

Several DM candidates have been proposed to either explain the entire DM abundance or help explain some of the gravitational phenomena which DM is otherwise thought to account for. Most well-motivated candidates strive to solve not just the DM mystery, but also other unsolved problems in modern physics. Although there are far too many candidates to cover in detail, Section 2.3 will discuss several of the DM hypotheses, and their properties.

Beginning with astronomical observations, several constraints can be placed upon DM candidates:

1. It must not absorb or emit photons, although annihilation into photons is permissible.
2. Its major component must be non-baryonic ($\Omega_b \ll \Omega_m$).
3. It must be cold, i.e. moving at non-relativistic speeds at the time of freeze-out⁴.
4. It must not have an electrical charge.
5. It must be stable, or else have a lifetime greater than the age of the universe.

Although some Standard Model (SM) particles may contribute to satisfying a few of these conditions, a major fraction of the DM must be made of new, undiscovered particles.

2.3.2 Standard model dark matter

There are strong restrictions on the amount of baryonic matter in the universe from primordial Big Bang Nucleosynthesis and observations of the CMB. However, several candidates for DM using particles from the SM can explain some of this abundance and are the subject of ongoing investigation. Although these solutions suffer from not meeting all the requirements for DM, they benefit from the fact that their existence has been confirmed.

One such set of candidates are the MACHOs (massive compact halo object). MACHOs are

⁴Freeze-out refers to the epoch when DM density fell to a point where DM-DM annihilation ceased.

non-luminous forms of baryonic matter floating in galactic halos. Examples of MACHOs includes brown dwarfs, neutron stars, and black holes. While this model helps describe the low mass-to-luminosity ratios observed in gravitational lensing experiments, the frequency of MACHOs is too small to be a major factor in the overall DM abundance. With an upper contribution limit of 8% (95% CL) to the mass of the galactic halos [20], their existence does not explain the flat rotation curves discussed in Section 2.2.4. Furthermore, the presence of anisotropies in the CMB as discussed in Section 2.2.4 are also not explained by MACHOs. Thus, although the existence of MACHOs is undeniable, they do not remove the need for non-baryonic forms of DM.

Another proposed solution from the early days of DM searches is the neutrino. This solution was appealing due to the neutrino being: long-lived, chargeless, weakly interacting, and having a confirmed existence. However, the SM neutrino suffers from two major drawbacks. First, its mass density on the cosmic scale is too low ($\Omega_\nu \leq 0.003$) to describe the invisible source of gravity dominating the universe. Second, its relativistic nature does not account for structure formation [21]. This means the problem of DM requires invoking particles from beyond the standard model (BSM).

More recently a new neutrino species known as the sterile neutrino has been proposed. The sterile neutrino (a.k.a. right-handed neutrino) is not considered a SM particle and exists as a proposed explanation for the asymmetry between traditional neutrinos (which are observed to be exclusively left-handed) and other leptons (which are capable of left and right-handedness). Due to this, sterile neutrinos are forbidden from coupling with W bosons (which is facilitated in traditional neutrinos by their left-handedness [22]). Sterile neutrinos may however have an oscillation probability to traditional neutrino flavours via a small mixing angle. This oscillation probability may be able to explain a 2018 result from MiniBooNE which observed a neutrino oscillation stronger than what is predicted by a neutrino model with three species [23]. Depending on the sterile neutrino mass, this solution may be capable of explaining all the DM in the universe [24].

2.3.3 WIMPs

The consideration towards neutrinos as a DM candidate helped to establish a new class of hypothetical particles, WIMPs (weakly interacting massive particles). A WIMP is any stable particle which experiences a gravitational interaction, and an interaction via another force at the weak scale. Travelling at non-relativistic speeds at the moment of “freeze-out”, WIMPs are a favoured CDM candidate.

WIMPs in the early universe would have behaved very differently to WIMPs today. In that hot, dense environment, all particles were in thermal equilibrium. The thermal equilibrium of WIMPs with ordinary matter would have been governed by the WIMP number density (n), the WIMP annihilation cross-section (σ_{ann}), and their relative velocity (v). Under the conditions of the early universe, Majorana⁵ variety WIMPs would have regularly annihilated with other WIMPs to produce SM particles. But as the universe expanded, and the WIMP density dropped, so too did this annihilation rate. Once the Universal expansion rate⁶ ($H(t)$) surpassed the WIMP annihilation rate⁷ ($\langle \sigma_{ann}v \rangle$), WIMPs “froze-out”. Meaning that WIMPs stopped annihilating, and their abundance became fixed.

Notably, a lower self-annihilation rate more quickly removes WIMPs from thermal equilibrium, increasing the predicted DM relic density. In order for the relic density to match the measured value of $\Omega_{DM} \approx 0.26$ requires:

$$\langle \sigma_{ann}v \rangle = 3 \times 10^{-26} \text{cm}^3 \text{s}^{-1} \quad (2.3)$$

this is roughly what is expected for a ~ 100 GeV particle which experiences a weak force interaction.

In a strange coincidence, supersymmetric⁸ extensions of the standard model of particle physics predict a new particle with these properties. This “WIMP miracle”, as it has come to be known, unites the disparate fields of cosmology and particle physics in an entirely surprising way. Given the promise of supersymmetry to unify the fundamental forces, and to solve the hierarchy problem⁹ in a grand unified theory, the WIMP has maintained the status of prime DM candidate for some time [25].

Apart from supersymmetry, valid WIMP DM candidates also arise from other theories. For instance, Universal Extra Dimensions (UED) [26], and non-thermal DM of huge masses, the so-called “WIMPzillas” [27]. These contributions extend the parameter range for WIMP

⁵i.e. Majorana refers to a class of particle which is its own anti-particle

⁶ $H(t) = \frac{a'(t)}{a(t)}$, where a is the scale factor which parameterises the relative expansion of the universe.

⁷Specifically, $\langle \sigma_{ann}v \rangle$ is the thermally averaged WIMP annihilation cross-section times the relative velocity v .

⁸Supersymmetry is a conjectured relationship in which every SM particle has an associated partner particle which has a spin that differs by a half-integer and has not been discovered.

⁹The hierarchy problem is a term given to the recognition that there is no natural explanation as to why the weak force is 10^{24} times stronger than gravity.

searches given to indirect detection, direct detection, and collider experiments to probe.

The strong motivation for WIMPs has driven the development of many WIMP searches in the GeV–TeV range over the past few decades. The outcome of these experiments has been the exclusion of a large region of the WIMP parameter space. Many supersymmetry models have also been further constrained by the lack of evidence in collider experiments. The focus for future direct detection WIMP searches is on improving detector sensitivity and reducing detector thresholds to search for sub-GeV WIMPs. These experiments are described in Section 2.4.

2.3.4 Axions

The axion is another BSM DM candidate particle which is yet to be experimentally verified. What is particularly compelling about the axion is how it naturally arises from a proposed solution to an entirely different modern physics mystery, the strong Charge Parity (CP) problem [28].

As is the case with most problems in modern physics, the strong CP problem describes a state of disagreement between theoretical predictions, and experimental results. In this case, the Quantum Chromo-Dynamics (QCD) equations of motion predict a CP symmetry violation, which is not observed. Specifically, QCD predicts that neutrons should experience a non-negligible electric dipole moment, which either does not exist or is a factor of $\sim 10^{12}$ too small [29]. The strong CP problem disappears when Θ , the “constant” in the QCD Lagrangian governing the strength of QCD asymmetry, is zero. However, there is no theoretical rationale for this case.

In 1977, Peccei and Quinn proposed an intriguing answer to the strong CP problem. Their solution relied upon describing the value of Θ as a function of space and time¹⁰ [25]. This field parameterisation naturally causes Θ to fall to zero as the overall vacuum energy approaches its lowest possible energy configuration. But as is the case with any quantum field, an oscillation in this Θ field around its minimum value yields a particle. For the Θ field, this new particle was the axion.

A common concern for the axion as a DM candidate is its low predicted mass of between 10^{-6} and 10^{-3} eV/ c^2 [30], which is problematic depending on the axion number density.

¹⁰Such an approach is known as a field parameterisation.

However, this concern vanishes if axions were produced in prodigious enough numbers in the early universe. Thermal production of axions shortly after the Big Bang would have led to axions with relativistic velocities [31]. But the “re-alignment mechanism” describes how axions can be cooled to temperatures capable of facilitating the mass scale of observed large super-clusters. This mechanism reduces the kinetic energy of axions in the early universe [14].

Axion detection may be possible via its photon coupling [32], which can lead to the conversion of an axion to a photon in a magnetic field or vice versa. This process is known as the Primakoff effect and creates an opportunity for axion detection.

In addition to axions, the energy range of axion searches has been expanded to search for “axion-like particles” (ALP). These particles share some of the axion phenomenology, such as their coupling to photons and predicted low mass. However, they do not have the same level of theoretical motivation. The ALP mass and its coupling to photons are both free parameters in these models. As such, ALP searches occur in an enlarged parameter space compared to the original QCD axions.

2.4 Direct detection of dark matter

2.4.1 Concept

In addition to the DM candidates listed in the previous section, there exist many more with varying degrees of theoretical motivation. This phenomenological richness provides a variety of experimental approaches in the search for DM. One particular branch of DM search is known as direct detection. Direct detection aims to measure DM-nucleus interactions in Earth-based, low background detectors [14]. Over the past 30 years, this approach has gained popularity due to its potential to satisfy the “WIMP miracle”, putting the detection of a GeV range WIMP in particular focus.

The presence of Earth in the Milky Way’s WIMP halo is expected to result in a DM particle flux through our planet [6]. Despite the extremely low predicted probability of a WIMP-nucleon interaction, such an event should be observable in Earth-based detectors, if the WIMP flux and cross-section are large enough [33]. Direct detection experiments look for a signal above the expected background across a range of recoil energies. However, the signals produced in these experiments are quite small (1-100 keV). For direct detection experiments to be effective there are three essential requirements:

- a low recoil energy (E_R) detection threshold
- low noise
- stable operation with excellent reproducibility of calibration signals

How this is achieved will often depend on the individual detector's technology or implementation. However, what all these experiments share is a need to maximise their potential interaction rate.

2.4.2 Interaction rate

The differential energy spectrum of a WIMP induced nuclear recoil ($\frac{dR}{dE_R}$) is built out of three main considerations. First are astrophysical such as the local WIMP density (ρ_{DM}), galactic escape velocity (v_{esc}), minimal detectable WIMP velocity (v_{min}), and velocity distribution ($f(\mathbf{v})$), which are discussed in Section 2.4.3. The second set of considerations take into account particle physics, which are governed by the event cross-section (σ_A) and discussed in Section 2.4.4. The final set of considerations are nuclear form factor corrections ($F^2(E_R)$) and will be discussed in Section 2.4.5. When combined, the differential energy spectrum is expressed as [34]:

$$\frac{dR}{dE_R} = F^2(E_R) \frac{\rho_{DM} \sigma_A}{2M_{DM} \mu_A^2} \int_{v_{min}}^{v_{esc}} \frac{f(\mathbf{v})}{v} d^3v \quad (2.4)$$

where E_R is the nuclear recoil energy, μ_A is the reduced mass of the WIMP-nucleus system ($\mu_A = M_{DM} M_A / (M_{DM} + M_A)$). The symbol M represents mass; with the subscripts DM and A indicating whether this is the mass of dark matter or target atom respectively.

2.4.3 WIMP halo

WIMPs are assumed to reside in isothermal, spherically symmetric galactic halos, with an average velocity vector of zero. Two features of our local WIMP halo are of particular interest to experimentalists: the density (ρ_{DM}), and its velocity distribution. To streamline the comparison of experimental results, the DM community makes canonical assumptions about the value of these features.

The local DM density, which itself is a subject of ongoing inquiry, has had fluctuating estimates of its value over the last 30 years [35]. These estimates arise out of two competing models, one which is flat at large galactic radii, and one which has a cusp. Recent data from the GAIA (Global Astrometric Interferometer for Astrophysics) space mission [36], which

inform these models, indicates a local DM density of between 0.2 and $0.6 \text{ GeV}\cdot\text{cm}^{-3}$, with a value of $0.5 \text{ GeV}\cdot\text{cm}^{-3}$ to be most realistic [37]. However, given the variability in this field a canonical value of $0.3 \text{ GeV}\cdot\text{cm}^{-3}$ has been chosen and is likely to remain in use by experimentalists for the foreseeable future.

The velocity of the DM halo follows a Maxwellian distribution ($f(\mathbf{v}) \sim \exp(-v^2/v_0^2)$), where v is the WIMP speed, and $v_0 = 220 \text{ km}\cdot\text{s}^{-1}$ is the average halo WIMP speed [38]. This formulation takes into account the fact that, although the average WIMP velocity vector is zero, the velocity squared (hence the dispersion) is not. The integral of this distribution is a necessary component of calculating the expected detection rate. It is commonly expressed in the form [34]:

$$\int_{v_{min}}^{v_{esc}} \frac{f(\mathbf{v})}{v} d^3v = \begin{cases} \frac{1}{v_0 y} & \text{if } z < y \text{ and } x < |y - z| \\ \frac{1}{2N_{esc} v_0 y} [\text{erf}(x + y) - \text{erf}(x - y) - \frac{4y}{\sqrt{\pi}} \cdot \exp(-z^2)] & \text{if } z > y \text{ and } x < |y - z| \\ \frac{1}{2N_{esc} v_0 y} [\text{erf}(z) - \text{erf}(x - y) - \frac{2}{\sqrt{\pi}}(y + z - x) \cdot \exp(-z^2)] & \text{if } |y - z| < x < y + z \end{cases} \quad (2.5)$$

where $x = \frac{v_{min}}{v_0}$, $y = \frac{v_{Earth}}{v_0}$, $z = \frac{v_{esc}}{v_0}$, $N_{esc} = \text{erf}(z) - \frac{2z \cdot \exp(-z^2)}{\sqrt{\pi}}$, $v_{esc} = 544 \text{ km}\cdot\text{s}^{-1}$ (Milky Way escape velocity), $v_{Earth} = 245 \text{ km}\cdot\text{s}^{-1}$ (velocity of the Earth with respect to the DM halo), and v_{min} is the lowest WIMP speed capable of causing a threshold recoil energy for a given detector.

2.4.4 Nuclear scattering cross-section

Before calculating the cross-section, a choice of scattering model must be made. To facilitate the comparison of different direct detection experiments, two types of WIMP-nuclear interactions have traditionally been focused on, spin-independent (SI) and spin-dependent (SD). In a SI interaction the WIMP couples to the mass of the nucleus, whereas in an SD interaction the WIMP couples to the spin of the nucleus [14]. Alternative interaction models exist other than SI and SD; for instance, the Effective Field Theory (EFT) model which utilises 28 operators (of which SI and SD are special cases) will be discussed in Section 2.6. However, to introduce readers to the topic, this description will focus on the canonical SI cross-section, and briefly discuss the difference to the SD case.

For SI scatters, the WIMP-nucleus cross-section (σ_A^{SI}) is expressed in terms of the WIMP-

nucleon cross-section (σ_n^{SI}). The relation between these two cross-sections is given by [39]:

$$\sigma_A^{SI} = A^2 \frac{\mu_A^2}{\mu_n^2} \sigma_n^{SI} \quad (2.6)$$

where μ_n is the WIMP-nucleon reduced mass¹¹. An A^2 enhancement to σ_A^{SI} can be seen as a result of Eq. 2.6. This feature leads to an increase in the sensitivity of SI scattering for heavy target nuclei and is a key motivation for their use.

SD searches on the other hand, do not benefit from this A^2 enhancement. In this case the WIMP-nucleus cross-section (σ_A^{SD}) depends on a spin constant (λ_S) unique to the target isotope (which can be looked up in [40]), and the total angular momentum of the target nuclei (J) [14]:

$$\sigma_A^{SD} = \frac{3}{4} \lambda_S^2 J(J+1) \frac{\mu_A^2}{\mu_n^2} \sigma_n^{SD} \quad (2.7)$$

Notably, contributions to nuclear spin are cancelled out between pairs of nucleons. So only isotopes with an odd number of nucleons are sensitive to SD WIMP interactions.

In Figure 2.6, Eq. 2.4 is used to calculate the differential SI event rates between WIMPs and a target nuclei. Figure 2.6a shows this for WIMPs of mass 20 GeV (blue), 35 GeV (red), and 50 GeV (black) in ^{131}Xe . Figure 2.6b shows this for a 100 GeV WIMP in ^{131}Xe (purple), ^{40}Ar (blue), and ^{72}Ge (black). All rates assume a WIMP-nucleon cross-section based the most recent WIMP-nucleon cross-section limits set by XENON1T [41].

2.4.5 Nuclear form factor

Momentum (q) transfer in the WIMP-nucleus interactions is described by $F^2(E_R)$. The variation of SI form factors for three common WIMP targets is shown in Figure 2.7. An interesting feature is apparent when paying attention to the rate in $F^2(E_R)$ reduction for higher recoil energies, particularly for targets with larger values of A . Specifically, a loss of scattering coherence for higher energy recoils is observed, which is a key feature that the SI form factor must describe.

This loss of coherence increases as the de Broglie wavelength ($\lambda_B = h/q$) of the WIMP approaches the size of the target nucleus. This results in a sharp drop in the interaction probability for a specific recoil energy based on the λ_B of the target isotope. The dominance

¹¹note that this formalism assumes WIMP interactions between protons and neutrons is equivalent.

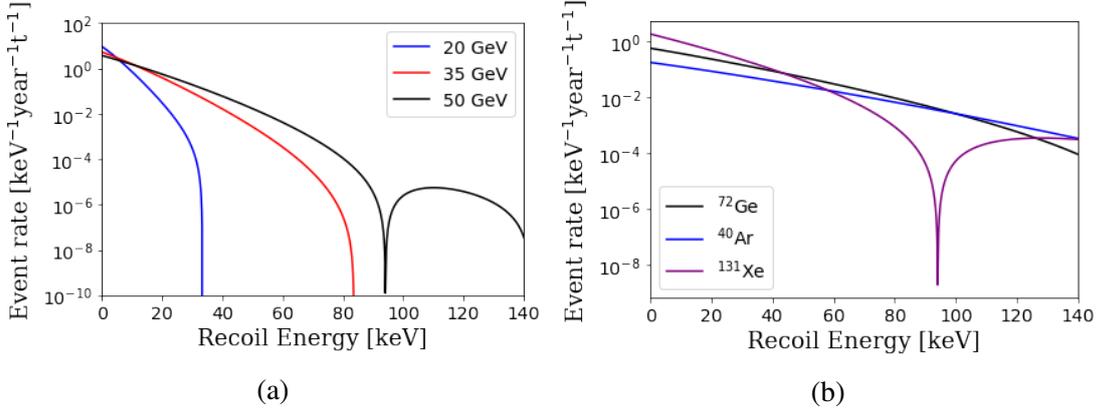


Figure 2.6 Predicted WIMP event rate based on XENON1T nucleon cross-section limits [41], for various target materials and WIMP masses. (a) shows event rate in ^{131}Xe for three WIMP masses: 20 GeV, 35 GeV, and 50 GeV. (b) shows event rate for a 100 GeV WIMP in three target materials: ^{40}Ar , ^{72}Ge , and ^{131}Xe .

of this effect over other factors in Eq. 2.4 at high recoil energies stands out in Figure 2.6b, where at ~ 90 keV a resonance appears for ^{131}Xe . The precise position of this resonance will depend highly on the target isotope, even for isotopes of the same element. Figure 2.6b also shows that the impact of the form factor only dominates Eq. 2.4 at high values of E_R . At low values of E_R the situation is reversed, with A^2 enhancement giving the heaviest target (^{131}Xe), the highest event rate.

The SD form factor has largely been disregarded from this discussion. However, it is worthwhile to note that the SD form factor must also describe the spin distribution within the nucleus.

2.4.6 Detection signals and technology

Aside from A^2 enhancement and loss of coherence, there are several other factors to consider when selecting a target material. For instance, the scalability of liquid noble based experiments is much better than for experiments using a crystal target. Alternatively, if one is interested in sub-GeV WIMP searches, one must consider the various energy detection thresholds that exist in different materials. The choice of target will also affect the type of signal which is produced by a WIMP-nucleus interaction. Particle interactions with nuclei or electrons have three methods for transferring energy: ionisation, excitation, and heat [42]. Utilising one or a combination of these signals, an event rate of DM interactions as a function of recoil energy can be determined. Some signal combinations which are typically employed

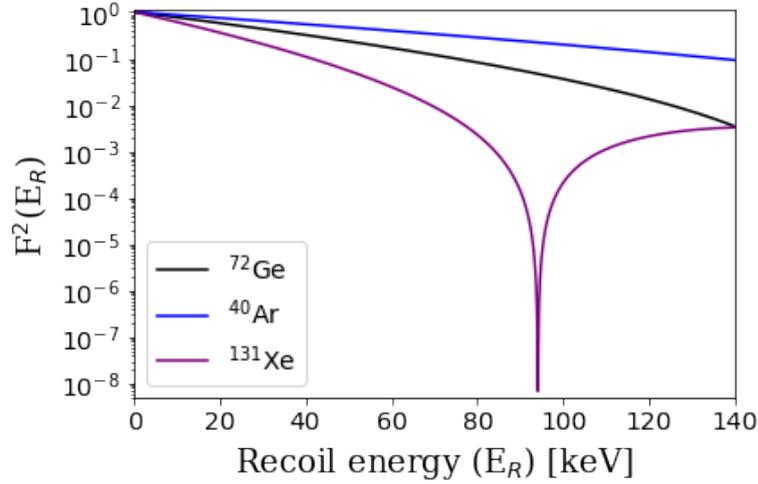


Figure 2.7 Nuclear form factor for ⁴⁰Ar, ⁷²Ge, and ¹³¹Xe

include ionisation and excitation (Xe, Ar), ionisation and heat (Si, Ge), and excitation and heat (NaI, CaWO₄).

An example of a type of detector which uses ionisation, and phonon signals are cryogenic solid-state detectors such as EDELWEISS and CDMS. These experiments led the way in the early days of direct detection. However, because they employ crystals as their target material, scaling up their exposure by adding target mass has proven difficult. A modern incarnation of this detection method is SuperCDMS. SuperCDMS is based in the SNOLab (Canada) and employs a germanium target sensitive to ionisation and phonon signals. An electric field applied across the detector modules extracts ionised electrons to the detector surface, with phonon and ionisation sensors located on the faces of the detectors. Using this technology in 2018, SuperCDMS reported a world-leading WIMP-nucleon cross-section limit of $1.4 \times 10^{-44} \text{ cm}^2$ for low mass WIMPs (1.6 to 5.5 GeV) [43].

Other experiments such as CRESST are able to read out light and phonon signals. Phonons are detected by utilising CaWO₄ crystals as detector modules, and transition-edge sensors (TES). Using silicon calorimeters, scintillation light is detected, which is converted into phonons, and are in turn is also read out by TES. With the rise of the CRESST-III experiment, this collaboration now focuses entirely on sub-GeV WIMPs. In 2019, CRESST-III demonstrated an energy threshold of 30.1 eV for one of their detector modules, which allows them to place world-leading sub GeV WIMP-nucleon cross-section limits down to 0.16 GeV [44].

Above a 10 GeV WIMP mass, noble liquid detectors are the current sensitivity leaders,

using ionisation and scintillation channels to define their signals. Although by comparison noble elements have a much lower density than solid state detectors, their domination in sensitivity at this energy range is due to the relative ease by which this technology can be scaled. The most popular utilization of noble elements for WIMP searches is the xenon dual-phase time projection chamber, an approach which will be described in Chapter 3. This technique was pioneered by ZEPLIN-II and XENON10 [45, 46] in 2006, and was also employed by LUX in 2013 when it was the first to achieve sub-zeptobarn sensitivity for the SI WIMP-nucleon cross-section [47]. This technique is currently employed by several collaborations (LZ [48], XENON [41], PandaX [49]), as they race to increase their target mass, and reach new levels of sensitivity. In the future this approach will be employed by DARWIN in the use of a 40 tonne target mass. The unprecedented sensitivity of DARWIN will for the first time lead to an irreducible background from neutrinos (a.k.a. the neutrino floor) [50]. At this point if WIMPs have not been found, new measures will have to be taken to address neutrino backgrounds¹². Dual-phase Xe TPCs are the focus of this thesis, and will be elaborated on in Chapter 3.

Single phase detectors are also used in DM searches, an example of this kind of detector XMASS (Xenon detector for Weakly Interacting MASSive Particles) located at the Kamioka Observatory in Japan [51]. Notably, XMASS is only sensitive to scintillation light, which makes event position reconstruction much more difficult. XMASS consists of a fiducialised 97 kg of Liquid Xeonon (LXe) in a spherical vessel. The containing vessel is instrumented with PMTs which cover more than 62% of the inner surface. In 2018, after 705.9 live days, XMASS published a SI WIMP-nucleon cross-section sensitivity of $2.2 \times 10^{-44} \text{ cm}^2$ for a WIMP mass of 60 GeV. This result however is far outpaced by the dual-phase contemporary experiments of XENONnT and LZ. Subsequently, such single-phase detectors have largely been replaced by a more sensitive devices.

A common alternative to xenon is argon. Although argon is plagued by radioisotopes such as ^{39}Ar , it has several appealing qualities. First, it is a much cheaper material than xenon¹³, which has traditionally made it more appealing for larger scale experiments. Second, Liquid Argon (LAr) has the capability to use pulse shapes to achieve a better rejection of backgrounds from Electron Recoils (ER) compared to LXe based detectors. This is due to the longer time delay between singlet and tripled states of ER and Nuclear Recoil (NR) inter-

¹²Directional detectors may have the capacity to differentiate between a WIMP and a neutrino. However, the volume of such a detector would be far larger than anything currently conceived.

¹³If we ignore the additional cost of purification and underground extraction of argon over xenon.

actions in LAr¹⁴. This is due to the longer time delay between the singlet and triplet states of ER and NR interactions in LAr (as compared with LXe). Third, in LAr, ³⁹Ar can be mitigated by sourcing argon from underground sites, where cosmogenic activation has not produced it, or just by identifying these radioactive events during data analysis [52]. Notably, ³⁹Ar is still present in xenon detectors, however due to its different chemical composition a high level distillation is possible. An example of argon being used in this way is the SNOLab based DEAP-3600 experiment located in Canada. This is a single-phase argon detector consisting of 3.6 tonnes (1 tonne of fiducialised) LAr, and is contained in a spherical vessel similar to XMASS. In their most recent result, DEAP-3600 published a limit of $3.9 \times 10^{-45} \text{ cm}^2$ for a WIMP mass of 100 GeV using an exposure time of 231 days. Much like XMASS, their result is far from the sensitivity limit, which DEAP-3600 will not be capable of reaching [53]. For this reason the DEAP collaboration has joined the Global Argon Dark Matter Collaboration along with other liquid argon (LAr) experiments such as DarkSide, CLEAN, and ArDM. Their goal is to develop a LAr detector with over 20 tonnes of active target. This project, named DarkSide-20k, is planned for operation in Italy's Laboratori Nazionali del Gran Sasso. In parallel, research and development efforts are going towards a next generation detector (ARGO) with a multi-hundred tonne LAr target mass designed to reach the neutrino floor. ARGO is planned to operate at SNOLab, with no timeline in place as of this moment.

To date, no evidence for DM from an individual direct detection experiment has ever been seen by another experiment. In this climate, data is treated statistically to give an upper limit to the potential WIMP-nucleon cross-section. For instance, the 90% upper limit given by the green shaded region in Figure 2.8 shows the current limit for this kind of search as a function of DM mass. The limit can be seen to peak at masses which are similar to heavy nuclei (between ~ 10 -100 GeV). Limits are suppressed at large DM masses, due to an increase in WIMP mass decreasing the WIMP number density (for a fixed WIMP mass density). Sensitivity is also suppressed at low DM masses, due to the effect of detector thresholds on low E_R signal efficiency. However, for all WIMP masses, sensitivity is suppressed by backgrounds. Subsequently, the limitation of backgrounds and signal/background discrimination play an essential role in increasing DM sensitivity.

2.4.7 Status of field

At present, the best SI WIMP sensitivity is held by different detectors depending on the energy range (see Figure 2.8). At sub-GeV DM masses, the leaders are cryogenic solid-state

¹⁴see Section 3.1 for a more in-depth discussion of the ER and NR signal generation.

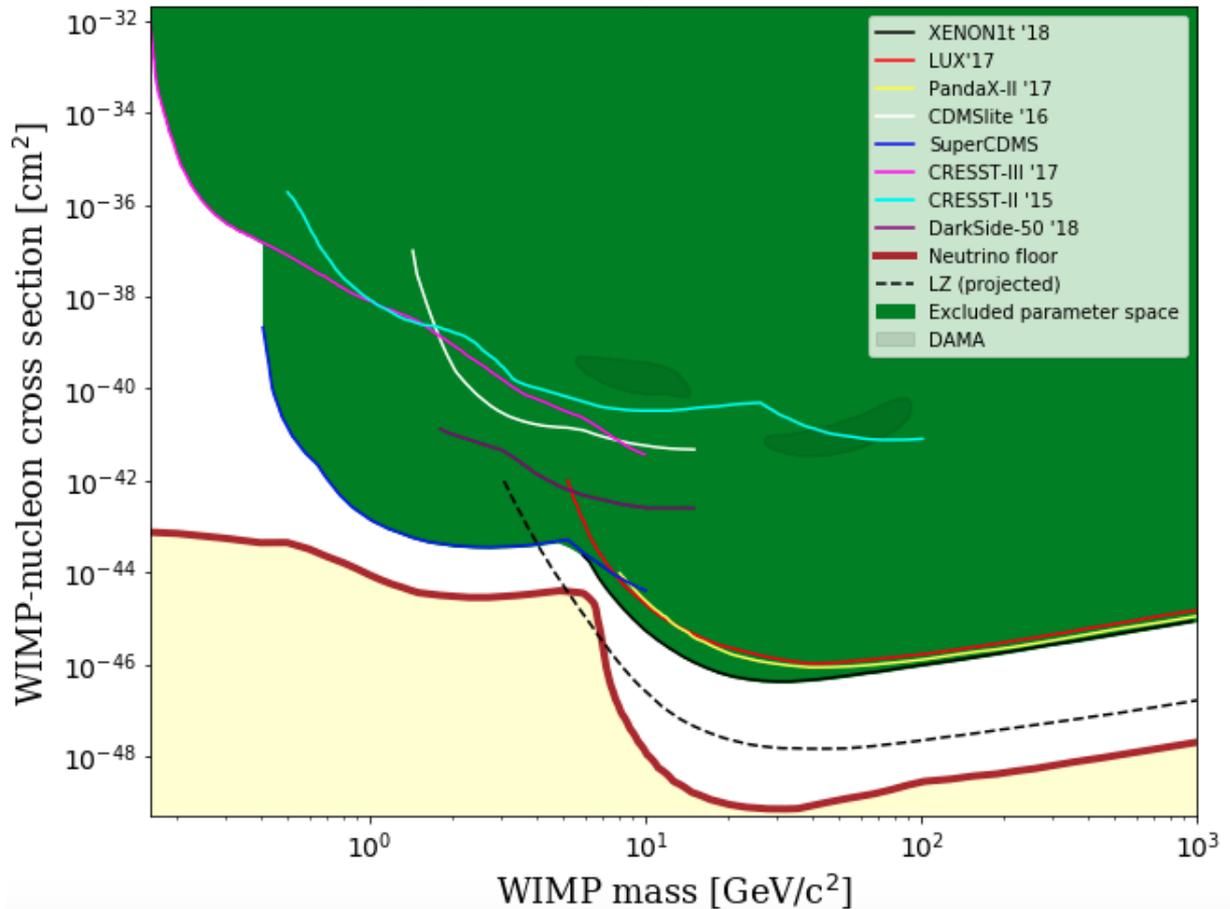


Figure 2.8 Limits on spin-independent WIMP-nucleon cross-section as a function of WIMP mass. This figure displays the 90% CL limit for WIMP-nucleon cross-sections from a variety of experiments up to 1000 GeV mass WIMPs. Experiments included in this figure are XENON1T (black) [41], LUX (red) [47], PandaX-II (yellow) [49], CDMSlite (white) [54], SuperCDMS [43] (blue), CRESST-III (magenta), CRESST-II (cyan) [44], and DarkSide-50 (purple) [52]. The projected sensitivity is also given for LZ [48]. The neutrino floor is also shown (brown) [55], which is the threshold at which experiments expect to coherently-detect a single neutrino. The green region represents the excluded parameter space, which includes the unexplained DAMA signal (grey contour) [34].

detectors such as SuperCDMS, and CRESST. While at higher DM masses (1-1000 GeV), xenon dual-phase TPCs lead the way. Positive DM detection has been claimed by DAMA, based on an annually modulating signal [56]. However, this result falls in a region excluded by several other experiments and continues to fail external confirmation.

Significant sensitivity enhancement is projected from XENONnT [41] and LZ [2], both of which are scheduled to go online later in 2020 pending COVID-19 incurred delays. These next-generation experiments will explore much of the parameter space between current sensitivity limits and the neutrino floor. Looking further ahead, generation three experiments such as DARWIN are also in development [50]. The promise of DARWIN is to have a sensitivity which reaches the neutrino floor. In the sub-GeV DM search cryogenic solid-state experiments will also continue to lower energy thresholds and become more sensitive in the low mass region. New techniques for data analysis will also allow dual-phase xenon TPCs to improve sensitivities in this region.

2.5 Other search methods

2.5.1 Collider searches

One of the most powerful tools available to the experimental physics community when it comes to the search for fundamental particles are the collider experiments conducted at particle accelerator facilities [25]. Their successes in discovering the proton, the anti-proton, and the Higgs boson stand as some of the most notable scientific discoveries of the 20th Century [57]. Above and beyond this, the capability of this approach to observe all of the particles in the Standard Model gives hope that they may also aid in discovering particles beyond the Standard Model - including DM.

The most powerful of this class of experiment, is the Large Hadron Collider (LHC) located at CERN in Geneva [14]. This facility uses a 27 km ring of superconducting magnets to accelerate two beams of protons to near the speed of light. These two beams travel in opposite directions, guided by separate beam pipes. These beams continue to be accelerated until their collision would produce a total energy of 13 TeV. Once up to speed the proton beams are “squeezed” together with additional magnets at positions in the ring corresponding to one of the four-particle detectors – ATLAS, CMS, ALICE, and LHCb. From the immense amount of energy released in these collisions (roughly 1.4×10^4 the proton rest mass) an assortment of particles can be produced, including all the particles which are already in the

SM, and potentially many more. If these collisions are also capable of producing DM, then their detection would represent a monumental achievement in particle physics.

Despite the fact that natural particle accelerators (such as supernovae) exist, collider experiments like those found at the LHC have several notable advantages [58]. First is that by building the entire experiment ourselves we already know the initial conditions of the collisions such as the type of particles involved in the collision, and their energies. Secondly, these experiments also provide the capability to produce large numbers of controlled collisions at predetermined times. As a result, collider experiments serve an essential role in DM searches. Specifically, in the event of any direct or indirect detection of DM, verification in the form of DM production at a collider experiment will need to occur.

A key feature of these collider experiments which DM searches rely upon is the alignment of the proton beams [59]. Specifically, both beams must travel directly down the beam pipe with no transverse momentum. This means the sum of transverse momenta of all decay products should also be zero. So while it is expected that a DM particle will not register in the sensors which make up the four detectors on the LHC beam-line, DM will still be detectable via some “missing” transverse momentum. Nevertheless, to date, no excess of events with missing energy, beyond what is expected from the SM, has ever been observed at the LHC or any other collider experiment.

2.5.2 Indirect detection

The indirect detection of DM aims to observe the annihilation or decay products of DM at sites in our universe which are expected to have a high DM density. Examples of these high density DM sites include: the centre of the Sun, the centre of nearby galaxies, and the centre of the Milky Way. DM annihilation is theorised to produce a range of particles such as: gamma-rays, neutrinos, positrons, as well as others which are either unstable or indistinguishable from standard cosmic rays. Looking for these remnants is often limited by our poor understanding of the many backgrounds present in astro-particle searches [60]. However, it is currently the only method for directly probing the WIMP annihilation cross-section $\sigma_{ann} v$. Referred to as the thermal relic cross section, the Fermi-LAT collaboration estimates this value to have an upper limit of $\sigma_{ann} v = 2.2 \times 10^{-26} \text{ cm}^3 \text{ s}^{-1}$ when $M_{DM} > 10 \text{ GeV}$ [61]. This result comes from the observation of gamma-ray data from 15 dwarf spheroidal satellite galaxies orbiting the Milky Way. The approach towards designing indirect detection experiments is varied, and depends on which decay products one wishes to search for [62].

Some experiments attempt to search for a monoenergetic line in the gamma spectrum. One such satellite-borne experiment is Fermi-LAT, which usually targets dwarf galaxies due to their relatively low astrophysical background [63]. Other experiments are ground-based, relying on interactions of cosmic gamma-rays in our atmosphere to create a cascade of secondary particles. This cascade results in a detectable Cherenkov radiation signal in ground-based imaging Air Cherenkov telescopes. These experiments require extensive simulations to estimate the numerous backgrounds present in their searches [64].

Other indirect searches attempt to observe the high-energy neutrinos which are expected to be created in DM annihilation. Examples of these experiments include: Super-Kamiokande [65], Ice-Cube [66], and ANTARES [67]. These experiments look for Cherenkov light generated in ice or water. Some DM models predict a flux of high energy neutrinos coming from the Sun which cannot be explained by SM solar processes. The idea is that if DM were to scatter in the Sun, they may lose enough energy to be captured in the Sun's gravitational potential well. As the DM sinks to the solar core its annihilation would produce an excess of neutrinos [68]. Due to a lack of other production processes, the detection of an excessive muon flux in these experiments would be a strong indication of DM. However, unlike direct detection limits, these results are strongly model-dependent.

The last indirect detection probe of DM is the search for excesses of positrons and anti-protons resulting from DM annihilation [63]. Interestingly, an unexplained positron excess in the interstellar medium has been observed by several experiments [69]. It is unconfirmed whether this excess is due to the annihilation DM. If DM is responsible this observation leads to a troublingly high WIMP annihilation cross-section, higher than astrophysical constraints would permit. In the DM interpretation of this signal, gamma-ray and cosmological observations constrain the WIMP to a mass range of 0.5 to 1 TeV. However, models which explain this excess originating from nearby pulsars have also been suggested [70]. Whether the explanation of this excess is DM, pulsars, or another source will be tackled by the next generation of satellite-borne experiments. Specifically, the experiments NUCLEON, CALET, and DAMPE will strive to provide clarity on the cause of this positron anomaly [71].

2.6 Effective field theory of dark matter detection

As described in Section 2.4 the main effort in the direct detection of DM has traditionally gone into SI and SD searches. However, almost nothing is currently known about non-gravitational DM-nucleon interactions. This leaves the parameter spaces of several possible

DM-nucleon interactions unrestricted. Subsequently, searches for DM interactions within EFT theories are performed to limit the strength of EFT operators.

The EFT search relies upon a non-relativistic effective theory which describes all possible WIMP-nucleon interactions, without the need for an underlying high energy model. It consists of a set of four-particle-contact interaction operators (\mathcal{O}), each describing a unique DM-nucleon interaction with respect to momentum conservation and Galilean-invariance. These operators are developed from a basis of four Hermitian quantities:

$$i\vec{q}, \quad \vec{v}^\perp \equiv \vec{v} + \frac{\vec{q}}{2\mu}, \quad \vec{j}_\chi, \quad \vec{j}_N \quad (2.8)$$

where \vec{q} is the momentum transferred from the nucleon to the WIMP, \vec{v}^\perp is the component of the relative incoming velocity between the WIMP and the nucleon perpendicular to that momentum transfer, j_χ is the spin of the WIMP, and j_N the spin of the nucleon. Combining these quantities yields fourteen independent EFT operators [72]:

$$\begin{aligned} \mathcal{O}_1 &= 1 \\ \mathcal{O}_2 &= (\vec{v}^\perp)^2 \\ \mathcal{O}_3 &= i\vec{j}_N \cdot (\vec{q} \times \vec{v}^\perp) \\ \mathcal{O}_4 &= \vec{j}_\chi \cdot \vec{j}_N \\ \mathcal{O}_5 &= i\vec{j}_\chi \cdot (\vec{q} \times \vec{v}^\perp) \\ \mathcal{O}_6 &= (\vec{j}_\chi \cdot \vec{q}) (\vec{j}_N \cdot \vec{q}) \\ \mathcal{O}_7 &= \vec{j}_N \cdot \vec{v}^\perp \\ \mathcal{O}_8 &= \vec{j}_\chi \cdot \vec{v}^\perp \\ \mathcal{O}_9 &= i\vec{j}_\chi \cdot (\vec{j}_N \times \vec{q}) \\ \mathcal{O}_{10} &= i\vec{j}_N \cdot \vec{q} \\ \mathcal{O}_{11} &= i\vec{j}_\chi \cdot \vec{q} \\ \mathcal{O}_{12} &= \vec{j}_\chi \cdot (\vec{j}_N \times \vec{v}^\perp) \\ \mathcal{O}_{13} &= i \left(\vec{j}_\chi \cdot \vec{v}^\perp \right) \left(\vec{j}_N \cdot \frac{\vec{q}}{m_N} \right) \\ \mathcal{O}_{14} &= i \left(\vec{j}_\chi \cdot \frac{\vec{q}}{m_N} \right) \left(\vec{j}_N \cdot \vec{v}^\perp \right) \\ \mathcal{O}_{15} &= - \left(\vec{j}_\chi \cdot \frac{\vec{q}}{m_N} \right) \left((\vec{j}_N \times \vec{v}^\perp) \cdot \frac{\vec{q}}{m_N} \right) \end{aligned} \quad (2.9)$$

The operator \mathcal{O}_2 is neglected because it cannot arise at leading order in a non-relativistic scenario. Each of these operators can in principle be coupled differently to protons vs neutrons giving 28 unique couplings.

These operators are used to build an interaction Lagrangian:

$$\mathcal{L}_{int}^{(N)} = \sum_i c_i^{(N)} \mathcal{O}_i^{(N)} \quad (2.10)$$

where (N) denotes the type of nucleon (n or p) for the interaction, and c_i is the coupling coefficient of operator \mathcal{O}_i . c_i is the parameter responsible for determining the likelihood of that particular DM-nucleon interaction occurring. The terms of this Lagrangian take the form:

$$\bar{\chi} \mathcal{O} \chi \bar{N} \mathcal{O}_N N \equiv \mathcal{O} \bar{\chi} \chi \bar{N} N \quad (2.11)$$

where N and χ are non-relativistic fields, the former for the target nucleon, and the latter for the incident WIMP. In this formulation, \mathcal{O} can be seen to couple DM to nucleons, with c_i governing the likelihood of the interaction.

Direct detection experiments are able to use their null results to place limits on c_i . However, in reality, DM coupling does not occur directly to a nucleon, but instead with the entire nucleus. Accounting for this begins with looking at six independent nuclear response functions, which depend heavily on the target isotope being used. Each nuclear response contributes towards the total scattering amplitude (\mathcal{M}). These response functions are denoted as $k = \{M, \Sigma', \Sigma'', \Delta, \bar{\Phi}, \text{ and } \Phi''\}$ and can be found in [72].

Arriving at \mathcal{M} from k follows a lengthy process discussed in [72, 73], which I will briefly summarise. It begins with the formation of nuclear response form factors ($F_k^{(N,N')}$) and interference form factors ($F_{(k_1,k_2)}^{(N,N')}$) from k . These terms sum over nucleon pairs $(N, N') = \{(n, n), (n, p), (p, n), (p, p)\}$ to account for two-body currents in the nucleus. Form factors ($F_{(i,j)}^{(N,N')}$) are then formed out of linear combinations of $F_k^{(N,N')}$ and $F_{(k_1,k_2)}^{(N,N')}$. $F_{(i,j)}^{(N,N')}$ is expressed as $F_{(i,j)}^{(N,N')}(v^2, q^2)$ on account of their velocity and momentum dependence. These values depend on the target medium being used [74]. In using $F_{(i,j)}^{(N,N')}(v^2, q^2)$ to calculate \mathcal{M} this series of form factors is combined according to Appendix A.2 in [72]. These form factors, $F_{ij}^{(N,N')}(v^2, q^2)$, are scaled based on the value of the EFT coupling constants $c_i^{(N)}$ in the calculation of the scattering amplitude for the WIMP-nucleus interaction:

$$\frac{1}{2j_\chi + 1} \frac{1}{2j_A + 1} \sum_{\text{spins}} |\mathcal{M}|^2 \equiv \frac{m_A^2}{m_N^2} \sum_{i,j=1}^{15} \sum_{N,N'=p,n} c_i^{(N)} c_j^{(N')} F_{ij}^{(N,N')}(v^2, q^2), \quad (2.12)$$

where j_A is the spin content of the target nucleus, j_χ is the WIMP spin, c_i is the coefficient of operator \mathcal{O}_i in the Lagrangian. Due to the presence of two couplings in each term, there is a possibility for destructive interference to occur. Typically this possible interference is made manageable by assuming one coupling is dominant over all others. This leads to the formula for scattering event rate:

$$\frac{dR}{dE_R} = \frac{\rho_{DM}}{32\pi m_{DM}^3 m_N^2} \int_{v>v_{min}} \frac{f(\vec{v})}{v} c_i^{(N)^2} F_{i,i}^{(N,N)}(v^2, q^2) v, \quad (2.13)$$

Figure 2.9 uses Eq. 2.13 to determine the recoil spectra in xenon for a set of EFT operators using three different WIMP masses [74]. As shown, momentum-independent interactions $\frac{dR}{dE_R}$ peaks at $E_R = 0$ due to kinematics. However, momentum-dependent operators will peak at higher values of E_R . The consequence of this result is that for the EFT search the energy range over which backgrounds need to be reduced is much larger than it was when compared to the SI and SD searches.

2.7 Summary

The existence of dark matter is well accepted by most physicists, with several well motivated models used to explain various astronomical and cosmic anomalies. Several methods for the detection of DM are being pursued, with maturation in detector technology continuing to limit backgrounds, and to drive improvement in sensitivity. This leaves hope for the field of experimental DM searches to provide strong evidence for DM in the form of a positive signal.

The remainder of this thesis will focus on the direct detection of DM via two different dual-phase xenon TPC experiments. Chapter 2 will give an overview of this technology, with a focus on the LUX and LZ detectors. The remaining chapters will then be dedicated to original work related to the direct detection of DM.

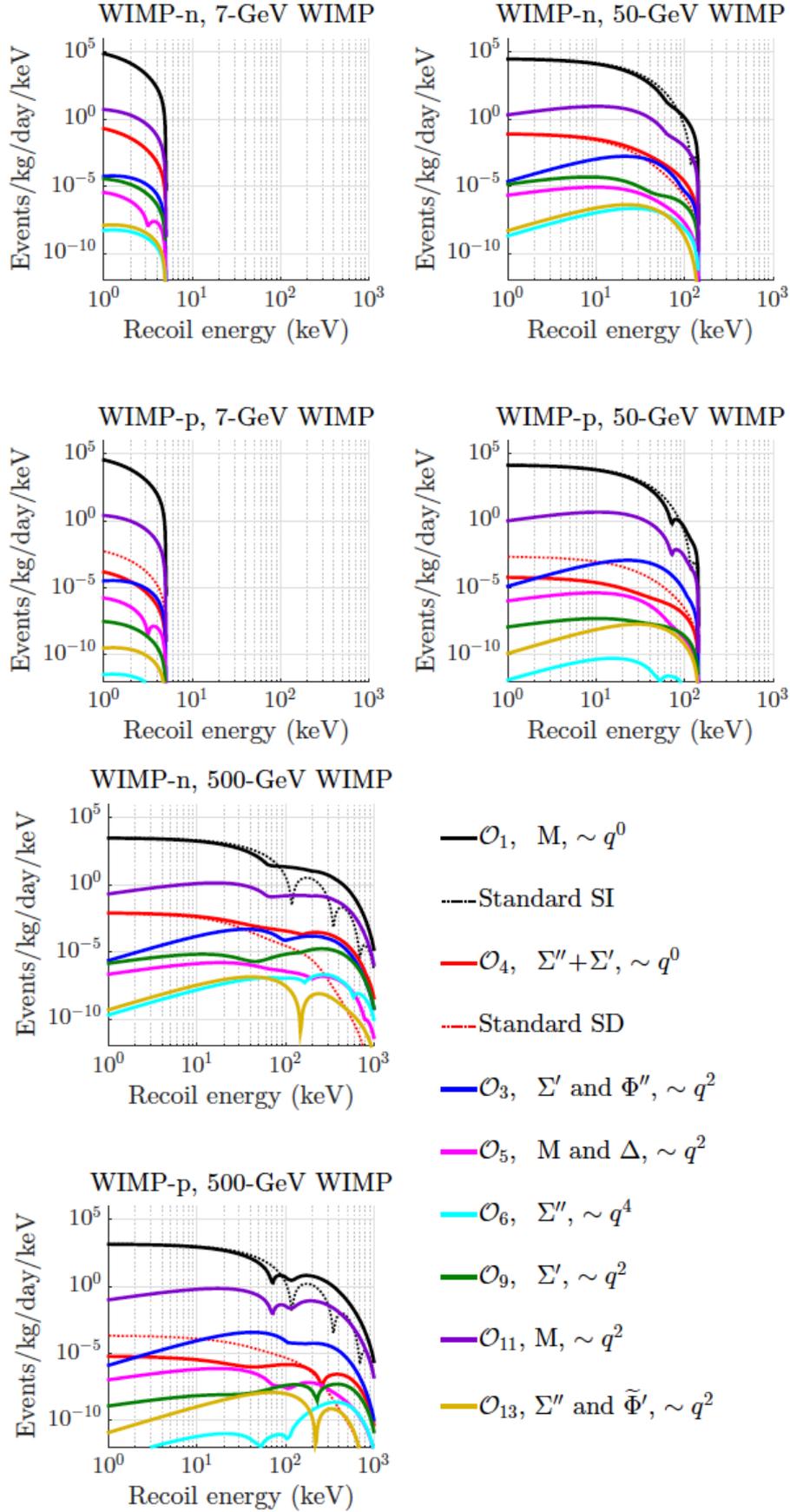


Figure 2.9 WIMP-nucleon recoil spectra in xenon for a representative sample of EFT operators, weighted by the natural abundance of isotopes [74]. For comparison, the recoil spectra for the standard SI and SD WIMP-nucleon interactions are displayed.

Chapter 3

The LUX and LZ experiments

3.1 Dual-phase xenon time projection chambers

3.1.1 Concept

As mentioned in Chapter 2 the best sensitivity limits for GeV range WIMPs are currently held by dual-phase Xe TPCs [42]. These devices use a liquid xenon (LXe) volume as a target and look for interactions from the incident WIMPs in our galactic halo discussed in Section 2.4.3. Within the LXe of these detectors, there are two types of recoils produced by particle interactions, nuclear recoils (NR), and electron recoils (ER) [75]. The type of recoil depends on the allowed interactions between an incident particle and the target. Most background radiation (i.e. gamma-rays from beta decay) will result in ERs. However, WIMPs along with a small fraction of backgrounds (i.e. neutrons, and the coherent neutrino-nucleus scattering) will produce NRs.

Fortunately, ERs and NRs produce different signatures which can be distinguished from one another. This is because of the differing track structures of recoiling electrons and nuclei in a LXe medium. The shorter tracks of recoiling nuclei in NR events yields a higher density of collisions with the medium as compared to ER events. Furthermore, a significant fraction of an NR event's energy is dissipated via heat¹. The result of these differing track structures is a different ratio of energy being produced via the ionisation and scintillation channels. This feature of dual-phase Xe TPCs is used to powerful effect as a background reduction measure, the implementation of which will be discussed in Section 3.1.6. However, before discussing this, several topics must first be introduced. This will begin with an introduction to the observed signals of a dual-phase Xe TPC in Section 3.1.2, followed by

¹Although dual-phase Xe TPCs are blind to heat more than enough scintillation and ionisation signals are generated to make events observable.

a more detailed discussion of the scintillation signal in Section 3.1.3, and the ionisation signal in Section 3.1.4. How these signals are used to reconstruct energy depositions will be discussed in Section 3.1.5, before moving onto ER/NR discrimination in Section 3.1.6. This section will finish with a description of another one of LXe’s powerful background reduction techniques called self-shielding in Section 3.1.7.

3.1.2 Ionisation and scintillation signals

Figure 3.1 shows a signal production diagram for an incident particle interacting with a Xe nucleus [2] in a dual-phase Xe TPC. The diagram shows the production of both a prompt scintillation signal ($S1$) and a delayed ionisation signal ($S2$). As indicated, the $S1$ is detected almost instantly by two sets of PMT arrays: one located at the bottom of the TPC; and the other at the top, above the LXe level in a gaseous Xe phase. $S2$ detection on the other hand only occurs once the ionised electrons have “drifted” into the gas phase, where they are accelerated by an extraction grid to produce an electroluminescent signal. The emission spectrum of this signal is similar to the $S1$ and is also detected by the PMTs. In this idealised figure, the drifting of ionised electrons is facilitated by an unillustrated vertical electric field with strength² $\mathcal{O}(100)$ keV/cm. This electric field is created and maintained by a cathode grid located just above the bottom PMT array; an anode grid slightly below the phase barrier; and field shaping rings surrounding the TPC.

3.1.3 $S1$ production

A closer examination of $S1$ shows that it is in fact the combination of two different signals: direct excitation, and recombination. However, since the time delay between these two scintillation signals is unresolvable they are not considered independently.

Direct excitation of Xe occurs because of an ER or NR event. The process leads to a Xe atom becoming electrically excited (Xe^*) by an interaction with a recoiling nucleus or electron (X), as illustrated in Eq. 3.1. The Xe^* then combines with a neighbouring neutral atom to form a diatomic excited molecule (Xe_2^*) called a dimer (Eq. 3.2). Here, a vibrational excitation and relaxation occurs which is not represented in the equation. This relaxation results in a largely kinetic emission³ signified as “heat” in Eq. 3.2. The creation of Xe_2^* leads

²The field must be strong enough to prevent ionised electrons from recombining with their ions, but weak enough to prevent spontaneous ionisation.

³A small amount of this emission is expected to occur in the near-infrared range, but this is not considered to be part of the $S1$.

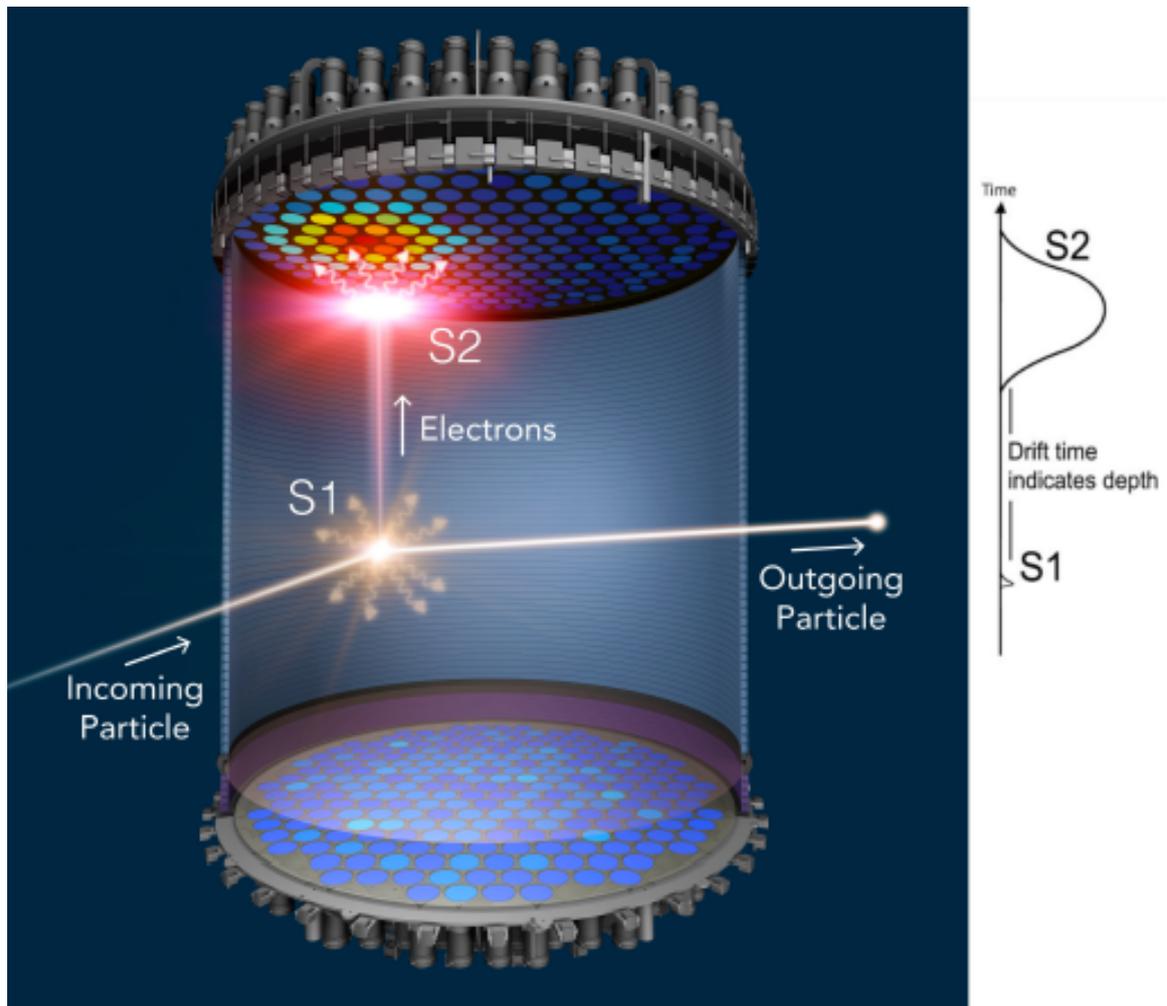


Figure 3.1 Operating principle of a dual-phase Xe TPC [2]. A particle interacting in the LXe target volume produces two signals: a prompt scintillation ($S1$), and a delayed signal caused by ionised electrons generating electroluminescent photons in the Xe gas phase ($S2$). Dual signal detection ($S1$ and $S2$) facilitates accurate 3D vertex reconstruction, and discrimination between nuclear and electron recoils.

to the production of a VUV scintillation photon ($h\nu$) upon nuclear de-excitation (Eq. 3.3):



The decay time of the Xe_2^* which produces the VUV photon depends on which of the two possible transitions produced it: the singlet state ($^1\Sigma_\mu^+$) to ground ($^1\Sigma_g^+$), or the triplet state ($^3\Sigma_\mu^+$) to ground. For the singlet state, this decay time⁴ is 4.3 ns, and for the triplet state, it is 22 ns [76]. Regardless of whether the Xe_2^* transitions via the triplet or singlet state the process dissociates the Xe_2^* into its constituent Xe atoms. This accounts for Xe being transparent to its own scintillation light since the probability of finding another Xe_2^* is low.

Recombination is the other process leading to VUV emission and occurs when an ionised electron (e^-) reforms with its dimer. In this process, the recoiling nucleus or electron does not excite the Xe atom but instead ionises as shown in Eq. 3.4. A dimer is then formed when a Xe^+ ion interacts with a neighbouring Xe atom (Eq. 3.5). Following this, the electron from Eq. 3.4 does not drift away, and instead recombines with the dimer leading to an excited atomic state (Eq. 3.6). In this equation, the double star in the superscript indicates a higher electronically excited state of the atom than the excited state indicated by a single star. The excited atom then forms a molecular excimer with a neighbouring atom (Eq. 3.7 and Eq. 3.8). The process ends with the excimer decaying along the same de-excitation path as the dimer did during direct excitation, once again emitting a VUV photon (Eq. 3.9).



⁴LAr, by comparison, has a longer triplet lifetime ($\sim 1 \mu\text{s}$) and a shorter singlet lifetime by a factor of ~ 1000 . This difference in lifetimes explains why LAr experiment can use this feature for ER/NR identification. LXe detectors such as ZEPLIN-I have used this approach, but better ER/NR discrimination methods exist.

The fraction of a given $S1$ which consists of recombination light depends on three factors: the type of recoil (ER or NR), recoil energy, and the electric field strength. The first two of these factors can be explained by considering the kinematics of a recoiling particle. As mentioned in Section 3.1.1, when comparing the track length of ER and NR events of the same energy, one will see that the more massive nucleus of an NR event will not travel as far as the recoiling electron in an ER event. Similarly, if we control for the type of recoiling particle, lower recoil energies will result in lower recombination fractions. The simple explanation for this is that the further apart an electron is from its parent nucleus, the greater the chance is that the electric field will be strong enough to ionise it. The final factor affecting the recombination fraction is the electric field strength used to drift ionised electrons from interaction sites. The argument for the influence of this is simple, using a higher electric field strength increases the ionisation fraction, since more electrons will be “drifted” away from interaction sites and are prevented from recombining.

3.1.4 $S2$ production

The ionisation electrons which make up the $S2$ drift towards the gas phase under the force provided by an electric field [77]. For electrons to be extracted into the gas they must overcome a potential barrier at the phase boundary. To facilitate this an electric field strength several times higher than in the LXe bulk is applied by an electron extraction grid. The region affected by this grid is small, beginning slightly below the gas/liquid interface, and ending slightly above.

The subsequent acceleration experienced by electrons as they enter the gas phase produces electroluminescent photons, with an emission spectrum similar to the $S1$. This is caused by collisions between the electrons and Xe atoms in the gas phase, leading to the creation of excimers. As these excimers return to their ground state VUV photons are emitted [42]. The number of photons produced is dependent on the Xe pressure, the electric field strength in the gas, and the number of electrons traversing the gas phase. As these photons are emitted they are detected by PMTs, creating a signal proportional to the ionisation yield of the initial collision ($S2$ signal).

The $S2$ is several orders of magnitude larger than the $S1$, and experiences a delay proportional to the depth of the interaction event. This delay, along with the $S2$ -photo-hit-pattern in the top PMT array allows for XYZ position reconstruction.

3.1.5 Energy reconstruction

Energy reconstruction is based on the measured $S1$ and $S2$ of an event. However, to perform this calculation requires assuming whether the recoiling particle is an electron or a nucleus. This is because NR events are less effective at producing excitation and ionisation signals than ER events of the same energy, a phenomenon known as quenching. Since the type of recoiling particle cannot be known a priori two different methods for energy reconstruction are used covering both cases. To distinguish between them the units keV_{ee} and keV_{nr} are used for ER and NR reconstructions respectively. In this thesis where no subscript is applied keV_{ee} is assumed. Depending on whether the event is ER or NR, the energy deposited by an incident particle (E_{er} or E_{nr}) can be expressed as:

$$\begin{aligned} E_{er} &= W \cdot (n_e + n_\gamma) \\ E_{nr} &= \frac{W}{f_n} \cdot (n_e + n_\gamma) \end{aligned} \quad (3.10)$$

where f_n is the fraction of the recoil energy that is transferred to excitations and ionisations instead of atomic motion (a.k.a. quenching factor). f_n increases with respect to E_{nr} , from a value of approximately 0.15 at 1 keV_{nr} to 0.30 at 100 keV_{nr} [78]. W ($W = 13.7 \pm 0.2$ eV) is the average energy required to produce a single exciton (Eq. 3.1) or an ion (Eq. 3.4), n_e is the number of ionisation electrons which are extracted into the gas phase, and n_γ is the number of prompt scintillation photons [77]. In this thesis E_{nr} is largely ignored, thus E should be interpreted as E_{er} unless otherwise stated.

Photon and electron yields are also calculated with the use of the factors g_1 , in units of detected photons per emitted photon (phd/γ), and g_2 in phd per free electron (phd/e^-). The first factor, g_1 , is the product of the average photon collection efficiency and the average PMT quantum efficiency. The second factor, g_2 , is the product of the electron extraction efficiency at the liquid-gas surface and the number of photons detected per extracted electron respectively. Incorporating this into Eq. 3.10 gives:

$$E = W \cdot \left(\frac{S1}{g_1} + \frac{S2}{g_2} \right) \quad (3.11)$$

3.1.6 Electron recoil event discrimination

ER discrimination is done by looking at the differing scintillation and ionisation yields of ER and NR events as a function of $S1$. In Figure 3.2, the $S2/S1$ ratios of NR and ER events from LUX calibration studies are displayed as two bands. In this diagram, NR events are seen to

possess a lower charge-to-light ratio because of the shorter recoiling particle track length, as discussed in Section 3.1.3. This ratio of excitations to ionisations (N_{ex}/N_{ion}) is quite different for ER and NR events of the same energy. For an NR, this ratio is ~ 1 , whereas for an ER it is between⁵ ~ 0.06 and ~ 0.2 . This difference is the origin of the distinct $S2/S1$ bands for ER and NR events.

The result of this is that the fraction of $S2$ to $S1$ light depends on whether the event was an NR or an ER. Subsequently, the measurement of $S2/S1$ is used for the discrimination of ER events. ER rejection is an essential step in WIMP identification, since traditional WIMPs cannot scatter on electrons. In contrast to this most backgrounds (i.e. gamma-rays from beta decay) will lead to ERs (see Section 3.4 for more detail). So effective is ER rejection at background reduction that by itself dual-phase Xe TPC experiments have been able to achieve background rejection levels ranging from 99.5% for XENON10 [79], to 99.99% for ZEPLIN-III [80].

3.1.7 Xenon self-shielding

Critical to any low background experiment is the material screening and selection program to control the trace radioactivity of the detector components. However, for dual-phase Xe TPCs backgrounds emanating from detector components are also removed by the self-shielding offered by LXe.

Self-shielding relies on both the density of LXe (2.88 g cm^{-3}), and the high resolution of interaction sites in three dimensions [2]. The concept of self-shielding is to exclude signals detected in the outer layer of LXe from data analysis, because the outer layer is more affected by backgrounds present in detector components. The data events recorded in the remaining inner LXe benefit from an extremely low background contribution, because of the shielded by the sacrificed layer. The size of this non-fiducial layer need only be a few centimetres thick, which makes scaling up the volume of this technology highly effective. This is exemplified by the comparison of LUX to LZ, which have $\sim 40\%$ and $\sim 80\%$ fiducial mass respectively. The amount of LXe sacrificed to self-shielding depends on two factors. The first is the minimisation of external radioactivity penetrating into the fiducial region. This depends on the mean free path of these particles, and the effectiveness of external veto systems to tag these events. The second is the accurate position reconstruction of interaction sites, to ensure events outside the fiducial region do not “leak” in.

⁵ ~ 0.06 has been calculated for LXe, but 0.2 is found to be consistent with experimental data [42]. Often an average value of 0.13 is used as an estimate. For LAr calculations and experiment agree on a value of ~ 0.2 .

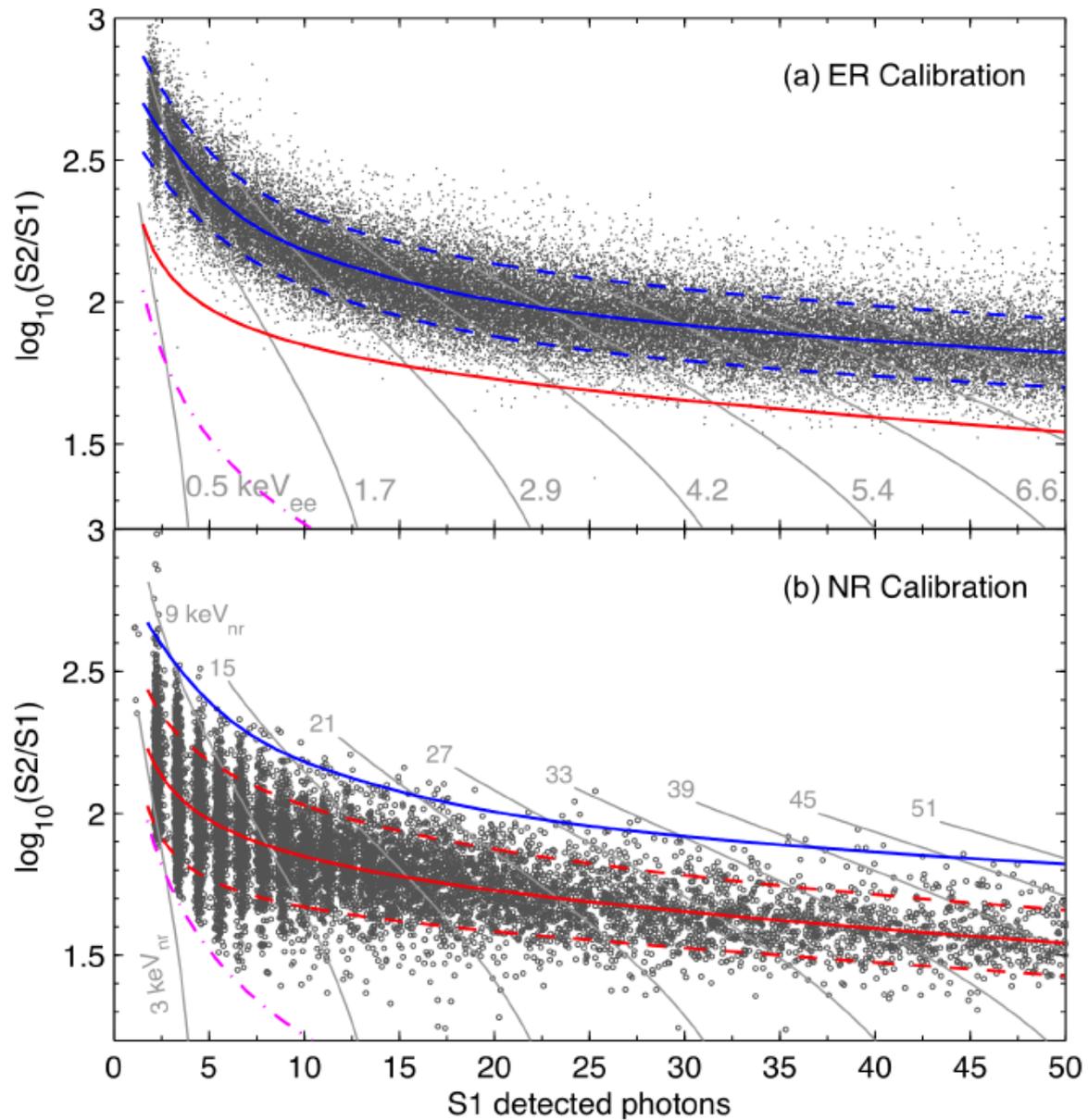


Figure 3.2 Discrimination parameter $\log_{10}(S2/S1)$ as a function of $S1$ obtained with LUX calibration data [2]. The upper figure shows the ER band, calibrated with beta decays from a dispersed tritium (^3H) source. The median is shown in blue, with 80% population contours indicated by the dashed blue lines. The lower figure shows the NR band populated by elastic neutron scattering from a deuterium-deuterium pulsed neutron source (D-D). The median is shown as a solid red line and 80% band width in dashed red lines. The grey lines in each plot show fixed energy depositions. In the upper plot this is for ERs in keV_{ee} , and in the lower plot NRs in keV_{nr} .

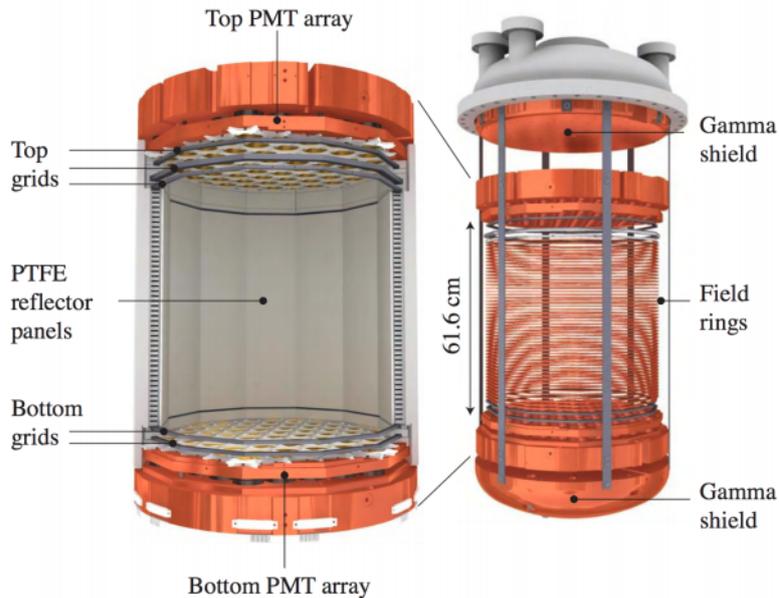


Figure 3.3 The LUX TPC. In this figure, the locations of PMT arrays, field rings, shielding, and the grids, are explicitly indicated. The ICV lid is included in the illustration on the right [81].

3.2 The LUX experiment

3.2.1 Overview

Located 1.5 km underground at the Sanford Underground Research Facility (SURF), the LUX detector (as illustrated in Figure 3.3) was a dual-phase Xe TPC designed for the detection of WIMPs. To achieve this, LUX consisted of 250 kg of active LXe (370 kg total, 100 kg fiducial), housed in a 47 cm in diameter Inner Cryostat Vessel (ICV), with 48 cm between the cathode and gate grids. For signal acquisition, 122 Hamamatsu R8778 PMTs were split into two arrays at the top and bottom of this volume. The TPC was lined with Polytetrafluoroethylene (PTFE) to improve the internal reflection of $S1$ light. An outer Cryostat Vessel (OCV) was used to hold a vacuum and insulate the inner vessel. Both cryostats were made from titanium.

To increase the shielding the outer cryostat was housed in a water tank. The presence of the water significantly reduced the flux of neutrons and gamma-rays from sources outside the detector. The water tank was so successful in its role that detector components were left as the dominant background. The tank held 3×10^5 litres of water around the outer cryostat and was outfitted with 20 Hamamatsu R7081 PMTs for active muon vetoing [82]. Additional

shielding from rock gammas was provided by an inverted steel pyramid placed beneath the water tank. This structure was ~ 42 cm wide at the top, narrowing to ~ 2 cm at the base; and constructed from 6 octagonal steel plates, each ~ 5 cm thick [47].

3.2.2 Timeline and results

Data collection was divided into four acquisition runs. The first two runs (Run01, and Run02) were above-ground tests to ensure the detector was working correctly [83], the second two (Run03, and Run04) were WIMP search runs conducted underground [47]. Run03 collected 85 days' worth of data, and Run04 collected 332 days for a combined 3.35×10^4 kg days' worth of data. In 2016 LUX published a world-leading upper limit on the SI WIMP-nucleon interactions, with a cross-section of 1.1×10^{-46} cm² for a 50 GeV/c² WIMP.

3.2.3 Electric field

The presence of a time-varying electric field in Run04 complicated analysis efforts [84]. The introduction of a time dependence on the electric field strength appears to have begun during a grid conditioning campaign conducted between Run03 and Run04. Grid conditioning was conducted to improve the extraction efficiency of electrons from the liquid into the gas, by burning off impurities on the cathode, gate, and anode grid. To do this, the voltage was set to a level slightly above the point of discharge and left for an extended period. It was hoped that the increased field strength from this effort would improve S2 identification, particularly for low energy events, and improve the detector's threshold with an increased field through the gas phase. At first, the conditioning was seen as a success. With the extraction field increasing from 2.9 kV·cm⁻¹ to 3.5 kV·cm⁻¹, leading to an improvement in electron extraction efficiency of $49 \pm 3\%$ to $73 \pm 4\%$ [47]. However, as shown in Figure 3.4 ^{83m}Kr calibration data revealed that the near-vertical electric field lines from Run03 had been significantly altered. Unexpectedly, drifting electrons now appeared to be receiving a significant radial "push" towards the centre of the TPC. Furthermore, the magnitude of this effect appeared to have an azimuthal dependence and varied with time. This effect was rectified with detailed mapping of the electric field, and shown to be consistent with the build-up of negative charge on the PTFE walls [85]. Although this effect did not end up preventing LUX from publishing a world-leading WIMP sensitivity limit, it did hamper efforts. Subsequently, this example motivates future experiments to pay extra attention to electron trajectories.

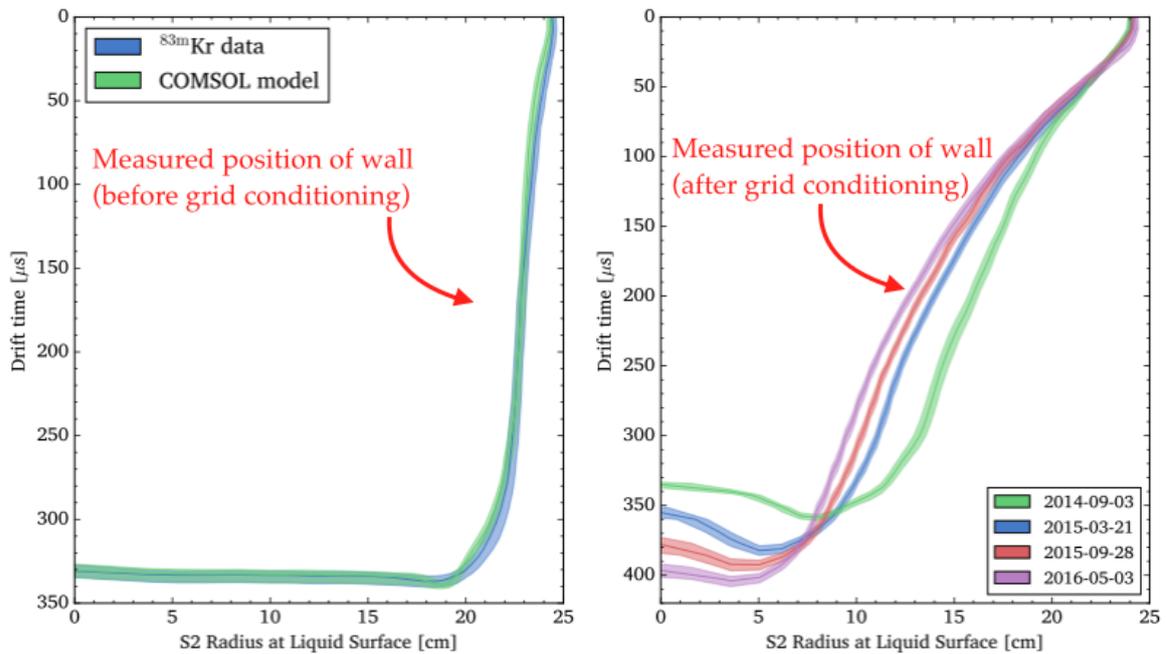


Figure 3.4 LUX wall position based on $^{83\text{m}}\text{Kr}$ data. The plot on the left shows Run03 data (before grid conditioning). The plot on the right shows various wall positions through 2014-2015 (after grid conditioning) [84].

3.3 The LZ experiment

3.3.1 Overview

LZ is the successor experiment of LUX scheduled to go online in 2020 (pending COVID-19 related delays). Although LZ and LUX share a common detection principle, LZ is ~ 20 times larger than its predecessor and possesses several new developments. LZ aims to achieve a SI WIMP-nucleon cross-section sensitivity of $1.4 \times 10^{-48} \text{ cm}^2$ for a $40 \text{ GeV}/c^2$ mass WIMP⁶ (90% CL) [48]. This section will describe LZ by focusing on the differences from its predecessor. As shown in Figure 3.5, the crucial update to LZ is the increased target mass. 10 tonnes of LXe will now be used, with 7 tonnes being actively monitored, and 5.6 tonnes in the fiducial volume. LZ will also have larger PMT arrays, with 253 Hamamatsu R11410-22 PMTs in the top array and 241 in the bottom [2]. Other key upgrades include a gadolinium-loaded (Gd) liquid scintillator anti-coincidence neutron veto system (described in Section 3.3.2), and a LXe skin region (described in Section 3.3.3).

LZ also aims to analyse data in real-time. As such, significant effort has gone into de-

⁶A factor of $100\times$ better than LUX.

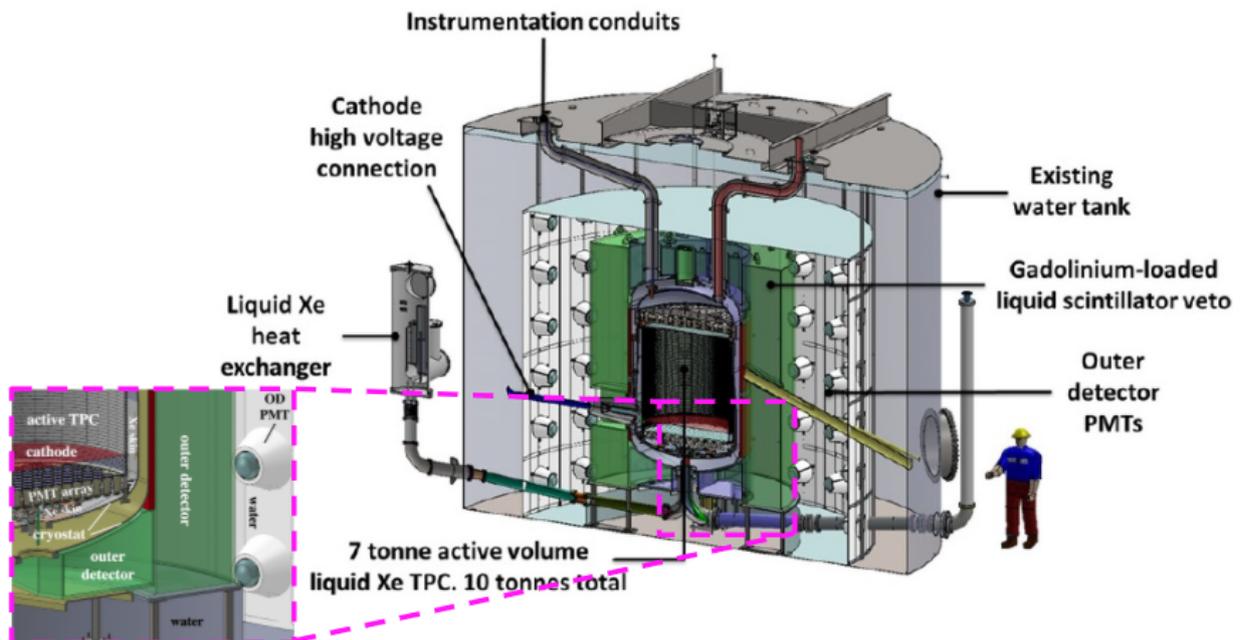


Figure 3.5 Overview of the LZ experiment showing the main components. Key differences to LUX are the size of the TPC, the LXe skin, and a liquid scintillator veto placed inside the water tank and the cathode high voltage connection [2].

veloping LZ's data analysis package (LZap). This was tested with several Mock Data Challenges (MDC) between 2017 and 2020. These two aspects (LZap and the MDCs) were a significant focus for many LZ researchers, and are described in Sections 3.3.4 and 3.3.5.

3.3.2 Outer detector

To further shield against external neutrons from detector components, LZ has implemented an outer detector (OD) system. The system consists of 20.8 tonnes of Gd-loaded linear-alkylbenzene (LAB) scintillator, contained in segmented clear acrylic tanks that surround the OCV. As seen in Figure 3.5, the entire system is located within the water tank previously used by LUX. The LAB scintillation signal is provided by a wavelength shifter and fluor which were added to the solvent. This scintillation is detected by a PMT array of the water tank, located outside the acrylic tanks. This allows the OD to both tag muons as they produce Cherenkov light in the water, and tag the neutron capture on Gd (or other materials) [2].

The Gd dissolved in the LAB facilitates the detection of neutrons. Roughly 15% of this Gd are the isotopes ^{157}Gd and ^{155}Gd . These isotopes have thermal neutron capture cross-sections of 254 kb and 61 kb respectively, making the probability of neutron capture around 1 order

of magnitude more likely than it would be for hydrogen⁷.

For an incoming neutron to mimic the NR of a WIMP in LXe, it must have an energy between roughly 0.5 and 5.0 MeV [2]. For LZ, most neutrons in this energy range are from (α, n) processes⁸, originating from materials around the edges of the Xe. A problematic neutron will enter the LXe TPC, scatter once, then leave the LXe again. Since most of these neutrons will traverse the intervening material, we expect a high proportion to thermalise with hydrogen before being captured by the Gd of the OD. We expect this since the length scale for thermalisation and capture is just a few centimetres, with a typical capture time of $\sim 30 \mu\text{s}$. In order to veto these events, any energy deposition in the LXe within 500 μs of an OD detection will be removed.

3.3.3 LXe skin region

The LXe skin is a 2 tonne region of LXe surrounding and optically insulated from the active target LXe mass. It is instrumented with PMTs to provide a veto in conjunction with the OD. The side skin region is divided into two sections. The first consists of a 4 cm to 8 cm layer of LXe located between the field cage and the ICV. The second is the dome skin, with an even thicker LXe layer, and is situated below the bottom PMT array.

3.3.4 LZap

LZap is the analysis software purpose-built for the event reconstruction of LZ's raw data acquisition. LZap is designed to take in raw waveform data, and return physics-ready reduced quantities (RQs). The tasks which LZap performs are:

1. To read in raw waveforms, and translate from raw Analog to Digital Conversion (ADC) counts to units of phd/ns.
2. Perform pulse finding on the summed waveforms for each PMT.
3. Pulse classification (e.g. $S1$, $S2$, single electron, etc.).
4. Evaluate pulse Reduced Quantities (RQ) (e.g. area, amplitude, width, etc.).
5. Interaction type identification (e.g. single scatters, multiple scatters, etc.).

⁷Although the cross-section for Gd is $\sim \times 1000$ larger than hydrogen, the difference in concentrations makes it only one order of magnitude different in the end.

⁸Also from spontaneous fission of ^{238}U , but these neutrons can be tagged because they are generated in conjunction with gamma-rays.

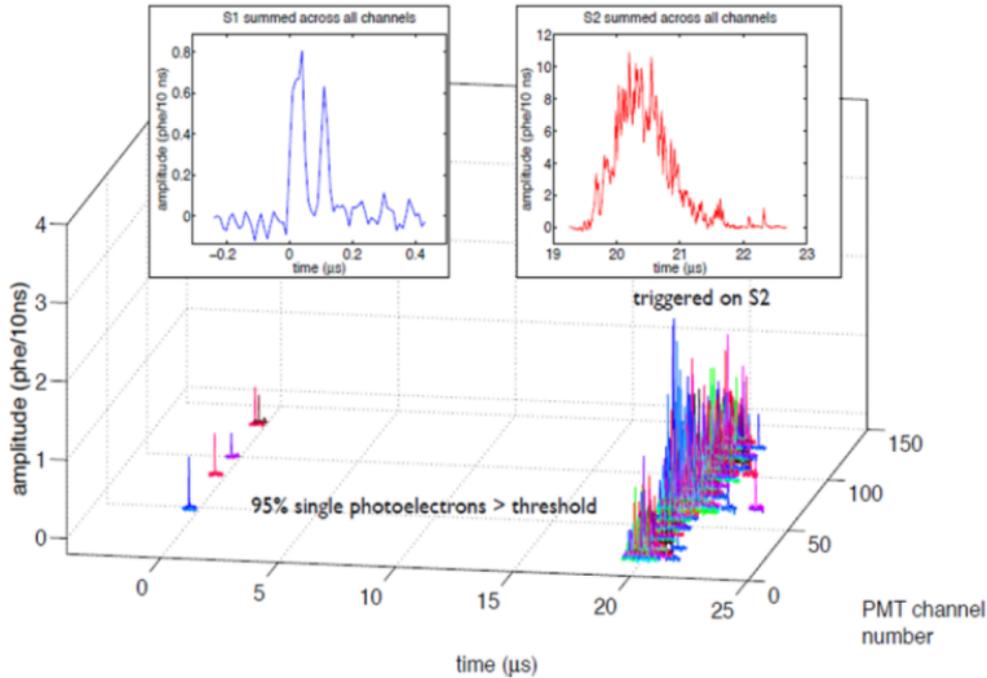


Figure 3.6 A 1.5 keV ER event taken from a three month above ground commissioning of LUX [2]. The image shows a five-fold PMT hit coincidence for $S1$, and the corresponding, and larger, $S2$ delayed by 20 μs .

6. 3D position reconstruction.
7. Apply relevant corrections.
8. Return RQs for each identified event.

The difference between the types of pulses that an $S1$ and an $S2$ can generate in the PMTs can be seen in Figure 3.6. In this example, $S1$ can be seen to have a high amplitude-to-duration ratio, whereas the $S2$ despite having a larger amplitude and duration has a lower amplitude-to-duration ratio. $S1$ and $S2$ classification is more involved than this and moves beyond the scope of this thesis, but the example serves to illustrate some of the basic heuristics which are used to perform pulse shape identification. Aside from $S1$ and $S2$, other pulse classifications are single electron (identifiable via short and quick pulse), and electron trains (usually a tail which follows an $S2$ pulse).

3.3.5 Mock data challenges

To ensure that LZ is physics-ready at the beginning of data taking, a series of Mock Data Challenges (MDC) are being conducted. These MDCs tested and developed the infrastructure

required, and trained LZ users on data analysis. Three MDCs were scheduled for LZ.

MDC1 was conducted between June and September of 2017 and primarily consisted of 1 month of simulated background data. Additional features included a time-varying Xe purity, and a $^{131\text{m}}\text{Xe}$ calibration source of uniform 1 Hz decay rate. During MDC1 Monte Carlo (MC) truth information was available to analysts. The main goals of MDC1 were to test LZap's data processing end-to-end and to produce full waveforms in the LZ DAQ output format.

MDC2 was conducted between April and September of 2018 and consisted of 6 months' worth of simulated WIMP search data. MDC2 included several new features including a radially dependent electric field and position-dependent electron drift velocities. MDC2 utilised time-dependent $^{131\text{m}}\text{Xe}$, $^{83\text{m}}\text{Kr}$ injections throughout the 140 simulated days of live time. It also included AmLi, Deuterium-Deuterium (D-D), and ^{22}Na source calibrations. Another difference to MDC1 was the blinding of analysts to the nature of physics signals that would be present. Although MC truth information was eventually made available to analysts, this was only as a final cross-check. The overarching goal of MDC2 was to test a near-complete set of analysis tools used to set a limit/discovery results.

MDC3 ran from summer 2019 until spring 2020 with the task of analysing realistic commissioning and initial science run data for LZ. This MDC included all planned internal and external calibrations. It also blinded analysts from MC truth information and released mock data in real-time. A bias mitigation scheme of including salt events was also conducted in MDC3. "Salting", as it is known, involves the injection of fake signal events into the data stream. The goal is to check that data analysis cuts do not disproportionately remove these salt events. The goal of this final MDC was to test the completed LZap, providing a dry run of LZ's first live-data physics run.

3.4 Backgrounds

3.4.1 Cosmogenic backgrounds

Dual-phase Xe TPCs are capable of detecting cosmic muons. Their high interaction rate is unlikely to be mistaken for a WIMP, but neutrons produced in muon particle showers are a concern. To mitigate this LUX and LZ are located 1.5 km underground in the Davis cavern of the SURF. This facility provides the experiments with a rock overburden of 4300

Table 3.1 Summary of backgrounds in LZ, showing the number of counts expected in 1,000 live days in an indicative 5.6-tonne fiducial mass in the region of interest with all cuts applied. A comprehensive set of numbers can be found in [2]

Item	ER counts	NR counts
Detector Components	6.2	0.07
Dispersed radionuclides	911	-
Laboratory and cosmogenic	4.3	0.06
Fixed surface contamination	0.19	0.37
$^{136}\text{Xe } 2\nu\beta\beta$	67	-
Neutrinos ($\nu - e, \nu - A$)	255	0.72
Total	1244	1.22
Total with 99.5% ER discrimination, 50% NR efficiency	6.22	0.61
Total ER+NR background events	6.83	

m.w.e, reducing the muon flux by a factor of roughly 3×10^6 compared to the surface [2]. Despite this, ~ 300 muons are expected to reach the OD system every day. The OD systems of LUX and LZ are designed to handle this, with a veto system in the water tank set up to identify muons via their distinct Cherenkov radiation signal. With these measures in place, it is unlikely for a muon-induced neutron to make an energy deposition without also observing a muon, muon-induced cascade, or multiple scatter somewhere else in the detector. Taking these measures into account, the upper limit on the expected background rate from muon-induced neutrons in LZ for its projected 1000 day run is 0.06 [86] and contributes to the fourth row of Table 3.1 [2].

3.4.2 Intrinsic xenon backgrounds

Despite the many benefits of LXe as a target material, it contains several medium to long-lived radioisotopes which dissolve in the LXe target [2]. One particularly long-lived isotope is ^{85}Kr with a half-life ($T_{1/2}$) of 10.7 years, which has the potential to create an ER event anywhere in the target volume via beta decay. ^{85}Kr is removed via a column of activated charcoal which chromatographically separates it from Xe [87]. Various radon isotopes are also present in the LXe, such as: ^{219}Rn , ^{220}Rn , and ^{222}Rn ⁹. Of these ^{222}Rn is the most stable ($T_{1/2} = 3.82$ days). However, radon is not eliminated by underground storage since it will continue to emanate from detector surfaces, and be produced by ^{235}U , ^{238}U , and ^{232}Th decays. Being noble elements, Ra and Kr are difficult to distil from Xe. Extensive screening and purification regimes are conducted to mitigate these backgrounds. These backgrounds

⁹The short half-life of ^{219}Rn ($T_{1/2} = 4$ secs) means it is unlikely to diffuse from detector surfaces and emanate into the LXe.

contribute to row three in Table 3.1.

Cosmogenic activation of the LXe during above-ground transportation or storage can also produce electromagnetic backgrounds if not carefully managed. An example of this from previous experiments was tritium (^3H), as it is a long-lived isotope ($T_{1/2}=12.3$ years) and has a high production rate of $15 \text{ kg}^{-1}\text{day}^{-1}$ on the surface of Earth [88]. However, LUX and LZ are both able to purify their Xe to remove tritium during operation [2]. Other concerns include the Xe isotopes of ^{127}Xe , $^{129\text{m}}\text{Xe}$, and $^{131\text{m}}\text{Xe}$. Unlike most backgrounds which can be mitigated by the self-shielding discussed in Section 3.1.7, these isotopes will mix homogeneously with the target material. However, with half-lives of 36.4, 8.9, and 11.9 days respectively, these backgrounds fall to negligible levels soon after transportation underground.

3.4.3 Detector material backgrounds

Despite a significant radioassay effort in the selection of all detector components, several radioactive isotopes still reside in these materials. This includes ^{238}U , ^{232}Th , and ^{40}K . The largest contributors of these backgrounds come from the most massive detector volumes (e.g. the ICV), or volumes that have higher activity levels (e.g. the PMTs). Although these backgrounds are handled by fiducialisation (discussed in Section 3.1.7), understanding them is essential for the work conducted in this thesis.

One particular type of material background known as plate-out is in fact caused by the presence of radon in the air during detector component manufacturing and storage. Plate-out occurs when radon progeny such as ^{206}Pb , produced in the alpha decay of ^{210}Po , becomes embedded in the PTFE of the TPC cage housing the LXe. Due to some signal loss from partial energy depositions in the PTFE, the observed signal in the LXe from ^{206}Pb has a broad energy spectrum. This makes the reconstruction of energy depositions in XY space essential for the mitigation of this, and all other wall backgrounds. This includes electromagnetic backgrounds which can be created by cosmogenic activation in the detector materials. One example of this is ^{46}Sc , which is produced in the titanium cryostat and results in the gamma-ray emission of 1120 keV and 889 keV ($T_{1/2} = 84$ days). Copper, steel, and Kovar components throughout the detector can also contain cosmogenically activated ^{60}Co , producing gamma-rays of energy 1332 and 1173 keV ($T_{1/2} = 5.3$ years). These backgrounds contribute to the 2nd, 4th and, 5th rows of Table 3.1.

3.5 Background modelling

3.5.1 Overview

To correctly characterise backgrounds LUX and LZ developed an accurate background model based on Monte Carlo methods. Backgrounds from various sources were simulated using LUXSim, and its successor code BACCARAT [89]. These simulation packages were based on GEANT4, a standardised toolkit for the simulation of particles through matter [90]. The Noble Element Simulation Technique (NEST) was also implemented to determine detector response to energy depositions [91]. To help facilitate the parsing of information between BACCARAT and NEST, analysts had the option of using the LUX Legacy Analysis Monte Carlo Application (LLAMA) [92]. The analysis described throughout this thesis relies heavily upon this software. Therefore, it will be helpful to introduce, and briefly describe, each of these packages.

3.5.2 LUXSim and BACCARAT

LUXSim is a GEANT4 based framework primarily developed for the simulation of background events in the LUX detector [90]. Designed to provide a useful interface for low background experiments, LUXSim used a component-centric approach for simulations, facilitating the distribution of radioactive sources throughout detector volumes. This framework was used as a basis for the development of BACCARAT (Basically a Component-Centric Analog Response for AnyThing), which reduced the software's dependence on the LUX geometry to make the code applicable to other low background experiments. Aside from this minor difference both LUXSim and BACCARAT use GEANT4 for particle transport and to model the geometry of the detector and its surroundings.

These packages call upon a selection of custom generators, used to create various backgrounds in the LUX and LZ detectors. These generators facilitate simultaneous and time-delayed emission of primary particles for any source in any detector component. These particles go on to interact in the detector materials, producing resultant particles that also go on to create further interactions. These generators have been the subject of significant development since the creation of this software. Command-line arguments are used to define particle sources and the volume in which they will be distributed based on an input activity level. To illustrate this, one can look at the macro command used to begin a simulation in BACCARAT:

```
/Bacc/source/set A B C D
```

where:

A = geometry component (e.g. LiquidXenonTarget)

B = source type (e.g. SingleDecay_40_19)

C = activity (value)

D = activity (units)

3.5.3 NEST

The Nobel Element Simulation Technique (NEST) is an empirical model used to simulate the scintillation and ionization yields of recoiling particles in noble element detectors [91]. Calculations performed by NEST take into consideration dependence on both electric field strength and recoil energy. NEST is based on the calibration results of various liquid noble detectors and is intended to provide a standardised approach to scintillation and ionisation yields. NEST is built to be easily integrated into any GEANT4 based package, facilitating the prediction of detector responses such as *S1* and *S2* via standard MC techniques.

In addition to this, the NEST team has also developed several other sub-NEST packages such as *NEST_light*. *NEST_light* is an optical ray-tracer used by LUX. Although full MC optical processes are facilitated by NEST, for the work conducted in Chapter 6 this was too slow making *NEST_light* a valuable tool. The speed of *NEST_light* comes down to it not relying on full MC propagation of photon trajectories. Instead, it randomly chooses a photon's initial direction, and only updates the trajectory when reflecting or refracting.

3.5.4 LLAMA

The LUX Legacy Analysis Monte Carlo Application (LLAMA) is a software package designed to help interface NEST and BACCARAT to produce easily readable detector response signals. Although the Detector Electronic Response (DER) model which had already been built performs the same task, LLAMA has several advantages for analysts in certain circumstances. First, electric field models can be updated more easily. Second, it facilitates the use of several different clustering algorithms (e.g. DB scan, and agglomerative hierarchical). Development of LLAMA and the DER occurred in parallel (the former for LUX, and the latter for LZ). As such, LLAMA is the standardised approach to simulations in LUX and is an option for analysts in LZ.

3.6 Calibration for position reconstruction

Various calibration efforts are made by LUX and LZ to ensure the detector response to incident particles of known energy is understood. This includes using: D-D for NR response; ^{14}C and tritium for high and low energy ER response respectively; and $^{83\text{m}}\text{Kr}$ and $^{131\text{m}}\text{Xe}$ for the accurate position reconstruction of events inside LZ's TPC. Of all these sources $^{131\text{m}}\text{Xe}$ is of particular concern due to possible background events leaking into the NR band, as will be discussed in Chapter 4. As such, it is worthwhile to describe the rationale, strategy, and procedure for $^{131\text{m}}\text{Xe}$ position reconstruction.

Calibration for position reconstruction is performed to reject spatially varying external backgrounds and to precisely define the fiducial volume. More specifically this requirement demands an understanding of the TPC's response to events as position (XYZ), and time varies. It is a critical procedure since TPCs have several features that may impact the detector signal response to $S1$ and $S2$ production. Factors such as detector geometry, electric field strength, and light collection efficiency can cause variation in the XYZ response of $S1$, and $S2$. Additionally, variations in Xe purity can affect the $S2$ drift time due to the attachment of ionised electrons to these impurities. If these variations are not understood, further analysis will be troubled by uncertainty over the origin of particular events. This could see an ER event mis-reconstructed in the NR band, wall events reconstructed within the fiducial region, or even an external particle incident inside the fiducial region (potentially a WIMP) incorrectly observed to be outside of the fiducial volume.

This calibration strategy calls for a homogeneously distributed source to be deployed throughout the active LXe. While LUX with its smaller geometry was able to rely on $^{83\text{m}}\text{Kr}$ with a half-life of 1.83 hours¹⁰, the larger geometry of LZ may not allow $^{83\text{m}}\text{Kr}$ to fully homogenise given the longer mixing timescale of the LXe in LZ (2.5 days) [93, 2]. As such, LZ will use $^{131\text{m}}\text{Xe}$ as a longer lived alternative, specifically for achieving the goal of homogeneity. By comparison $^{131\text{m}}\text{Xe}$ decays at a higher energy¹¹ (163.93 keV), and has a longer half-life (11.9 days), ensuring homogeneity for the purposes of calibrating the position reconstruction.

The hardware for calibration in LZ will involve Xe carrier gas flowing over a parent isotope (^{131}I) that is emanating $^{131\text{m}}\text{Xe}$ [2]. ^{131}I will completely decay into ^{131}Xe with a half-life of

¹⁰ $^{83\text{m}}\text{Kr}$ decays produce two lines, 32.2 keV and 9.4 keV separated by 157 ns. As such, the 2 $S1$ signals are sometimes seen separately.

¹¹Low energy calibration is performed with tritium. But this cannot be performed in situ. Thus periodic tritium calibrations are performed, and the evolution of detector response over time is measured with $^{131\text{m}}\text{Xe}$.

8 days. However not all decays go directly to the ground state, with 0.39% first decaying to the metastable isotope ^{131m}Xe . ^{131}I is a readily available source in the medical industry, coming in a variety of forms. The most convenient form for LZ's purposes is in a solid pill.

While it may not seem ideal to introduce radioisotopes into the detectors sensitive region, this technique was previously validated by the LUX experiment [1]. The use of external radioisotope sources may seem a more cautious approach; however, it is entirely prevented by the strong self-shielding of LXe, an extremely powerful feature used for removing external ER backgrounds from the sensitive region. The 163.93 keV energy of ^{131m}Xe decays is much larger than typically expected DM signals, ensuring it is easily excluded from any direct detection analysis. Furthermore, the mono-energetic nature of this decay ensures it occupies an easily identifiable region in $S2$ vs. $S1$ space as seen in Figure 3.7. Here the ^{131m}Xe energy deposition is clearly seen clustered between $200 \text{ phd} < S1 < 1000 \text{ phd}$, quite separated from our WIMP search region of interest (ROI) of $2 \text{ phd} < S1 < 120 \text{ phd}$. In this mock data the following parameters were used, $g_1 = 0.128$, $g_2 = 92.26$ (a higher value of g_2 compared to past experiments [1]). ER discrimination was not applied to this dataset. Subsequently, a suite of background events which populate the ER band can be seen at low $S1$ ($S1 < 300$) and low $S2$ ($S2 < 200000$)¹².

LZs current ^{131m}Xe deployment plan consists of an initial 35 mBq injection, followed by weekly 23 mBq injections to maintain an activity > 12 mBq in the TPC (> 1000 events/day). However, as shall be discussed in Chapter 4, before ^{131m}Xe can be used as a calibration source its decay and performance in the detector, must be well understood.

3.7 Data analysis cuts

When using low background detectors the identification and removal of known backgrounds is a critical step in establishing the necessary environment for WIMP detection. Given the similarities between LUX and LZ, there is a strong overlap in the cuts they employ when performing any kind of data analysis. In this section the most common of these cuts will be described, facilitating the discussions in later chapters over specific analyses related to LUX and LZ data.

¹²Incidentally, this figure also provides a demonstration of energy reconstruction as discussed in Section 3.1.5, with the ^{131m}Xe energy being reconstructed via the blue line using Eq. 3.11.

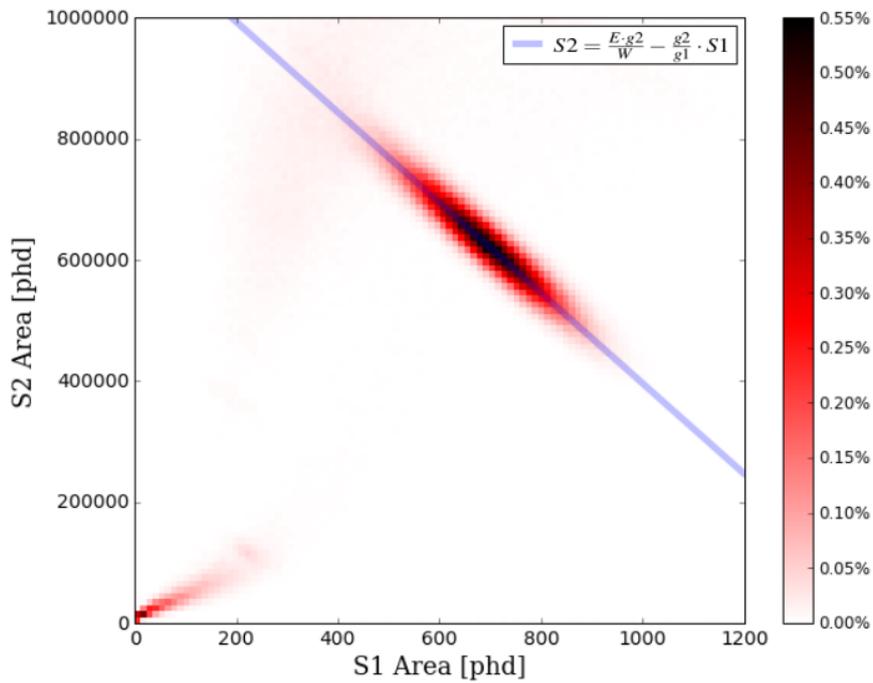


Figure 3.7 The red-black heat map shows where simulated mock data taken from LZ’s MDC2 are proportionally located in $S2$ vs. $S1$ space. The blue line represents the reconstructed energy of the calibration source $^{131\text{m}}\text{Xe}$ ($E = 163.93$ keV), using Eq. 3.11 with: $g_1 = 0.128$, $g_2 = 92.26$ (a higher value of g_2 compared to past experiments [1]). Also visible in this plot are a suite of other background events with low $S1$ (as seen on the lower left corner).

3.7.1 Region of Interest

The Region of Interest (ROI) cut is designed to remove events that deposit energy outside the expected range for a particular DM search. The definition of this region depends on both the experiment (i.e. LUX or LZ), the search type (i.e. standard WIMP search, or EFT), and the expected mass of the DM particle.

For a standard WIMP search, the ROI will depend on the WIMP mass. The ROI energy range for LUX is commonly quoted as $3.4 - 25 \text{ keV}_{nr}$, where “nr” denotes that the energy was deposited by a nuclear recoil [94]. This range was chosen because it captures 80% of all WIMP interactions for a 100 GeV WIMP at standard galactic DM halo parameters as described in Chapter 2. This NR energy range has an equivalent ER energy range of $0.9 - 5.3 \text{ keV}_{ee}$, where “ee” denotes an energy calibration for ER events. The energy window for ER events differs from that of NR events due to differences in scintillation and ionisation yield between these two types of events as discussed in Section 3.1.

For LZ this ROI is commonly quoted as $6.0 - 30 \text{ keV}_{nr}$, and $1.5 - 6.0 \text{ keV}_{ee}$ [2].

In LUX, the EFT searches operate with a broader energy range, extending up to 70 keV_{ee} [95]. For LZ a similar or extended range will be used.

3.7.2 Multiple scatter cut

Given current limits on the probability of a direct interaction of DM discussed in Section 2.4, it can be safely assumed that if DM were to be directly detected it would be via a single scatter rather than a multiple scatter event. This premise underpins a background cut designed to remove gamma-rays and neutrons which typically scatter multiple times in the detector. The multiple scatter cut relies on 3D position reconstruction to accurately identify scattering vertices in the active LXe region. If $S2$ vertices cannot be resolved, the shape of the $S2$ can also be used to determine if the event is a multiple scatter. Thus, the cut itself simply removes any event which appears to scatter more than once from consideration as a WIMP candidate.

3.7.3 Fiducial volume cut

The fiducial cut is also intended for the removal of gamma and neutron backgrounds originating from the TPC itself. This cut requires the exclusion of events within a few centimetres of the detector boundaries, the precise dimensions of the excluded volume is chosen to

maximise the signal to background ratio of WIMP candidate events.

In LZ the radial fiducial cut excludes events that occur up to 4 cm from the wall. The vertical fiducial cut excludes events up to 1.5 cm from the cathode (at the bottom of the TPC), or 13.5 cm from the gate¹³ (at the top of the TPC). With these exclusions in place, LZ is provided with an ultra-quiet 5.6-tonne LXe fiducial target, having reduced neutron and gamma backgrounds by a factor of 10. However, the precise size of the fiducial volume is determined by the accuracy of event position reconstruction, with NRs from the decay of radon progeny occurring close to PTFE surfaces [2].

When implementing these cuts for LUX, they need to be slightly altered to account for the differing geometry. The radial fiducial cut in LUX excludes events up to 3 cm from the wall. The vertical fiducial cut is defined in terms of electron drift time - that is the time delay between the $S1$ and $S2$. For this LUX excludes events with a drift time¹⁴ greater than 300 μs (~ 5 cm above the cathode), or less than 40 μs (~ 5 cm from the phase barrier) [94].

3.7.4 LXe skin cut

As an additional layer of defence against gamma and neutron backgrounds, LZ utilises a LXe skin region. The skin itself is a separate layer of LXe surrounding the active region, with dedicated PMT arrays recording signal events. This cut removes events in the fiducial region which have a coincidence in the skin region. In doing so the fiducial volume can be increased from 5.1 tonnes to 5.6 tonnes without a significant increase in background signal rates [2].

The use of a LXe skin is a new development for LZ and was not implemented in LUX.

3.7.5 Outer detector cut

Another cut introduced for LZ which was not used by LUX was the OD cut. Utilising the OD veto system described in Section 3.3.2, this cut removes fiducial events that occur in conjunction with an OD signal.

¹³A negatively charged grid used to prevent the electric field from having an effect on the PMTs.

¹⁴The average drift speed of electrons in LUX is ~ 1.5 mm/ μs .

Chapter 4

$^{131\text{m}}\text{Xe}$ simulation studies for LZ

4.1 $^{131\text{m}}\text{Xe}$ background production mechanisms

$^{131\text{m}}\text{Xe}$ is used in LZ to calibrate position reconstruction. Its characteristic decay energy of 163.93 keV makes it readily identifiable, and its half-life of 11.9 days ensures it will mix homogeneously with the target Xe before completely decaying. The $^{131\text{m}}\text{Xe}$ decay scheme is complex, with only 1.95% of decays resulting in a single gamma-ray, and the remaining decay modes involving multiple decay products. Such a decay scheme has the potential to create a background signal via a couple of mechanisms. The first mechanism is called a Partial Energy Deposition (PED) and involves some of these products escaping the sensitive LXe volume without depositing all their energy. The second mechanism is called a Multiple Scatter Single Ionisation event (MSSI) and involves some decay products depositing their energy in a region of the LXe without detectable ionisation signal¹. In each case, the total energy of the $^{131\text{m}}\text{Xe}$ decay is not changed, but the reconstructed energy is. It is plausible that over the course of LZ's planned 1000 day run some of these decays will lead to a false WIMP signal.

In order to determine the likelihood of such backgrounds occurring, four sequential steps were taken. First, a $^{131\text{m}}\text{Xe}$ generator was written in BACCARAT, which included all possible decay modes. This generator had to be built to account for GEANT4's inability at the time to correctly produce all decay products of $^{131\text{m}}\text{Xe}$. Second, this code was used to conduct a full chain Monte Carlo (MC) simulation of $^{131\text{m}}\text{Xe}$ in LZ. Third, these simulated events were analysed, specifically looking for PEDs at the MC truth level. Finally, these same simulations were examined for MSSI events in a similar fashion.

¹These same events go by the name gamma-X within the LUX collaboration.

In this chapter, the ^{131m}Xe simulations and studies will be discussed. This chapter will consist of five parts. First, a description of ^{131m}Xe , and the physics governing its decay scheme. Second, a discussion of the BACCARAT generator, and its validation. Third, a summary of the ^{131m}Xe simulations. Fourth, the results from the PED analysis. Finally, a description of the MSSI analysis will be given.

4.2 Metastable ^{131}Xe

^{131m}Xe is a metastable state of ^{131}Xe caused by the excitation of the nucleus. The term “metastable” refers to nuclear isomers with half-lives greater than 5×10^{-9} seconds, ^{131m}Xe falls under this category with a half-life of 11.9 days [96]. The long half-life of metastable states is due to the large nuclear spin difference between the metastable state and the ground energy state. This large spin change causes suppression of a single gamma-ray emission decay mode. For ^{131m}Xe , the metastable state has a $-11/2$ spin which must decay to a ground state of $+3/2$.

4.3 ^{131m}Xe emissions

The gamma decay mode, with an energy of 163.93 keV, has a branching ratio of only 1.95%. The internal conversion process occupies the remaining probability, resulting in a vacancy in one of the atomic sub-shells [98]. The internal conversion process of ^{131m}Xe is characterised by 3 emission types: internal conversion (IC) electrons, Auger electrons, and characteristic X-rays.

IC electrons arise due to the chance of any inner shell electron having its wave-function overlap and penetrate into the atomic nucleus [99]. This allows that electron to couple with an excited nuclear energy state, and directly take the nuclear transition energy without an initial gamma-ray emission. The emitted electron will have a kinetic energy equal to the nuclear transition energy, minus the electron’s binding energy to the atom. These IC electrons typically originate in the K-shell ($1s$) since these electrons are the most likely to be found in the nucleus. This does not exclude s state electrons from the L, M, or N shells ($2s$, $3s$, or $4s$) from being ejected via the IC process. When this happens they are called L, M, or N IC electrons. In some cases p electrons (from shells L and higher) can also be ejected by the IC process; however, this is much less likely due to the lower probability of the wave-function intersecting the nucleus.

Table 4.1 Emission spectra for ^{131m}Xe as it decays to ^{131}Xe [97]. The various decay modes are formed by combinations of conversion electrons, Auger electrons, and X-rays. The probability of each mode along with the electron and photon energies are given here.

Decay Mode No.	Decay Probability	Electron Energies (keV)	Photon Energies (keV)
1	0.0195	-	163.93
2	0.0529	129.37, 0.931	33.629
3	0.0297	129.37, 0.998	33.562
4	0.0169	129.37, 0.158	34.402
5	0.0445	129.37, 0.828	29.452, 4.28
6	0.1125	129.37, 3.43, 1.678	29.452
7	0.29153	129.37, 3.43, 1.349	29.781
8	0.00042	129.37, 24.5, 1.5	4.28, 4.28
9	0.0179	129.37, 24.5, 3.43	4.28
10	0.0196	129.37, 29.2, 1.08	4.28
11	0.02588	129.37, 24.5, 3.43, 4.09, 2.54	-
12	0.0021	129.37, 33.4, 1.16	-
13	0.14112	158.79, 4.09, 1.05	-
14	0.0061	158.79, 4.84	0.3
15	0.0061	158.79, 4.84, 0.3	-
16	0.13268	158.79, 3.43, 1.71	-
17	0.0657	162.89, 1.04	-
18	0.0134	163.74, 0.19	-
19	0.00147	163.91, 0.02	-

Auger electrons can appear after an IC electron emission when a vacancy has been left on an inner electron shell. This may lead to an electron from a higher energy level falling into the vacancy, resulting in the atom releasing a quantum of energy. When this quantum takes the form of an electron it is called an Auger electron. The kinetic energy of the Auger electron is the difference between the energy of the initial electron's transition to the vacancy and the ionisation energy of the electron shell from where the Auger electron was ejected [100].

An alternative to an Auger electron is a characteristic X-ray. Here the ejected energy quantum is not in the form of an electron, but rather a photon. The characteristic X-ray derives its name from the fact that each element has a unique energy emission signature. The emission spectra for ^{131m}Xe are given in Table 4.1. This table includes both photon and electron energies, with each decay mode given a unique decay mode number for the purpose

of distinction. To illustrate a typical process, decay mode number 11 is shown in detail in Figure 4.1 [97]. Notably, $^{131\text{m}}\text{Xe}$ has many more decay products than are listed in Table 4.1. The unlisted products consisted of X-rays, and Auger electrons with energy < 1 keV. In these instances, decay products were appropriately combined together into either a single electron or a single photon with a summed energy.

The energy difference between the metastable state and the ground state of $^{131\text{m}}\text{Xe}$ is 163.93 keV, which is thus the expected de-excitation energy of all events in the $^{131\text{m}}\text{Xe}$ BACCARAT generator. 1.95% of the time this decay takes the form of a single gamma-ray. The remaining cases see a combination of conversion electrons, Auger electrons, and X-rays from atomic relaxation, as illustrated in Figure 4.2. This Figure shows the photon and electron production in $^{131\text{m}}\text{Xe}$ decays, with 5 different electron peaks labelled A-E, and 5 photon peaks labelled E-A. Peaks with the same index occur in coincidence, and sum to 163.93 keV. The decay scheme included in BACCARAT has been taken from [97]. Each individual electron and photon peak is built from several quantised decay products which can be seen in Table 4.1.

4.4 $^{131\text{m}}\text{Xe}$ Monte Carlo Generator

4.4.1 Primary particle energy validation

A $^{131\text{m}}\text{Xe}$ generator must be constructed in BACCARAT since at the time of simulation native GEANT4 sources did not properly produce the complex decay schemes with Auger electrons, and X-rays which are described in Section 4.2. Once constructed, the generator had to be validated to assess three criteria: that primary particle emission matches Figure 4.2, that all $^{131\text{m}}\text{Xe}$ decays emit 163.93 keV, and that $^{131\text{m}}\text{Xe}$ decays are randomly distributed in all LXe volumes.

The construction of the $^{131\text{m}}\text{Xe}$ generator in BACCARAT began with a careful examination of the allowed decay processes, shown in Figure 4.2. These quantised decay modes were built into the $^{131\text{m}}\text{Xe}$ generator by sampling the probabilities from Table 4.1. Once built, the energies of the primary decay products were compared to the calculated values given by Ringbom [97], which can be seen in the margins of Figure 4.3. This figure distinguishes between the electron and the photon decay products.

The margins of Figure 4.3 shows a comparison of Ringbom's calculations (red crosses) [97], and the BACCARAT generator, built based on Ringbom's empirical descriptions of $^{131\text{m}}\text{Xe}$

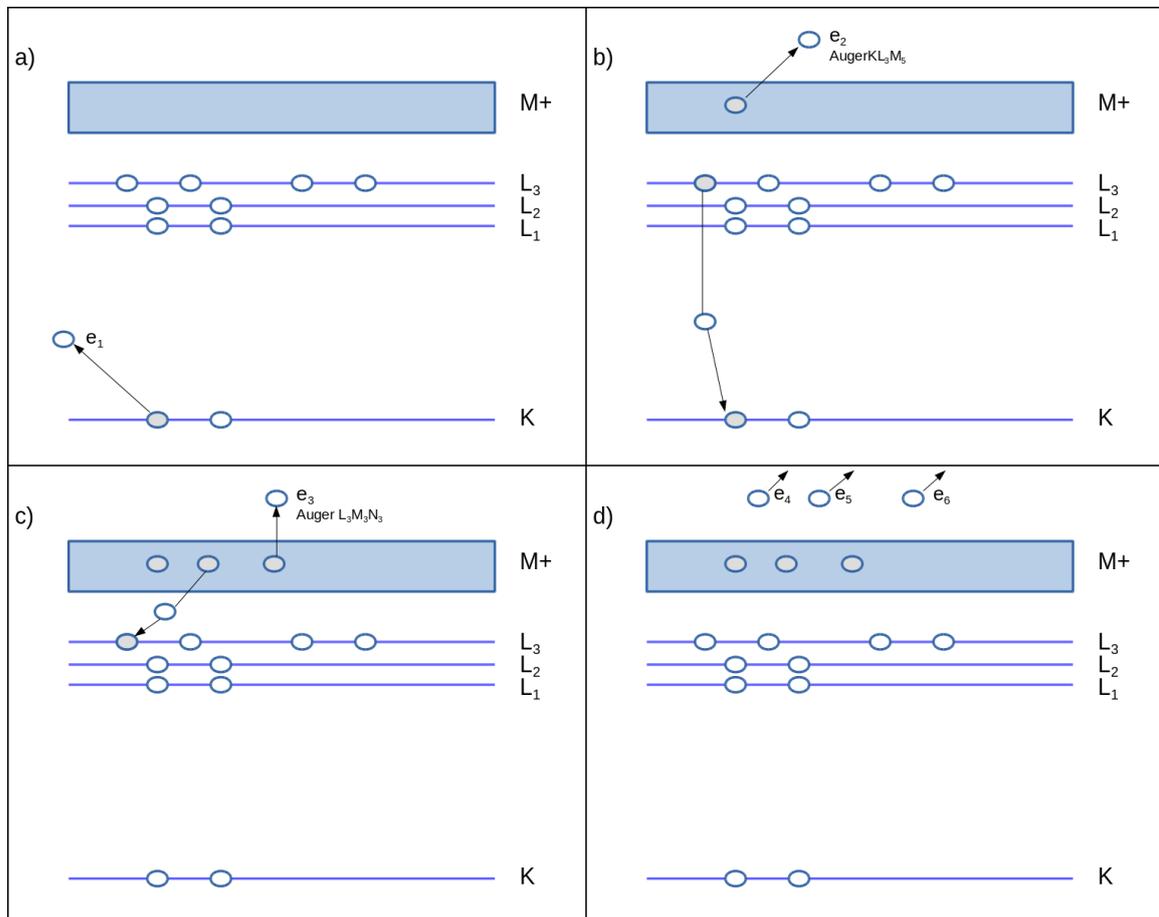


Figure 4.1 A decay mode for ^{131m}Xe taken from [97]. To illustrate the process, electrons are drawn as white circles and vacancies in grey. Step (a) begins with a K-conversion electron emission, followed by an electron filling this K shell vacancy and corresponding Auger electron emission in step (b), followed by a similar transition in step (c). The energies of the electrons emitted in the last step (d), are equal to the binding energies of each vacancy in the M+ shell (where the + indicates a valence band) and are emitted to ensure the full decay comes matches the excitation energy of ^{131m}Xe . For this particular decay path, where only electrons are emitted, the total electron energy equals the excitation energy of ^{131m}Xe . All parts of this decay occur instantaneously and are only shown in four stages for the purpose of clarity.

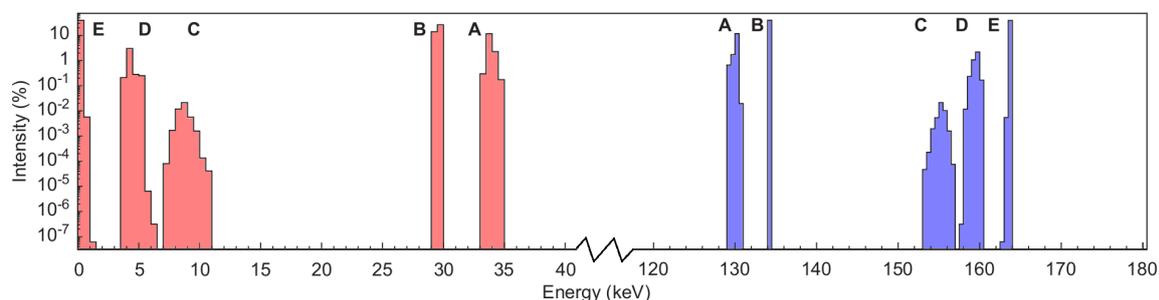


Figure 4.2 Spectral output $^{131\text{m}}\text{Xe}$. In red are the photon peaks (excluding the 163.93 keV peak), and in blue are the electron peaks. Intensity is given on a logarithmic scale [97].

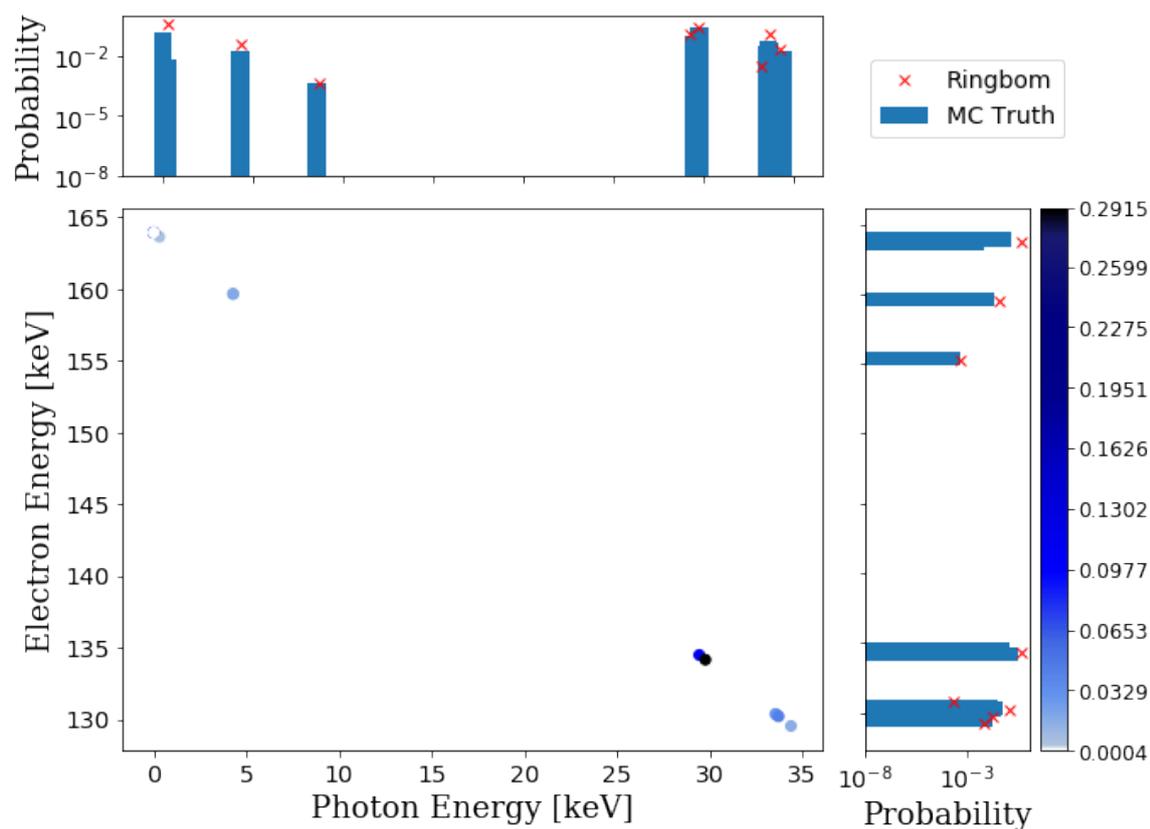


Figure 4.3 Generator validation plot. In the centre, the figure shows the coincidence between the simulated primary photon and electron particle energies. For each case, the sum comes to 163.93 keV. In the margin simulated electron and photon peaks (blue bars) are compared with the same peak values quoted in [97] (red crosses) for $^{131\text{m}}\text{Xe}$ decays.

decays (blue bars). Inspection of this shows the results are in close agreement with each other. Where small discrepancies appear, namely an excess of ~ 30 keV gamma-rays in the blue bars at the expense of a host of lower energy photons, appears to arise as a result of Ringbom's calculations not matching his empirical description. While this may be a result of an undescribed decay mode, incorporating this into the generator without an empirical justification would not be prudent.

4.4.2 Primary particle spatial distribution

It was also necessary to ensure the spatial distribution of $^{131\text{m}}\text{Xe}$ events was random within a specified volume, to mimic the homogeneous distribution of these decays in LZ. Figure 4.4 shows the original position of 10^5 simulated events, called in all LXe volumes. This figure showed that the generated events covered all LXe volumes when called to do so. This is made explicit by highlighting the separation of the skin volume (outlined in green) from the sensitive LXe volume (outlined in red), while leaving the non LXe volumes blank.

4.5 $^{131\text{m}}\text{Xe}$ simulation

Once the $^{131\text{m}}\text{Xe}$ generator was validated, BACCARAT was used to simulate 6.12×10^9 $^{131\text{m}}\text{Xe}$ decays in all LXe volumes of the TPC. The simulations were conducted in late 2016, which used default values for g_1 and g_2 of 0.075, and 59.2 at the time. The simulation also assumed a uniform vertical electric field in the FFR of 290 V/cm.

This simulation represented 3.5×10^6 days of data at a presumed 0.02 Hz average decay rate²; however, this decay rate is still subject to change by the LZ collaboration. Given LZ's current deployment plan, assuming a 0.02 Hz average $^{131\text{m}}\text{Xe}$ decay rate is rather conservative, and is expected to be much lower [101]. However, this is the preferred option when estimating the background caused by $^{131\text{m}}\text{Xe}$.

Once raw energy depositions were obtained, they were clustered into spherical bins of radius 0.4 mm and processed with NEST to determine $S1$ and $S2$ yields. This facilitated the search for WIMP-like signals after applying all data analysis cuts. Since no WIMP events were present in the original simulation, any WIMP identification would represent a serious concern for the use of $^{131\text{m}}\text{Xe}$ as a calibration source. $^{131\text{m}}\text{Xe}$ has two mechanisms by which it can potentially create a false WIMP signal: PEDs and MSSIs. Separate analyses were

² $^{131\text{m}}\text{Xe}$ will be injected periodically so as to have a 0.02 Hz decay rate on average.

conducted on these simulations to determine the impact of these pathologies on LZ. These two studies will now be discussed in the following sections.

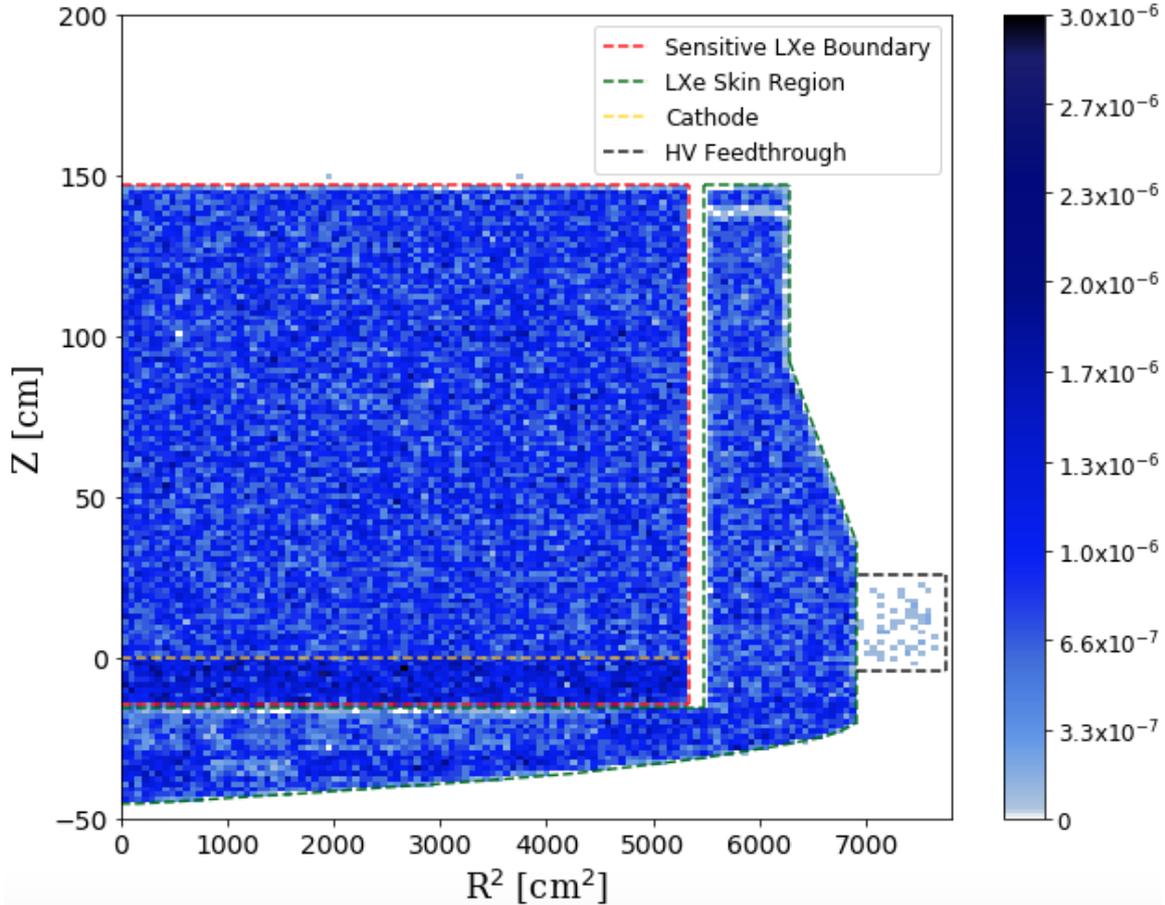


Figure 4.4 Simulated primary particle initial positions for 10^5 events. Particles are randomly distributed throughout all the LXe volumes which the generator was called in. Volume encircled in red represents the sensitive LXe target. In green is the skin. The cathode position is represented with the yellow line at $Z = 0$. A small contribution can also be seen coming from the LXe in the high voltage feed through.

4.6 Partial Energy Deposition Analysis

A PED occurs when one (or more) $^{131\text{m}}\text{Xe}$ decay product escapes the LXe without depositing all of their energy. As a result, the reconstructed energy is less than the characteristic 163.93 keV; meaning, the event will not be identified as a $^{131\text{m}}\text{Xe}$ decay. The worst-case scenario for the standard WIMP search occurs if a 163.93 keV emitted gamma-ray escapes detection after a fiducialised Compton scatter in the ROI. While the escape of other individual

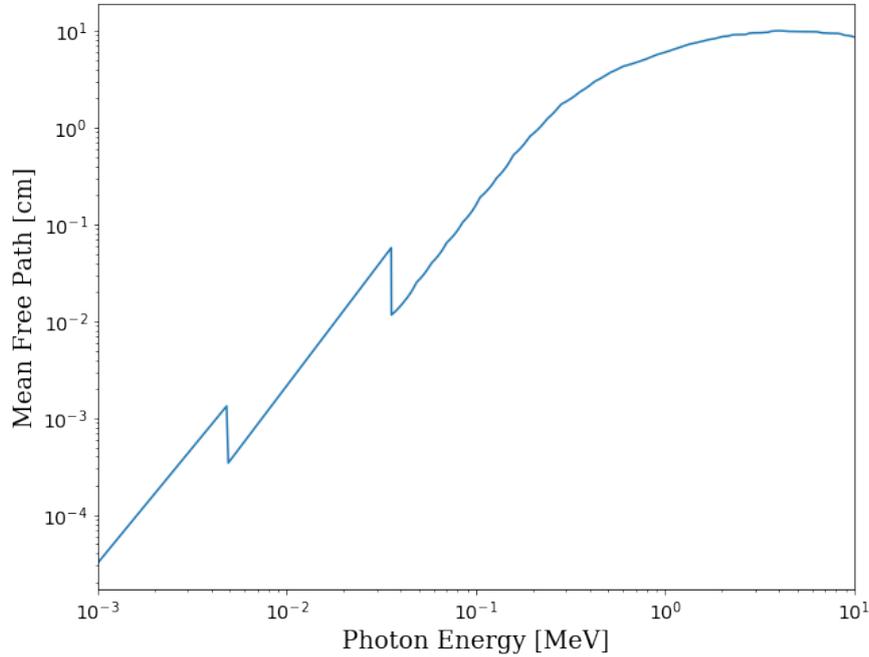


Figure 4.5 Photon mean free path in LXe of density 2.88 g/cm^3 .

decay products from Table 4.1 would not lead to a signal in the WIMP search ROI, they could be a problem for other high energy searches, such as the EFT DM search. Of these remaining de-excitation modes, the largest concern is for decays containing $\sim 30 \text{ keV}$ X-rays. This is because after the 163.93 keV gamma-ray, these X-rays have the greatest chance of leaving a PED given their Mean Free Path (MFP) of $\sim 0.35 \text{ mm}$. This is illustrated in Figure 4.5, which shows how photon energy affects its MFP in LXe. Figure 4.5 takes empirical photon attenuation data from [102], and uses the formula:

$$\ell = (\mu \cdot \rho)^{-1} \quad (4.1)$$

where ℓ is MFP (cm), μ is the attenuation coefficient (cm^2/g), and ρ is the LXe density (2.88 g/cm^3).

This informs the hypothesis that the vast majority of PEDs ought to occur near the edge of the TPC. As such, and given the expected accuracy of LZ's position reconstruction, this hypothesis assumes PEDs will be excluded by fiducialisation. In addition to this, there is a small chance that some of these PEDs might also be removed by the MS, OD, or skin cuts. The rationale being that escaped decay products could leave energy deposits in locations that trigger these cuts. These additional cuts are only likely to affect a 163.93 keV gamma-ray

which happens to deposit a small amount of energy in the LXe. Lower energy photons are unlikely to pass through the PTFE or have a MFP long enough to be registered as a MS.

This analysis aims to test these hypotheses and ensure that this calibration source is not introducing a new background. In doing so the number of $^{131\text{m}}\text{Xe}$ PED events which LZ would observe over 1000 days, for both the traditional WIMP search and the EFT DM search will be determined. With this value, the number of PEDs that lead to a false DM signal will be calculated.

To find the position of these PED events, and to estimate the fraction of PEDs leading to false WIMP signals, the simulated $^{131\text{m}}\text{Xe}$ decays from Section 4.5 were analysed for PED events. This involved looking for decays that did not deposit all of their energy in the LXe. As indicated in Table 4.2, of the 6.12×10^9 simulated decays mentioned in Section 4.5, only 357442 ± 598 events fit this description of a PED. As shown in Figure 4.6, the spatial distribution matched expectations with all these events occurring near the edge of the TPC. Specifically, most PEDs can be seen to be taking place by the wall, while a subdominant fraction occurs by the phase boundary.

Table 4.2 highlights the effectiveness of various data analysis cuts for both the standard WIMP search and the EFT DM search. In both searches, the ROI cut mentioned in Section 3.7.1 (1.5 - 6.0 keV for the WIMP search, and 1.5 - 70 keV for the EFT search used by LUX) reduce the number of PED events by an order of magnitude. However, this alone is not effective at eliminating the entire background. Of the remaining data analysis cuts, the skin cut can be seen to be capable of rejecting a few more of these events. However, it should also be noted that the presence of $^{131\text{m}}\text{Xe}$ in the skin will also contribute to the overall PED background. The MS cut can also be seen to be contributing to the removal of PED events, albeit a rather inconsequential one. However, this is slightly better than the OD cut, which does not help at removing backgrounds such as $^{131\text{m}}\text{Xe}$ PEDs, since their decay products cannot pass through the OCV to make energy depositions in the OD.

The most effective cut from Table 4.2 is clearly the fiducial cut, which reduces the PED background to zero for both the EFT, and standard WIMP search. In fact, regardless of ROI, or any other data analysis cut, fiducialisation alone was successful at removing all PED events. This result is to be expected, since the most penetrating decay products, the 163.93 keV gamma-ray, has $\text{MFP} = 5.5$ mm, and the ~ 30 keV X-ray has $\text{MFP} = 0.35$ mm. As such neither of these are likely to escape the 4 cm of radial length from the wall excluded by the fiducial cut.

Table 4.2 Result from PED background analysis with various data analysis cuts applied. The cut definition can be found in Section 3.7. The first column indicates which cuts have been applied for each of the WIMP and EFT searches. The second column shows how many PED events remain after each cut. The final column normalises these values to a 1000 day run assuming an average 0.02 Hz decay rate. WIMP search ROI is 1.5 keV to 6.0 keV. The EFT ROI extends up to 70 keV.

	<i>Cuts</i>	<i>Raw simulation</i>	<i>1000 day normalisation (.02 Hz mean rate)</i>
	Total (no cuts)	357442 ± 598	100.7 ± 0.2
<i>WIMP Search</i>	ROI	18004 ± 134	5.07 ± 0.04
	ROI+MS	17975 ± 134	5.06 ± 0.04
	ROI+MS+OD	17975 ± 134	5.1 ± 0.04
	ROI+MS+OD+Skin	14981 ± 122	4.2 ± 0.03
	ROI+MS+OD+Skin+Fid	0 (<2.44)	0.0 (< 6.88 × 10⁻⁴)
<i>EFT Search</i>	ROI	34162 ± 185	9.62 ± 0.05
	ROI+MS	34107 ± 185	9.61 ± 0.05
	ROI+MS+OD	34107 ± 185	9.61 ± 0.05
	ROI+MS+OD+Skin	28397 ± 169	8.00 ± 0.05
	ROI+MS+OD+Skin+Fid	0 (<2.44)	0.0 (< 6.88 × 10⁻⁴)

The effectiveness of the fiducial cut is clear from Figure 4.6, which shows the location of all PED events before the application of data analysis cuts. In the central plot, the position of all PEDs is shown in R - Z space, with the position of the LXe boundary, and fiducial volume indicated. In the horizontal margin is the abundance of PEDs with respect to radial position, and in the vertical margin is the distribution with respect to LXe depth. The plots in the margins have had their axes reduced to focus on the distributions and their distance to the fiducial boundary. It can be seen that PEDs are much closer to the radial fiducial cut than they are to the vertical fiducial cut. The reason for this is due to increased statistics, with the TPC wall having a much larger surface area than the phase boundary.

Since fiducialisation alone removes all simulated PEDs from analysis, both the EFT and standard WIMP searches have a background PED contribution of zero. However, for conservatism, an expected rate of zero will not be used. Instead, an upper confidence limit based on a null result is derived by using the original confidence belt definition construction by Neyman [103]. This gives an upper expectation value of 2.44 events with 90% confidence. When normalised to 1000 days, this reduces to 6.88×10^{-4} PED events, assuming a mean $^{131\text{m}}\text{Xe}$ decay rate of 0.02 Hz. This rate may vary later in the experiment however. As such it is useful to express this upper PED limit as a function of a decay rate. This yields the equation:

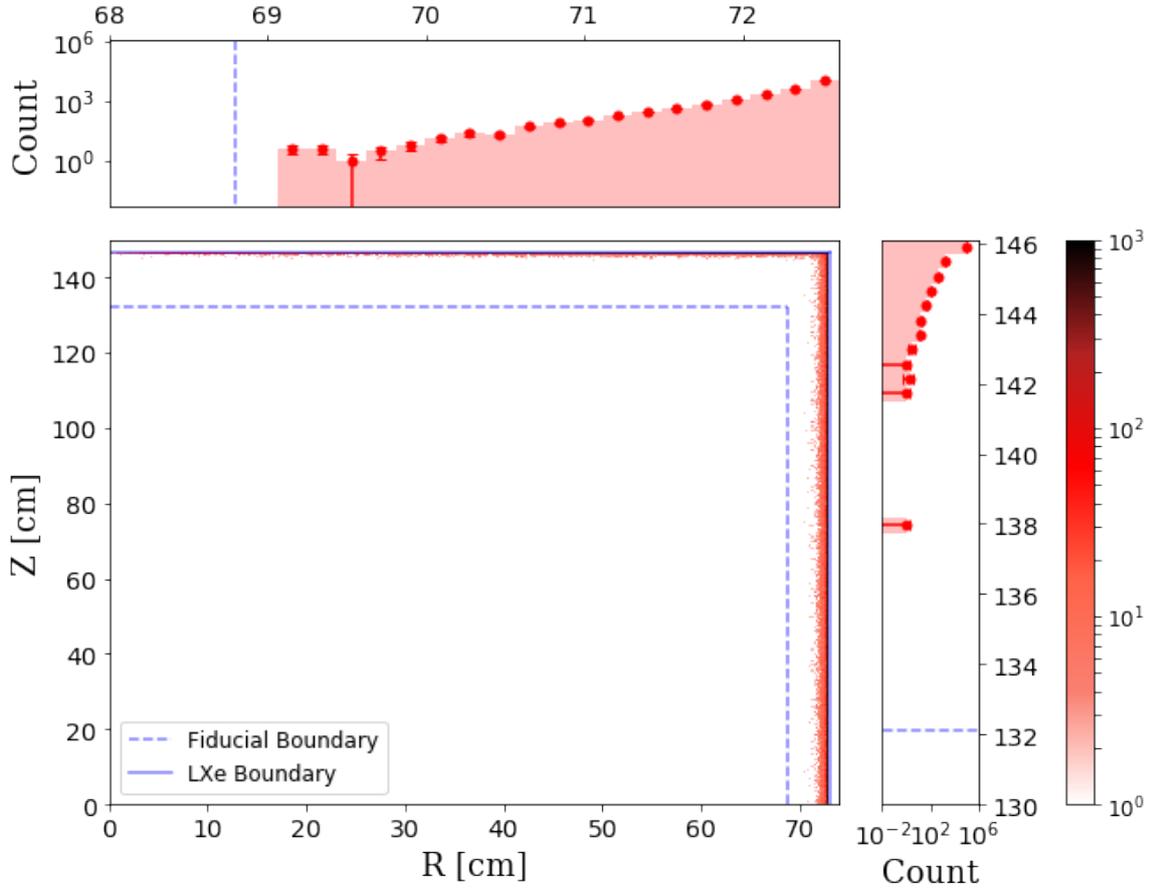


Figure 4.6 Position of all simulated $^{131\text{m}}\text{Xe}$ partial energy depositions. All events are removed by the fiducial cut. The central plot shows all events in R - Z space. The plot in the horizontal margin shows a close up of events near the wall. The plot in the vertical margin shows a close up of events near the phase boundary.

$$N_{PED} = 0.034 \times R_{\mu} \quad (4.2)$$

where N_{PED} is the 1000 day normalised expectation of $^{131\text{m}}\text{Xe}$ PED events at 90% confidence, and R_{μ} is the mean decay rate of $^{131\text{m}}\text{Xe}$ in Hz.

4.7 Multiple Scatter Single Ionisation Analysis

Another potential background from $^{131\text{m}}\text{Xe}$ decays are MSSI events (a.k.a. gamma-X events). These MSSI events have the potential to alter the observed $S2/S1$ ratio, causing an ER event to fall in the NR band. $^{131\text{m}}\text{Xe}$ is thought to be capable of producing an MSSI event when some of its decay products deposit energy in the LXe above the cathode, and some deposit

below the cathode. Such a MS event would yield an $S1$ and $S2$ for the energy deposition above the cathode. But below the cathode, where the direction of the electric field is reversed, only an $S1$ signal will be produced. As such, energy depositions below the cathode will only be registered with an $S1$ signal. For a nominal MSSSI event, where there is only one scatter above (E_{UP}), and one below the cathode (E_{LO}), three signals will be produced: $S1_{LO}$, $S1_{UP}$, and $S2_{UP}$. This negates the MS cut since the two $S1$ s will merge into a single observed signal, and only one $S2$ will be seen³. Problematically, the enhanced $S1$ which is observed will alter the $S2/S1$ ratio. Such a lowered $S2/S1$ ratio could potentially shift an event into the NR band.

The same simulated ^{131m}Xe decays from Section 4.6 were used to search for MSSSI events. This was done to determine the number of MSSSI events which LZ would observe over 1000 days, for both the traditional WIMP search and the EFT WIMP search. An event was identified as an MSSSI if it had at least one $S1$ producing energy deposition in the LXe beneath the cathode, and one $S1$ and $S2$ producing energy deposition in the LXe above the cathode based on MC truth.

Table 4.3 Result from MSSSI background analysis with various data analysis cuts applied.

	<i>Cuts</i>	<i>Raw simulation</i>	<i>1000 day normalisation (.02 Hz mean rate)</i>
	Total (no cuts)	27965 ± 167	7.88 ± 0.05
<i>WIMP Search</i>	ROI	19 ± 4	0.005 ± 0.001
	ROI+MS	19 ± 4	0.005 ± 0.001
	ROI+MS+OD	19 ± 4	0.005 ± 0.001
	ROI+MS+OD+Skin	19 ± 4	0.01 ± 0.001
	ROI+MS+OD+Skin+Fid	0 (<2.44)	0.00 (< 6.88 × 10⁻⁴)
<i>EFT Search</i>	ROI	6539 ± 81	1.84 ± 0.02
	ROI+MS	6529 ± 81	1.84 ± 0.02
	ROI+MS+OD	6529 ± 81	1.84 ± 0.02
	ROI+MS+OD+Skin	6529 ± 81	1.84 ± 0.02
	ROI+MS+OD+Skin+Fid	93 ± 10	0.026 ± 0.003
	ROI+MS+OD+Skin+Fid+NR	29 ± 5	0.008 ± 0.002

The result of this analysis for both the EFT and standard WIMP search can be seen in Table 4.3, which shows the production of 27965 ± 167 MSSSI events from the simulation described in Section 4.5. The same data analysis cuts from the PED analysis were applied to

³There is some capacity to identify these events based on their $S1$ -photo-hit-pattern in the bottom PMT array, but this is not considered in this analysis.

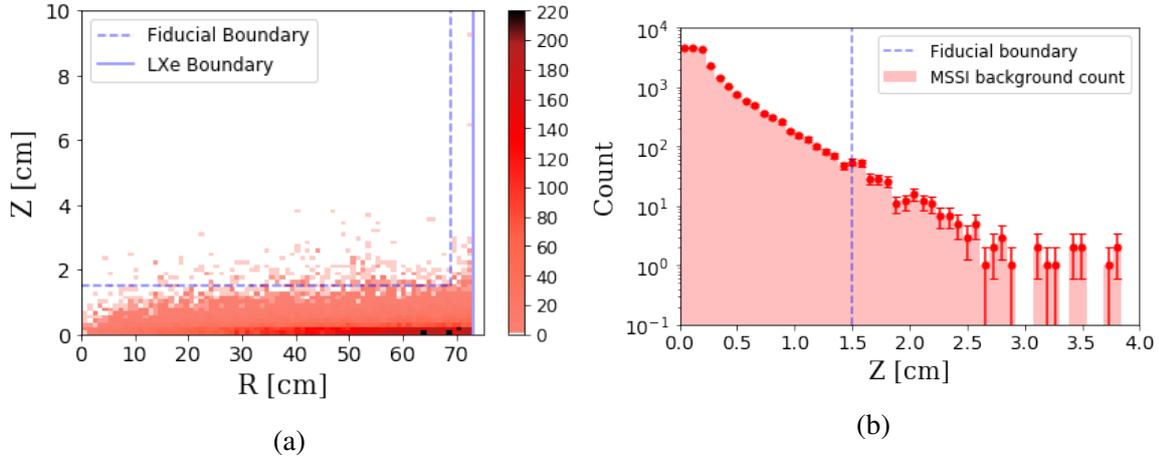


Figure 4.7 Position of MSSI backgrounds originating from $^{131\text{m}}\text{Xe}$ events. Figure (a) shows the R - Z distribution of MSSI events after MS cuts. Figure (b) shows the distribution of events after the MS cut and radial fiducial cut.

these MSSI events. However, one notable difference exists in the final row of the EFT search - an application of an NR band cut. This was done because in this analysis not all events were removed by the fiducial cut. Until this point, the cut had been neglected since all events were already removed. But now that events were passing the fiducial cut the NR cut was applied to determine the number of background events in the signal region.

For MSSIs the standard WIMP search ROI removed $90\%_{-21\%}^{+10\%}$ of all events. This was a significantly higher proportion of events than was the case for PEDs; indicative of the higher energy range MSSIs. This point is echoed by the MSSI removal rate of the expanded EFT search ROI which only removes $77\% \pm 1.8\%$ of MSSIs. The MS cut is also seen to be rather ineffectual, with zero, and ten events being removed respectively for the standard and EFT WIMP searches. Expectedly, OD, and skin cuts do not contribute to MSSI reduction. The reason for the ineffectiveness of the skin cut for MSSI events compared to PED events has to do with each population's R - Z distribution. Given the clustering of PED by the wall in Figure 4.6, their reduction by the skin cut was understandable. However, MSSI events do not distribute themselves in the same way. Figure 4.7a shows the R - Z distribution of all MSSI events. It can be clearly seen that these events tend to cluster by the cathode rather than by the wall. Furthermore, the fact that these events are caused by a MS of a particle that seems to be travelling vertically, it is less likely that such events would trigger the skin sensors.

The same distribution explains the effectiveness of the fiducial cut, which once again significantly reduces the MSSI background. A key difference in fiducialisation for MSSI

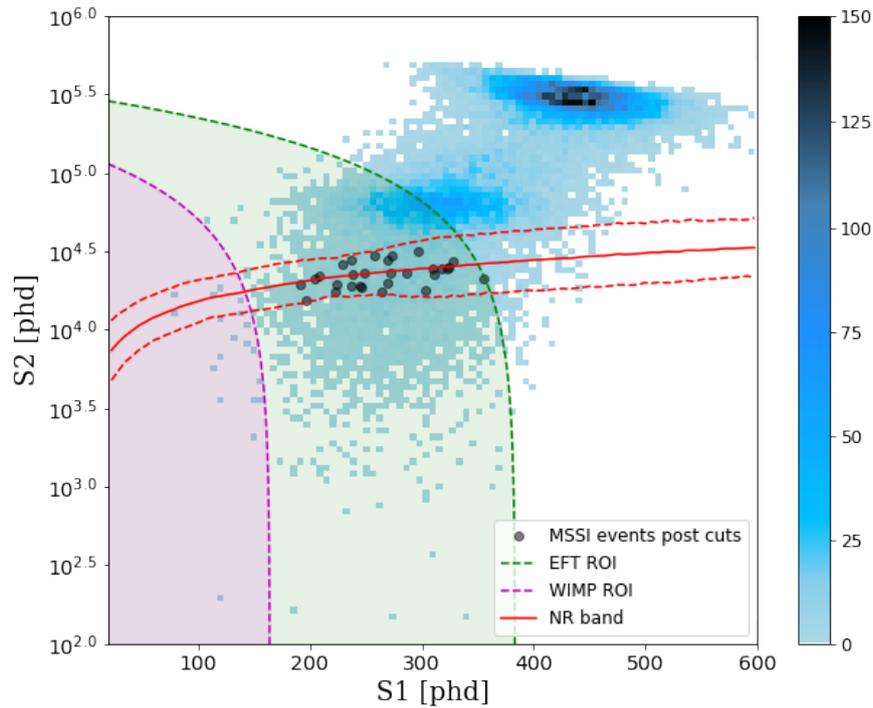


Figure 4.8 $^{131\text{m}}\text{Xe}$ MSSSI events in S_2 vs. S_1 space. In blue are all simulated MSSSI events, with the final background selection for the EFT search shown in black. The green region and magenta regions represent the EFT and standard WIMP search energy ranges respectively. The lack of black events in the magenta region indicates no known $^{131\text{m}}\text{Xe}$ induced MSSSI background for the standard WIMP search.

events is that now it is not the radial cut responsible for removing the bulk of these events, but rather the Z fiducial cut above the cathode. This is highlighted in Figure 4.7b, which shows the vertical position of MSSSI events after the radial fiducial cut is applied. This figure shows that MSSSI events are capable of penetrating up to ~ 4 cm above the cathode, leaving an opportunity for MSSIs to be a background in DM searches. This may not be the case for the standard WIMP search, where the fiducial cut removes all remaining MSSSI events not cut by the ROI. But for the higher energy searches, such as the EFT search, a small fraction does remain.

The remaining MSSSI events are further reduced by the NR band cut, which is applied since it is only MSSIs which mimic an NR which are capable of being mistaken for a WIMP. To determine this fraction, a 2σ NR band was generated with NEST. This band was generated using the same g_1 , g_2 , and electric field strength described in Section 4.5, and are the red lines shown in Figure 4.8. This figure also highlights the EFT, and standard WIMP search ROIs which are shaded in green and magenta respectively. Also shown is the selection of

MSSI events in either ROI after all cuts. Here all cuts refer to MS, fiducial, EFT ROI, and NR band cuts. The final count of MSSI events in the EFT search is 29, which normalises to $8 \times 10^{-3} \pm 7 \times 10^{-3}$ after 1000 days assuming a 0.02 Hz $^{131\text{m}}\text{Xe}$ decay rate, as shown in the last row of Table 4.3.

For the standard WS ROI, the simulated rate was zero. Therefore, using the same rationale from Section 4.6, an upper limit on the expected MSSI count of 6.88×10^{-4} (90% CL) per 1000 days for the standard WIMP search.

A notable feature in Figure 4.7a is the presence of MSSI events up to ~ 9 cm from the cathode near the wall. Unusually for events so far from the cathode, these events do not appear to trigger the MS cut. It is thought that these anomalous events occur as a result of 163.93 keV gamma-rays which escape the LXe volumes, pass through the ICV, and scatter off of the OCV back into the LXe. Since the space between the ICV and OCV is a vacuum the MFP is elongated compared to what is presented in Figure 4.5. This facilitates the capability of high energy gamma-rays to penetrate further into the detector from the cathode in rare circumstances. The signature of these events in MC truth is a summed energy deposition of less than 163.93 keV (since energy loss from scatters beyond the LXe is not recorded). This appears to be the case for these events which have penetrated so high up into the LXe. The presence of this variety of MSSI events explains why MSSIs appear to penetrate deeper into the detector at larger radii than in the centre of the TPC. However, they are not a significant feature in the contribution to MSSI backgrounds since they are excluded by the radial fiducial cut in the same way that PED events are removed.

4.8 Conclusion

The $^{131\text{m}}\text{Xe}$ generator built for LZ has met all performance criteria relating to the generation of primary particle position, and energy distribution. Its use in simulating PED backgrounds showed that no PED events occur in the fiducial region after 3.5×10^6 days assuming a mean $^{131\text{m}}\text{Xe}$ decay rate of 0.02 Hz. This leads to an upper limit of 6.88×10^{-4} PED events at a 90% confidence level per 1000 days. This value was established regardless of ROI. The simulation of MSSI events from $^{131\text{m}}\text{Xe}$ decays also showed no events in the standard WIMP search after applying all data analysis cuts. However, for the higher energy EFT WIMP search, this was not the case. This meant that while the standard WIMP search also had the same 90% confidence level limit of 6.88×10^{-4} per 1000 days, the EFT search had the higher expected rate of $8 \times 10^{-3} \pm 7 \times 10^{-3}$.

Each of these background types will be revisited in later chapters. PEDs will be looked at again in Chapter 5, where the result of electric field variations is capable of causing a PED background to become fiducialised if unaccounted for. MSSI backgrounds from sources other than $^{131\text{m}}\text{Xe}$ will be looked at for the LUX detector in Chapter 6, where their impact on the LUX Run04 result will be examined.

Chapter 5

The Effect of Electric Field Non-Uniformity on Position Reconstruction in LZ

5.1 Rationale

Experience from non-uniformities in the electric field of LUX led to an increased effort in handling the trajectories of drifting electrons in LZ. This meant that realising the stated MDC2 goal of reconstructing interaction sites in XYZ required the development of two abilities in LZ's software. First was the ability for BACCARAT to handle the simulation of electron trajectories through a non-uniform electric field. The second was LZap's ability to correct the position once given knowledge of the electric field and the paths of drifting electrons. In order to meet these requirements an Electron Drifting module (ED module) was developed to simulate the path of electrons in LXe, under the influence of an electric field.

Before describing this ED module, and its usage by LZ, this chapter will begin with a description of LZ's most recent electric field simulations. This will be followed by the results of two studies that used the ED module: one examining the effect of the grids on the position of electron extraction, the other comparing two different methods for determining the effect of electron diffusion. The chapter will end by discussing the use of the ED module in MDC2 mock data production, and by LZap.

5.2 Electric field in LZ

LZ relies on an electric field in the LXe bulk in order to create an S2 signal. The field is responsible for transporting ionisation electrons from their point of origin to the phase boundary. Once there, electrons are extracted into the gas phase and undergo photo-luminescence, producing a photon signal proportional to the original ionisation yield. In turn, these photons are detected by the top PMT array, producing the observed S2.

The electric field used for MDC1 (produced by the LBNL group), and in Section 4.5 assumed an ideal situation in which there is a uniform field strength with no radial component. However, in reality, this is not expected to be the case. As shown in Figure 5.1 a detailed modelling by Reed Watson of the field created by the field shaping rings, cathode, and anode of LZ indicates local variances in the field strength [104]. A particular feature of this electric field worth focusing on is that the radial field strength is non-zero. In fact, it increases closer to the walls, and as indicated in Figure 5.1a is at its highest by the wall at the top and bottom of the TPC. The higher radial field strength near the wall is due to the proximity to the field shaping rings. Of these rings, the cathode and anode, (at each end of the TPC) have the highest voltage. The incorporation of this non-zero radial field for MDC2 and beyond creates a need for LZ simulations to be able to offset the extracted radial position ($R_{EXTRACTED}$) of electrons from their point of origin (R_{ORIGIN}) when producing mock data. It will also require LZap to be able to reconstruct R_{ORIGIN} from $R_{EXTRACTED}$. These needs stem from the description of fiducialisation in Section 3.1.7, which relies upon the correct reconstruction of interaction sites in the LXe.

Notably, the electric field simulation shown in Figure 5.1 assumes a radial symmetry, which is unlikely to be the case when LZ goes online. LZ can be assured of an eventual radial asymmetry in its electric field based on the experience of previous dual-phase Xe TPC experiments such as LUX [85]. However, since LZ has not been built yet this asymmetry has not yet been observed, and its precise effects are currently undetermined. Consequentially asymmetry in the electric field has not yet been built into these simulations.

Field simulations for LZ were produced with COMSOL, a simulation package for finite element analysis [105]. COMSOL facilitates the use of partial differential equations (PDEs), allowing for the simulation of conventional physical systems. The output of the COMSOL simulation is a *.txt* file, which lays all R - Z points out in a grid, of 0.1 cm resolution.

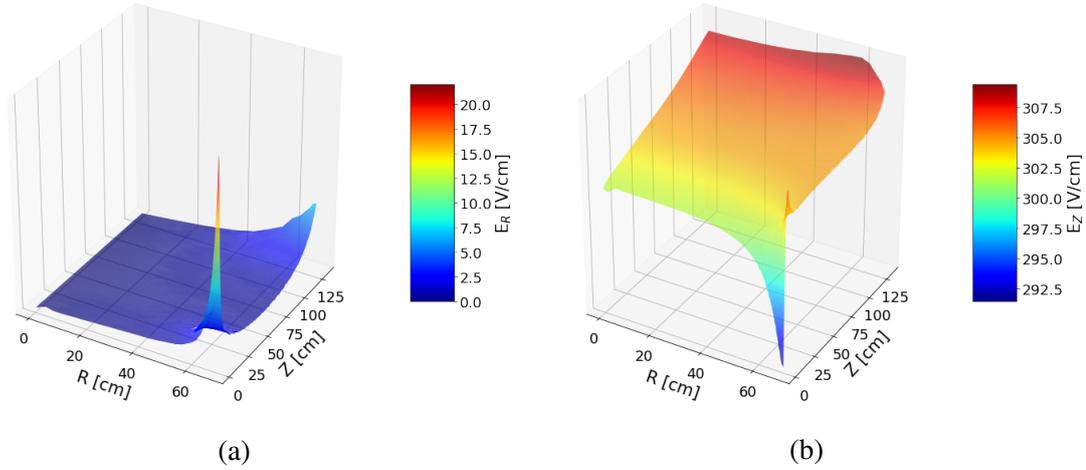


Figure 5.1 Electric field strength in LZ fiducial region based on up to date modelling taken from [104]. The plots on the left and right show the field strength in the radial and vertical directions respectively. In this figure, the cathode position is $Z=0$ cm.

5.3 Electron trajectory in LZ

The electric field in LZ's TPC is responsible for causing electrons to drift towards the phase boundary. However, another stochastic form of motion exists called diffusion.

Drift velocity is the non-stochastic component of electron motion. Based on the formula $\text{force} = \text{electric field strength} \times \text{charge}$ ($\mathbf{F} = \mathbf{E}q$), a constant electric field strength should yield a constant acceleration for free electrons. However, due to interactions with the LXe medium, a retarding force is applied to drifting electrons which results in a constant drift velocity. The situation is analogous to a mass falling in a gravitational potential such as the Earth, which also experiences a constant gravitational force but sees the body eventually reach a terminal velocity due to air resistance. A key difference between the falling body and the drifting electron is that drift velocity is achieved essentially instantaneously, making the acceleration phase inconsequential. The relationship between field strength and drift velocity can be seen in Figure 5.2 from [106], where the EXO-200 experiment has measured this value over a large range of electric field strengths.

The other type of motion experienced by drifting electrons is known as diffusion, which results in random dispersal both tangentially and longitudinally. The stochastic process of diffusion arises from the same interaction with the LXe medium which retards electron motion; and is modelled with a Gaussian distribution, centred on the final position of an electron after following a drift line without diffusion. The variance of this distribution is

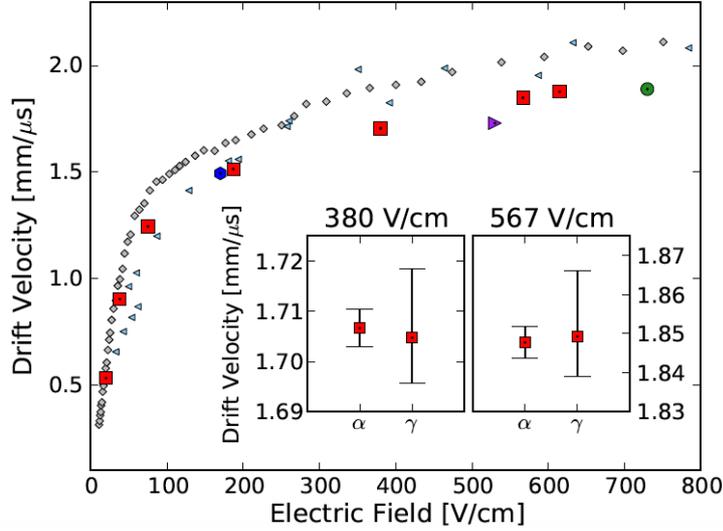


Figure 5.2 Electron drift velocity in LXe. Red squares were measured using EXO-200 at 167 K for two datasets: one using ^{228}Th calibration, and the other using alpha decays emanating from the cathode. The inset shows a comparison of the two methods of measurement at 380 V/cm and 567 V/cm in EXO-200. Also included are results from: EXO-200 at 163 K (cyan) [107] and at 165 K (grey) [108]; XENON10 at 177 K (green) [109]; XENON100 at 182 K (purple) [110]; and LUX at 175 K (blue) [111]. LXe operating temperatures are mentioned to account for possible variations due to temperature.

given by:

$$\sigma_{T|L}^2 = 2D_{T|L}\Delta t \quad (5.1)$$

where Δt is the amount of time that an electron has drifted for; and $D_{T|L}$ is the diffusion constant in the transverse and longitudinal directions respectively¹. Diffusion is highly anisotropic, which means that a distinction needs to be made for D_T and D_L [109]. The two diffusion coefficients have been well documented in literature and are commonly understood to be $D_T = 55 \text{ cm}^2/\text{s}$ [106] and $D_L = 12 \text{ cm}^2/\text{s}$ [109] in LXe.

Diffusion in the X , Y , and Z dimensions need to be modelled independently on a step-wise basis. This is done by considering the process as a random walk in each degree of freedom, for each time step dt . For instance, in the transverse planes, a random diffusion distance is sampled from independent Gaussian distributions with variance $\sigma^2 = 2D_T\Delta t$. It is common to consider the distribution in the 2-dimensional plane orthogonal to the electric field independently. This results in an initial point like distribution at $t = 0$, after time t consisting of a radial variance with [106]:

¹Transverse diffusion occurs independently in two dimensions (X , Y). The formula here describes motion in only one dimension.

$$\sigma_R^2 = \sigma_X^2 + \sigma_Y^2 = 4D_T \Delta t \quad (5.2)$$

Where σ_R^2 is the variance of the R distribution; and σ_X^2 and σ_Y^2 are the variances of the X and Y distributions respectively. Since X and Y are the two transverse dimensions $\sigma_X^2 = \sigma_Y^2 = 2D_T \Delta t$.

5.4 The Electron Drifter Module

The ED module was designed to determine the electron trajectory in LXe under the influence of an electric field. Electron paths are calculated using the Euler method for integration, where steps of either constant time or distance are taken. This section will describe the ED module, which itself consisted of several functions: `GetEntryFromTxt`, `GetErEzFromRZ`, `find_eDriftVelocity`, `GetVrVzFromErEz`, `eDiffusion`, and `Get_e_Line`. Note that the original names for functions are given so that readers from the LZ collaboration can closely follow this work and use it if needed.

The `GetEntryFromTxt` function was written to read the LZ field map. This field map was built with COMSOL, which outputs field maps in a text file format. Each line of the text file gives the total (E_T), transverse (E_R), and longitudinal (E_Z) electric field strength at each R - Z position in the detector. `GetEntryFromTxt` reads a specified line from a given field map, and returns a user-specified entry. At the time that this module was being developed the electric field was assumed to be radially symmetric.

The `GetErEzFromRZ` function determines the E_T , E_R and E_Z for a given set of R - Z coordinates. It does this by taking the user-specified R - Z position, locating the four closest points in the field map, and then performing a 4 point interpolation to estimate the electric field strength at the given position.

The `find_eDriftVelocity` function determines the electron drift velocity for a given electric field strength. This is achieved by using the EXO data in Figure 5.2 (i.e. the data represented by red boxes), and linear interpolating between the two closest values.

The `GetVrVzFromErEz` function gives the transverse (v_R), and longitudinal (v_Z) velocities of an electron. By combining the previous two functions, this function begins by taking the user-specified R - Z position inside the LXe, at determining the total velocity (v_T) of the electron based on E_T . It then divides this velocity into its constituent v_R and v_Z based on the

angle between the vectors E_R and E_z .

The `eDiffusion` function was designed to approximate electron diffusion. Since diffusion is a stochastic process it is not needed to determine the mean position of electron motion. This means that the calculation of $R_{EXTRACTED}$ for a cloud of S2 electrons diffusion can be disregarded. Since this $R_{EXTRACTED}$ was the immediate reason that this module was written, `eDiffusion` was designed so that it could easily be switched on or off at the user's discretion. Diffusion is modelled by considering motion independently in each spatial dimension. Since the ED module only considers motion in R - Z plane (i.e. with electrons either moving radially towards or away from the centre of the TPC)², information about X - Y positions are not retained from step to step. But given that diffusion requires the consideration of motion independently in 3 dimensions `eDiffusion` was only designed to model diffusion in 2 dimensions. Although this is not physical it allows for simplistic debugging and is a useful step in the development of electron diffusion in 3 dimensions. Diffusion was handled in a step-wise fashion, implemented after an electron had drifted at a constant velocity for a user-specified time (Δt). Alongside Δt , `eDiffusion` takes four other arguments. These arguments are two pairs of R - Z (where R and Z are in the horizontal and vertical planes of detector) coordinates corresponding to the initial (R_i, Z_i), and final (R_f, Z_f) position of an electron after drifting for Δt without diffusion (as illustrated in Figure 5.3). The effect of diffusion is then added. This is done by randomly sampling two numbers: the longitudinal drift length $\Delta Z'$, and the transverse drift length in X ($\Delta X'$)³. $\Delta Z'$ was obtained by sampling a Gaussian distribution with a mean of 0 and variance of $2D_L\Delta t$, whereas $\Delta X'$ was sampled from a Gaussian with a mean of 0 and variance of $2D_T\Delta t$. Finally, the diffusion had to be applied to the existing electron trajectory. To do this we must consider that longitudinal (transverse) diffusion occurs in the direction of (at right angles to) the path of the electron. As such it follows that applying this diffusion is not merely a matter of adding $\Delta Z'$ ($\Delta X'$) to Z_f (X_f) since X and Z are the coordinate system used by BACCARAT and LZap defined with respect to the TPC. Instead, the effect of the diffusion needs to be translated into the coordinate system used by BACCARAT and LZap (X_{NEW}, Z_{NEW}). To do this we apply a transformation and a translation:

$$X_{NEW} = \Delta X' \cdot \sin \theta_X - \Delta Z' \cdot \cos \theta_X + X_f \quad (5.3)$$

$$Z_{NEW} = \Delta X' \cdot \cos \theta_X + \Delta Z' \cdot \sin \theta_X + Z_f \quad (5.4)$$

²Operating in R - Z instead of 3 independent spatial dimensions speeds up the operation of the module.

³ R is assumed to be entirely in the direction of X . Notably, R and X are interchangeable due to the assumed cylindrical symmetry.

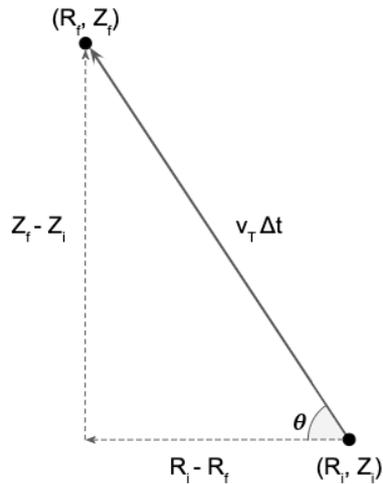


Figure 5.3 Diagram of an electron drifting in an electric field for time Δt . Electron begins at the coordinates (R_i, Z_i) , and finish at (R_f, Z_f) . This diagram represents a single step in the ED module algorithm described in this section.

where θ_X is the angle between the X axis of the TPC and the trajectory of the electron. The relationship between θ , R_i , R_f , Z_i , and Z_f are illustrated in Figure 5.3 (where R can be thought of as being entirely in the X direction for this simplification). As seen in Figure 5.1, the majority of cases will have a radial field strength of close to zero (i.e. $\theta \approx 90^\circ$). As indicated in Eq. 5.3 and Eq. 5.4, when theta is 90° , the translation terms cancel and we simply add $\Delta Z'$ ($\Delta X'$) to Z_f (X_f).

The `Get_e_Line` function combines the previous functions to simulate the path of an electron drifting through LXe under the influence of an electric field. The function takes several command-line arguments which are: an initial R - Z position, a step size (Δd), and an electric field file in `.txt` format. The function works by using finite steps of length Δd to move the electron from a user-specified initial position to a final position. At each step, the velocity is updated based on its new position.

When executing `Get_e_Line` the user needs to make three critical decisions: whether to drift electrons forward or backward in time, whether to apply diffusion or not, and whether to make steps with constant time or constant distance. Each decision will be based on whether the user wishes to simulate or reconstruct events, which will be discussed in Sections 5.6.

The result of this can be seen in Figure 5.4 which shows several drift lines generated at intervals of 10 cm, beginning from a 1 cm above the cathode, and ending at the phase boundary at $Z=146.1$ cm. Lines in black were generated without diffusion, while the red

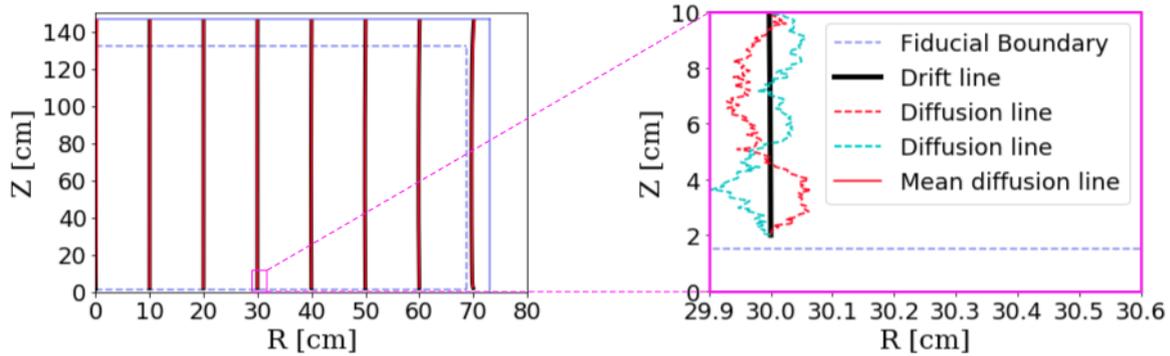


Figure 5.4 Electron trajectory in LZ field as indicated by a series of drift lines. Lines are generated at 10 cm intervals, beginning 1 cm above the cathode. The left plot shows drift lines without diffusion (black), and the average of 100 different drift lines with diffusion beginning from the same location (red). Red lines overlap with black for large sample sizes. The plot on the right shows a close up of one drift line, with two individual diffusion lines alongside. The field used for this simulation is the same as from Figure 5.1.

lines represent the average of 100 drift lines, each with diffusive effects on. The effect of diffusion is not noticeable at this scale, with the red lines not diverging from the black in any discernible manner. To visualise the effect of diffusion a zoom in on one of these drift lines is shown to the right of the first plot. This plot shows two individual diffusion lines, where the stochastic effect of diffusion can be seen more clearly. All lines used a constant step length of 0.5 mm, half of the electric field map resolution. The field used for this simulation is the one shown in Figure 5.1. As such this represents the most realistic visualisation of electron trajectory at the moment of writing.

5.5 ED module analyses

5.5.1 LZ grids analysis

Electric field lines were generated as an early test of the ED module's capabilities. In most regions of the LXe, these lines were fairly unremarkable, remaining largely vertical as exemplified by Figure 5.4. However, around the cathode and anode grids, a type of field line funnelling was observed. To demonstrate this effect, and determine its capacity to influence radial position reconstruction, independent simulations of electron motion around the top and bottom grids were performed.

The top grid analysis was performed by generating a large number of field lines with the ED module. These lines were generated at a regular spacing of $R = 0.001$ cm, originating

just beneath the grids, at a depth of $Z = 145$ cm, and used a step size of $1 \mu\text{m}$. The result of this can be seen in the two plots of Figure 5.5a. The lower plot in this figure shows the path of field lines through the grids, the upper plot shows a histogram of their final position. For the purpose of clarity, not all field lines used in the histogram are shown in the lower plot. Instead, a selection of field lines at 0.2 cm spacing was used to clearly illustrate how the field lines converge. The position of the grids is indicated by the underlying heat map of the electric field strength. This heat map provides the electric field strength in V/cm .

The bottom grid analysis followed a similar procedure. Drift lines were generated 0.1 cm above the cathode at regular radial intervals of 0.1 cm. However, these field lines were terminated at a depth of 2 cm above the cathode. This is illustrated in Figure 5.5b which also shows two plots. Once again, the lower plot shows the path of drifting electrons, and the upper plot shows their final radial position at $Z = 2$ cm.

With a radial wire density of 4 wires per centimetre in the anode, Figure 5.5a shows the expected clustering of final field line positions at regular 0.25 cm intervals. This effect is mimicked by the cathode at the bottom of the TPC, which also clusters field lines at the same interval. The difference between these two effects however is that the top grid will influence the $R_{EXTRACTED}$ of all detected electrons, whereas the cathode will not. This is due to extracted electrons having to pass through the top grid in order to reach the gas phase, whereas the influence of the radial motion of electrons from the cathode can be seen to dissipate before electrons reach the fiducial boundary at $Z = 1.5$ cm.

Given the density of grid wires, funnelling is not expected to interfere with the XY position reconstruction algorithm. This stands for two reasons. First, this algorithm uses the aggregate position of a cloud of $S2$ electrons which averages out the effect. Second, the algorithm has a significantly poorer expected XY resolution of 3 cm [112], meaning that even for an individual electron the effect is an order of magnitude smaller than the existing uncertainty. For these reasons, the shift of individual electrons by a quarter centimetre will not have a meaningful impact on position reconstruction for WIMP searches where the $S2$ ROI begins at⁴ 420 phd.

In order to simulate field lines through this region of the detector, electron trajectories

⁴The unit phd is the unit used to measure $S1$ and $S2$. It stands for photons detected, in reference to the number of photons seen in the PMT arrays.

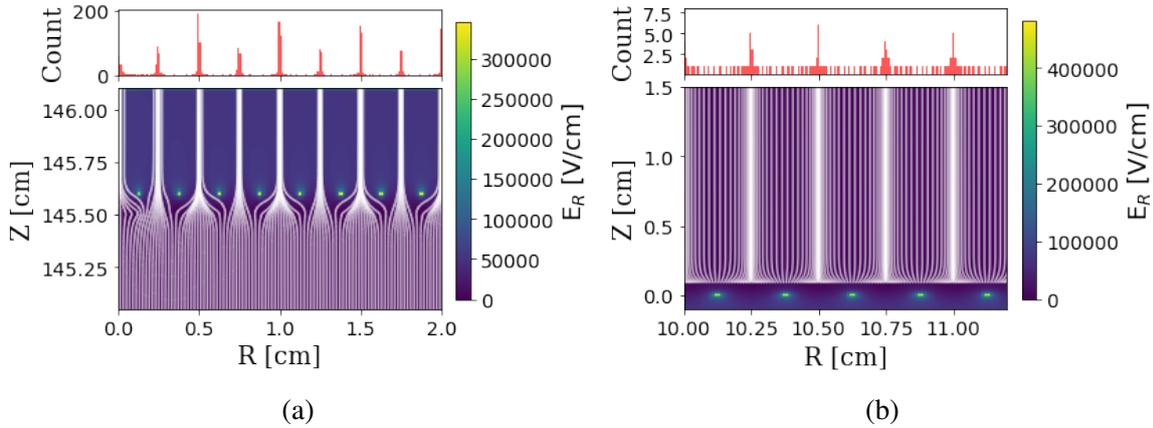


Figure 5.5 LZ's radial electric field (blue-yellow heat map) and field lines (white lines) near the grids and cathode. Figure (a) shows the region around the grids at the top of the TPC. The lower plot in (a) shows a heat map of the electric field strength, with yellow/green hot spots representing the high voltage of the grid wires. Field lines are overlaid on this heat map, with each field line represented in white. These field lines were generated with the ED module originating at $Z = 145$ cm with a radial spacing of 0.02 cm and terminate at $Z = 146.1$ cm. These field lines can be seen to converge in on one another as they pass through the grid. The hit pattern of electrons generated at $Z = 145$ cm with a tighter radial spacing of 0.001 cm can be seen in the histogram directly above. Figure (b) shows a similar analysis conducted around the cathode. This plot however uses a radial field line density of 0.01 cm, and the upper histogram shows the radial position at $Z = 2$ cm.

had to be finely produced. For this reason, Figure 5.5 stands as a benchmark for future tools which aim to produce field lines in LXe.

5.5.2 Diffusion analysis

In the lead up to MDC2 the ED module was selected to model electron trajectories in BAC-CARAT. The fact that diffusion was not modelled in 3 dimensions by the ED module was not seen as a significant issue for determining $R_{EXTRACTED}$, but for individual electrons diffusion needed to be taken into account to give the S2-photo-hit pattern a more realistic spread. To handle this, single electrons had diffusive effects approximated by using Eq. 5.2⁵, where the full drift time of the simulated S2 electron cloud was used for Δt .

In order to check the legitimacy of this parameterisation, nine electron trajectories from the bottom of the TPC were calculated, with diffusion turned off. This MDC2 parameterisa-

⁵In truth diffusion in X and Y were sampled independently. But since we are confining our description to the R - Z plane this is superfluous to this discussion.

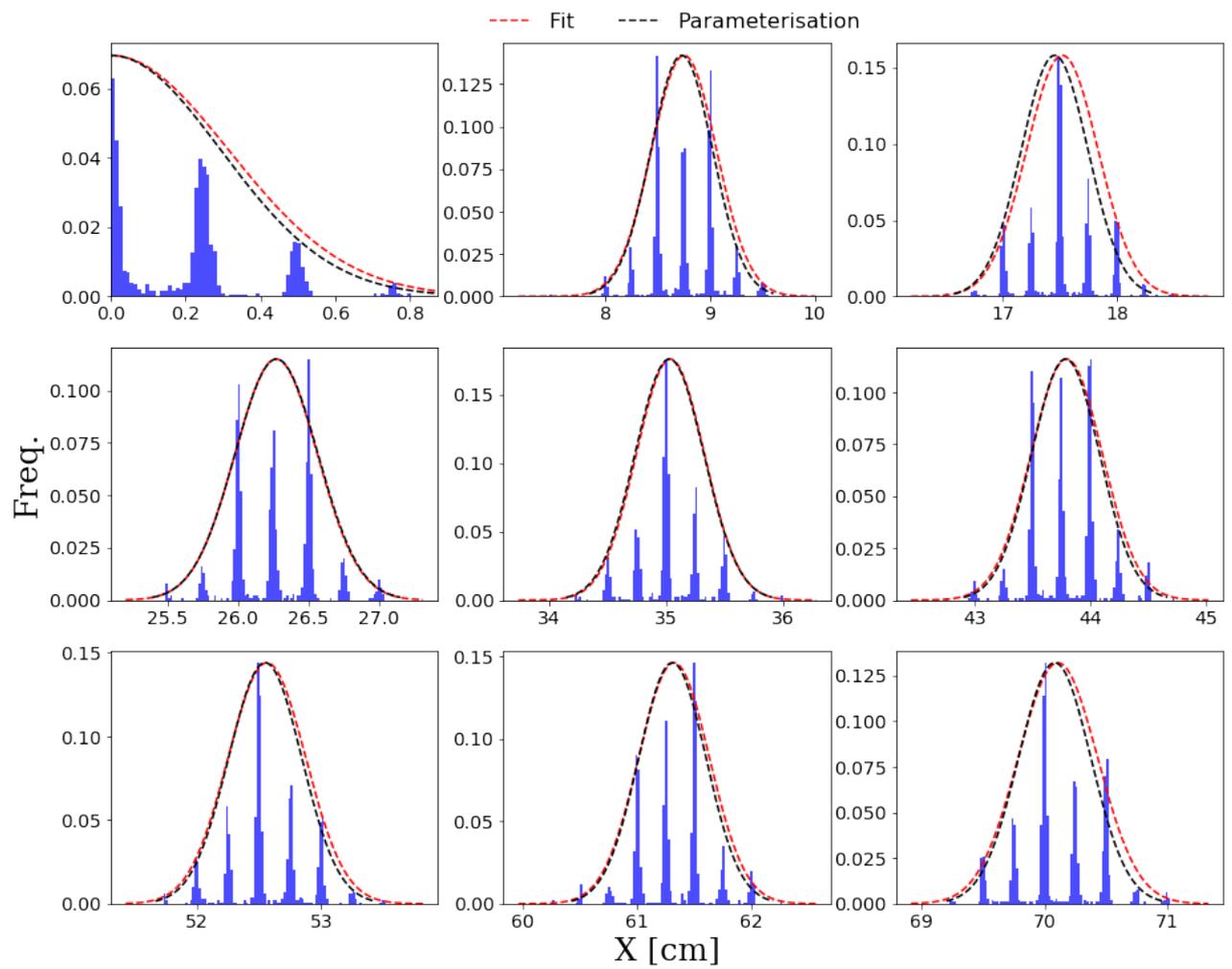


Figure 5.6 Comparison between the extracted position of electrons with diffusion turned on and off. The blue histogram in each plot represents the extracted position of electrons with diffusion turned on. The dashed red line is the Gaussian obtained by calculating the mean and standard deviation of the blue distribution. The dashed black line is the parameterised calculation.

tion was then applied to the output to create a distribution of events. This parameterised distribution was then compared to the distribution of $R_{EXTRACTED}$ positions found when the ED module simulated 1000 electrons from the same position, but with diffusion turned on. Since the ED module only applies diffusion in one transverse dimension, the diffusion parameterisation used by MDC2 had to be modified for the purpose of comparison. This was done by sampling from a Gaussian with a variance given by Eq. 5.1 instead of Eq. 5.2.

The result of this comparison can be seen in Figure 5.6, which shows the $R_{EXTRACTED}$ for nine different drift line analyses. In each case 1000 drift lines with diffusion turned on were generated, each of which had their $R_{EXTRACTED}$ illustrated in the blue histogram. Each drift line originated from a position 2 cm above the cathode and terminated at the phase boundary ($Z = 146.1$ cm). The lines presented all began between $R = 0$ cm, and $R = 70$ cm, at regular intervals of 8.75 cm. The figure shows a Gaussian obtained for each set of drift lines represented as a red dashed line. This Gaussian is arrived at by calculating the mean and standard deviations of the blue histograms.

A notable feature in Figure 5.6 is the presence of the “spikey” sub distributions which make up the histograms. The radial distance between these sub-peaks occurs at regular 0.25 cm intervals. This interval is consistent with the grid effects described in Section 5.5.1. Given this result, it would seem that grid funnelling is a likely explanation for this feature.

For all plots in Figure 5.6, parameterised diffusion, and step-wise diffusion appear to behave similarly. As a measure for the consistency between the ability of the two methods to determine $R_{EXTRACTED}$, the difference between the two Gaussian means ($\Delta\mu$) is calculated. The average $\Delta\mu$ for these nine drift lines is 0.290 mm. In order to determine if the spread of $R_{EXTRACTED}$ is being determined consistently between the two methods, the difference between standard deviations ($\Delta\sigma$) is used. For these nine drift lines, the average $\Delta\sigma$ is 0.140 mm. For the first plot where electrons originated at $X = 0$ cm, it is notable that there are no events with a negative X value. This is because the electric field map used to inform these calculations does not handle negative values and assumes a radial symmetry. As such, the parameter X is best thought of as a distance from the origin rather than a fully formed independent dimension.

The result of the ED module diffusion analysis lends to the legitimisation of electron diffusion parameterisation for MDC2 mock data. However, two considerations need to be made regarding this study. First, only transverse diffusion in one dimension is considered. The

actual parameterisation for MDC2 used Eq. 5.2 to determine the variance. Second, the MDC2 parameterisation does not rotate the diffusion according to Eq. 5.3 and Eq. 5.4. However, this does not seem to have a significant effect since this analysis only covers cases where the electric field is largely homogeneous throughout the electron trajectory (as indicated by Figure 5.1).

5.6 MDC2 application

5.6.1 Rationale

One of the goals of MDC2 discussed in Section 3.3.5 was to be able to handle a radially dependent electric field, and position-dependent electron drift velocities. To achieve this the ED module was used in two ways. First, it was used to calculate the drift time and the point at which any electron would be extracted into the gaseous Xe ($R_{EXTRACTED}$). This allowed for simulated $S2$ positions to be offset in the mock data from any origin (R_{ORIGIN}, Z_{ORIGIN}). How this was done will be discussed in Section 5.6.2. Second, since analysts would be blind to (R_{ORIGIN}, Z_{ORIGIN}), a module was built for LZap to take ($R_{EXTRACTED}$, drift time) and reconstruct the original coordinates ($R_{RECONSTRUCTED}, Z_{RECONSTRUCTED}$). How this was done will be discussed in Section 5.6.3.

5.6.2 Electron drift in BACCARAT

The electric field used for MDC2 was similar to the one shown in Figure 5.1, but with one critical difference - an exaggerated radial field strength. This stronger radial field would cause electrons to drift towards the centre of the detector as they ascended towards the gas phase. The increase to E_R was implemented not to meet a realistic expectation of field behaviour, but only to test LZap's ability to accurately reconstruct vertices under similarly extreme conditions as those experienced by LUX Run04 (as described in Section 3.2.3).

To implement this effect in BACCARAT, the ED module was called upon. This was necessary as BACCARAT itself did not have an inbuilt method for shifting the radial position of events. In lieu of this, the ED module generated a "drift map" for each 1 mm grid of the LXe volume in the TPC, based on the assumption that the radial field acted symmetrically. This map used step size of 1 mm, and deactivated step-wise diffusion to ensure stochastic effects did not influence the map. The drift map offset the position of $S2$ sites of in XY , and also provided a drift time based on the electron's trajectory.

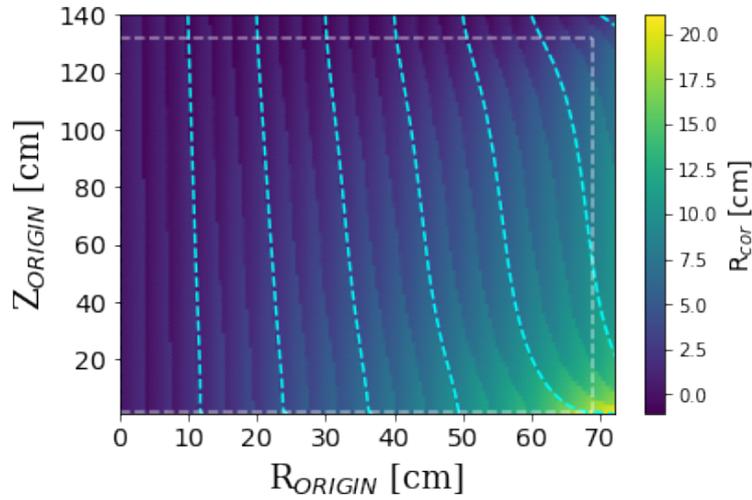


Figure 5.7 TPC map showing the effect of the electric field in LZ. The heat map shows the difference in $R_{EXTRACTED}$ and R_{ORIGIN} for each point in the LXe. The dashed aqua lines show the trajectory of selected electrons to the phase boundary. The fiducial volume is displayed, outlined in dashed white lines.

Diffusion was still an important feature to ensure S2-photo-hit-patterns had a realistic shape. However, it was added at the end of electron propagation instead of the step-wise manner described in Section 5.4. Diffusion implementation for MDC2 followed the same method used to create the parameterised Gaussian distributions in Section 5.5.2 - by applying Eq. 5.2 after the electron motion ends. In this implementation Δt is the full drift time of the event.

The radial effect of the drift map is illustrated in Figure 5.7. This image shows a heat map representing every 1 mm^2 of the LXe. At each position, the radial correction (R_{COR}) used for simulations is given by $R_{ORIGIN} - R_{EXTRACTED}$. Overlaid is an outline of the fiducial region in white, as well as several electric field lines in aqua. These field lines served a specific purpose which are the topic of focus for Section 5.6.3. It is notable that this particular selection of drift lines does not demonstrate the grid effects described in Section 5.5.1. However a close examination of the background drift field map reveals an apparent corrugation. This corrugation is caused by the grids.

5.6.3 LZap XYZ Position Corrector Module

Once the reconstructed position of mock data had been adjusted to account for the effect of the electric field, the need for the XYZ Position Corrector was apparent. Without the update of this module, LZap would reconstruct the position of events based on the assumption of a

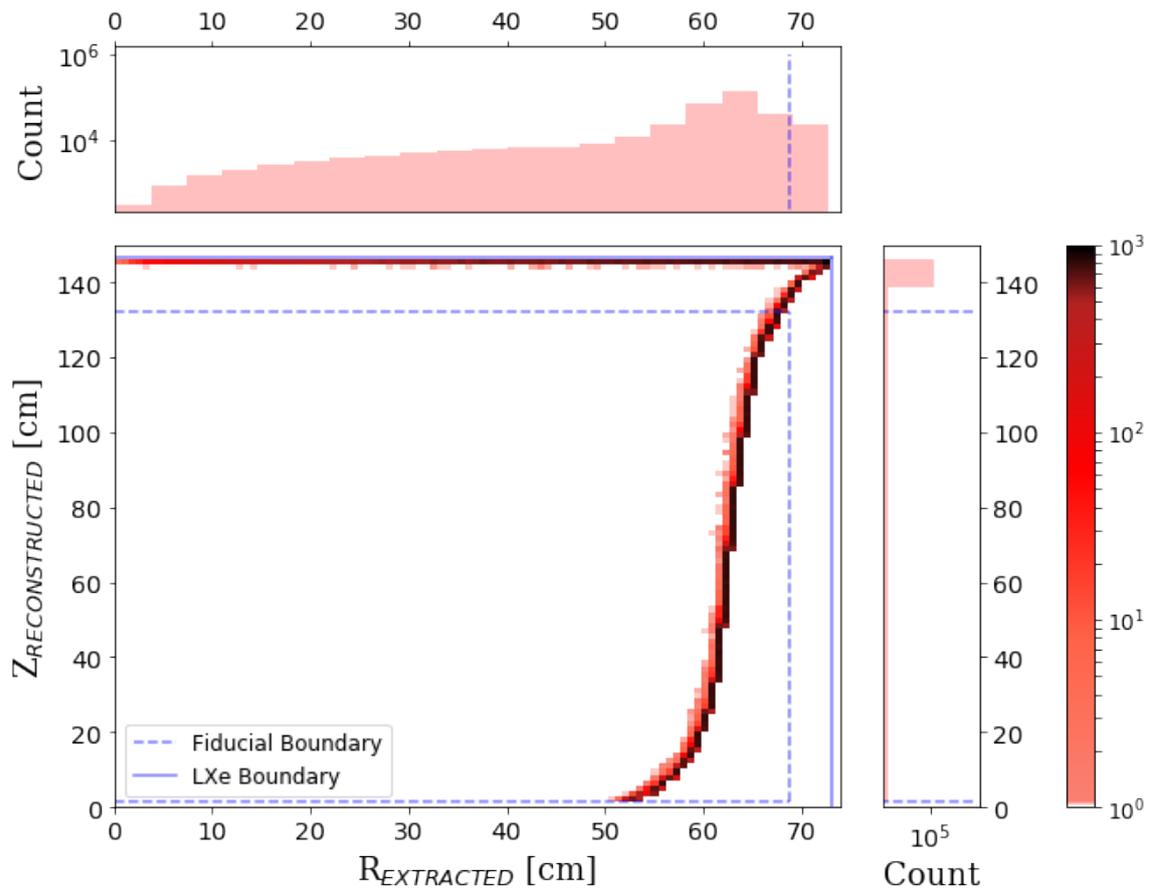


Figure 5.8 Position of all simulated $^{131\text{m}}\text{Xe}$ PED events from Chapter 4 offset by the MDC2 drift map used by BACCARAT. Central plot shows all events in R - Z space which have clearly experienced an inward shift which scales with depth. The two histograms in the margins shows the same events in a single dimension.

perfectly vertical electron drift, at constant velocity of $0.15 \text{ cm}/\mu\text{s}$. While this was suitable for the electric field used for MDC1, to use the same approach for MDC2 data would be inappropriate. This is illustrated in Figure 5.8, which shows the reconstructed position of the $^{131\text{m}}\text{Xe}$ PEDs from Figure 4.6 under the influence of the MDC2 field. What Figure 5.8 shows is that 64.4% of the $^{131\text{m}}\text{Xe}$ PEDs from Chapter 4, which were entirely removed by the fiducial cut, will be reconstructed inside the fiducial region if the MDC2 field is not taken into consideration. $^{131\text{m}}\text{Xe}$ PEDs are of course not the only backgrounds which would have been mis-reconstructed inside the fiducial region: wall events, external neutrons, and any other background which we primarily rely on the fiducial cut to remove will be impacted by this. However, the choice to focus on the effect this would have had on $^{131\text{m}}\text{Xe}$ PEDs is due to their familiarity to the reader on account of their introduction in Chapter 4, and their close mapping to the LXe boundary. Specifically, the heat map in the central plot in Figure 5.8 clearly shows the reconstructed position of the TPC wall, since it lines up perfectly with the position of $^{131\text{m}}\text{Xe}$ PEDs.

XYZ Position Corrector fixed this problem by taking the reconstructed *XY* position of an event based on its *S2* photo-hit-pattern to obtain an $R_{EXTRACTED}$, combined it with information about the electric field to determine R_{ORIGIN} . The module also determined the *Z* coordinate based on the drift time of the *S2* event. To achieve this the *XYZ Position Corrector* used a series of electron drift lines created by the *ED* module. A selection of these field lines are represented by dashed aqua lines in Figure 5.7. In this figure the majority of the field lines used by LZap are omitted in order to make clear both: the trajectory of electrons, and how these lines warp at increasing radii. Field lines were produced by creating an electron with the *ED* module (with diffusion turned off), at a position on the phase boundary. The *ED* module was then set to step backwards in time, so as to retrace the path of the electron which ended up at this particular position on the phase boundary. Each of these “reverse drift lines” extended from the top of the LXe, down to either the cathode, or any other LXe boundary if the electron trajectory was not entirely vertical. Drift lines were generated at 1 mm intervals, and used a constant step time of $0.1 \mu\text{s}$. The choice of drift line density and step time was made in order to maintain the existing resolution of the field map when transporting individual electrons. The effect of the grids discussed in Section 5.5.1, limits this resolution to 0.25 cm. The choice then of 0.01 cm line density ensures that the *XYZ Position Corrector* module can resolve this effect.

The *XYZ Position Corrector* used these lines to determine the initial position of events once provided with an *S2 XY* position and drift time. For events which fell precisely on one

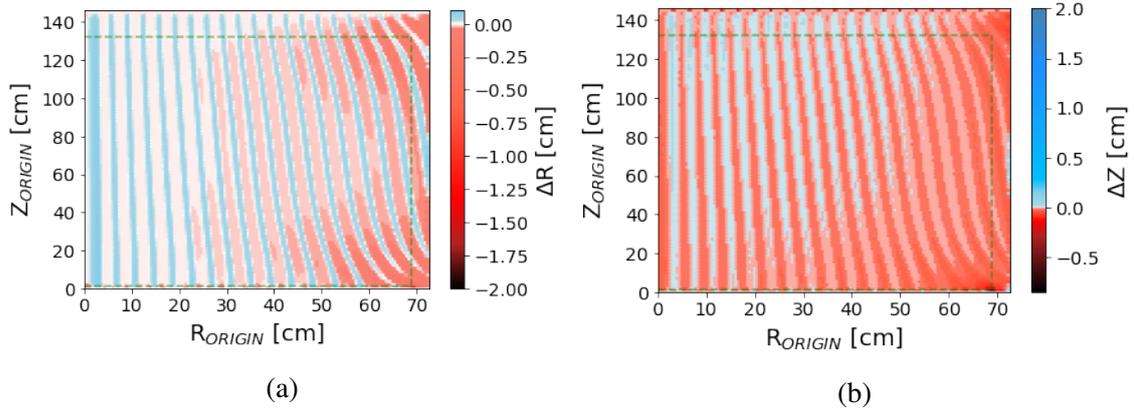


Figure 5.9 The difference between reconstructed and true positions of events as a function of their position in the TPC. (a) shows the difference between $R_{RECONSTRUCTED}$ and R_{ORIGIN} , and (b) shows the same for Z . The dashed green line in both figures represents the fiducial region.

of the input drift lines, the XYZ Position Corrector interpolated the initial position based on the drift time. For events which fell between drift lines event position was reconstructed via a 4 point interpolation between two adjacent drift lines.

5.6.4 Validation

In order to validate the usage of the ED module as described in Sections 5.6.2 and 5.6.3, the reconstructed R - Z coordinates for every 1 mm grid of the LXe was compared to its original R - Z . This procedure began with using ED module to generate a map of the TPC giving the $R_{EXTRACTED}$ and drift time for every 1 mm² of the TPC. Each datum in the map then had its R_{ORIGIN} and Z reconstructed using the XYZ Position Corrector.

Figure 5.9 shows the outcome of this test, with Figure 5.9a displaying the difference between $R_{RECONSTRUCTED}$ and R_{ORIGIN} (a.k.a. ΔR), and Figure 5.9b showing the same for Z . Both figures show close agreement between reconstructed and original positions with the mean ΔR of 0.0071 cm and mean ΔZ of -0.0016 cm (where $\Delta Z = Z_{RECONSTRUCTED} - Z_{ORIGIN}$). The figures are also relatively homogeneous with ΔR variance of 0.0011 cm², and ΔZ variance of 0.0002 cm². However, although rare, there are some notable exceptions in this map to both the agreement and homogeneity. If the parameter space is restricted to within the fiducial region these values drop to $\text{Exp}(\Delta R) = 0.0065$ cm, and $\text{Exp}(\Delta Z) = -0.0012$ cm. The variances of these distributions also drop when looking only at vertices inside the fiducial volume with $\text{Var}(\Delta R) = 0.0009$ cm², and $\text{Var}(\Delta Z) = 0.0001$ cm². This indicates that the most

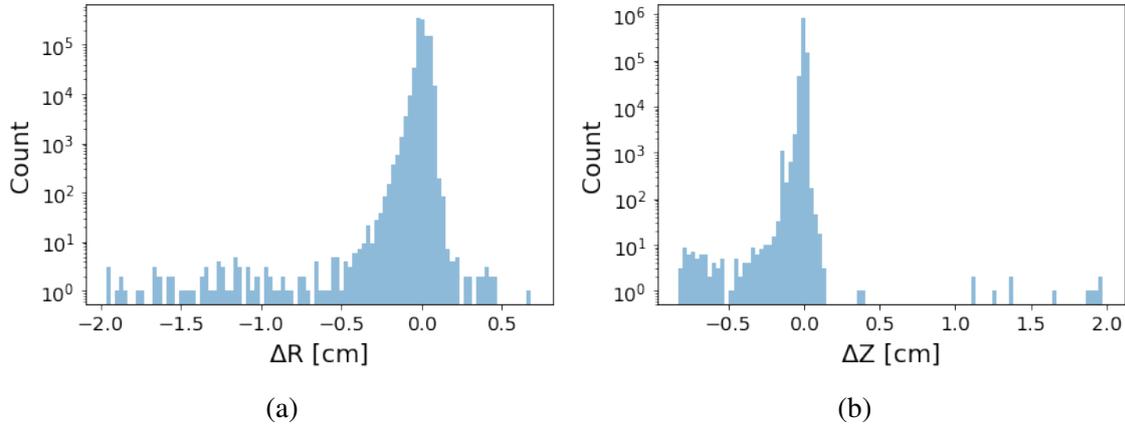


Figure 5.10 The difference between reconstructed and true positions for all events. (a) represents this difference in R , and (b) shows it for Z .

poorly reconstructed vertices are closer to the cathode, anode, and field shaping rings. This observation also bears out in Figure 5.9, where the magnitude of both ΔR and ΔZ increases closer to the field shaping rings.

Some of the inhomogeneity in Figures 5.9a and Figure 5.9b can easily be seen in the form of a corrugation across the LXe. This effect is created by grid funnelling discussed in Section 5.5.1. However, this effect does not account for the most poorly reconstructed events. In reality, 0.005% of events have a $\Delta R < -1.0$, and 0.0028% a $\Delta Z > 0.1$. These events do not stand out in Figure 5.9, but are displayed in Figure 5.10 which shows the histograms of ΔR in Figure 5.10a and ΔZ in Figure 5.10b for all events. What is clear from looking at these figures is that some rare cases events are being reconstructed up to 2 cm from their true positions, potentially leading to some non-fiducial events being reconstructed inside the fiducial volume.

To check this, non-fiducial events which have been reconstructed within the fiducial region were selected. Of the 1.05×10^6 events which make up this TPC map, only 34 fell into this category, and are displayed in Figure 5.11. Figure 5.11a shows the original position of these events, and Figure 5.11b shows them after being reconstructed. What these figures show is that despite the small step time/step size used by the ED module, events that should be removed by the fiducial cut still have a small chance of being mistakenly reconstructed inside the fiducial region. Whether this effect is being caused by the ED module's implementation into BACCARAT or the XYZ Position Corrector is not obvious. Nevertheless, the cause is clear, since all poorly reconstructed events rely on drift lines passing through regions with high electric field strengths. This points to the likely solution being an even smaller

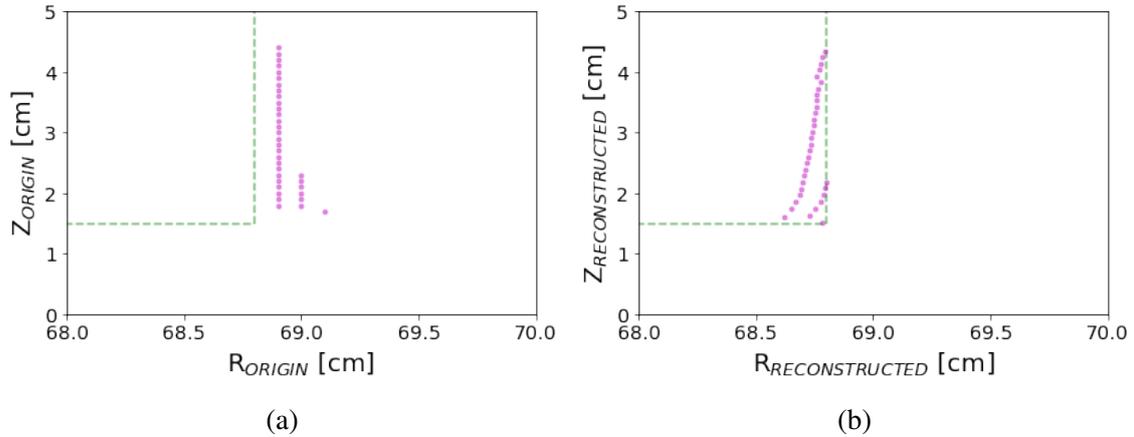


Figure 5.11 All non-fiducial events which are reconstructed inside the fiducial region. In both figures, the dashed green line represents the fiducial region, and the magenta dots the event positions. (a) shows the true position of these events, and (b) shows their reconstructed positions.

step time/step size when using the ED module. While this is a relatively easy fix, given the timescale for MDC2 it was not feasible. As such, these events were removed at an MC truth level to avoid a reduction in the fiducial volume. In any event, this observation will be of interest to future LZ analysts, and further highlights the need for attention to be paid to the reconstruction of electron trajectories, particularly as the electric field strength increases. For this not to be addressed when handling real data could lead to uncertainty over the origin of a potential DM signal, or a reduction in detector sensitivity.

5.7 Conclusion

The electric field used to drift S2 electrons towards the gas phase is an essential component of LZ's TPC. As such, its effect on the trajectory of electrons must be understood and modelled. This was first achieved by the ED module: capable of simulating electron paths in LXe with, or without diffusion.

The ED module was used in a couple of studies for LZ. First, to demonstrate the effect of the LZ grids to cause a funnelling of electrons passing the grids into discrete bins of 0.25 cm separation. However, this effect is not expected to affect the overall XY position resolution which is already at 3 cm [112]. Second, it was used to check whether the parameterised approach to electron diffusion implementation used for MDC2 matched the step-wise approach. This was done by comparing the extracted position for events using each method

on various drift lines originating at 2 cm above the cathode. The two approaches were found to have a close agreement; with an average difference of only a 0.29 mm, and a standard deviation in the difference of 0.14 mm.

The ED module was also used to determine event drift times and the extracted positions of electrons in the mock data of MDC2. It was further used as the basis of the LZap XYZ Position Corrector module for reconstructing the original position of events based on the drift time and the radial *S2* position determined from the *S2*-photo-hit-pattern.

The results of these studies justified the use of the ED module for these purposes. However, this will have to be implemented with finer resolution, which will require improvements to the running speed of the ED module. Furthermore, the need for a full implementation of step-wise diffusion in 3 dimensions, and the interpretation of an asymmetrical electric field, represents a significant future step in this module's development. Achieving this will demand additional upgrades to the ED module's handling and parsing of data.

Chapter 6

Gamma-X event simulation for LUX

6.1 Rationale

In the lead up to the LUX Run04 EFT search, a new gamma-X cut¹ was developed, with the stated goal of outperforming previous gamma-X cuts. Previously developed gamma-X cuts were assessed based on their ability to retain known single scatter events (i.e. efficiency) rather than their ability to remove known gamma-X events (i.e. background rejection rate). The reason for this approach follows a clear logical process: gamma-X and single scatter (SS) events cannot be distinguished, but gamma-X events are rare, so a cut that removes a high proportion of events from data is likely to be misidentifying single scatters as gamma-X events. Furthermore, cuts were developed based on apriori arguments about what a gamma-X event pathology should look like in a dual-phase Xe TPC, rather than from observations or simulations. This approach also stems from a lack of any known gamma-X dataset or simulations upon which cuts could be benchmarked.

To reach the aforementioned goal of outperforming previous gamma-X cuts, a set of gamma-X events were simulated. As was seen in Chapter 4, BACCARAT can produce gamma-X (a.k.a. MSSI) events as part of a large simulation. With this, previous gamma-X cuts could be benchmarked, and the parameter space of gamma-X events could be compared to single scatters in order to determine parameter spaces that would be suitable for developing a cut.

This chapter will describe this simulation effort and subsequent parameter space search.

¹Gamma-X events were introduced in Chapter 4.

Section 6.2 will begin by providing a detailed definition of a gamma-X event, then describe their identifying features, and finally discuss sources which are capable of creating gamma-X events. This will be followed by Section 6.3, which will give a description of previously developed gamma-X cuts, and discuss each cut's: similarities, differences, and expected performance. Following this will be Section 6.4 containing a description of the Run04 EFT data analysis cuts. These cuts were used in all LUX data selections throughout the following chapters. With this groundwork in place, the chapter will move onto Section 6.5, which describes the gamma-X simulations, and their validation. Finally, Section 6.6 will compare these simulations to calibration data to decide which parameters might be useful in a gamma-X cut.

6.2 Gamma-X events

6.2.1 Description

A gamma-X event is a phenomenon observed in Xe dual-phase TPCs. They can be considered a type of multiple scatter (MS), as they occur when an incident gamma-ray leaves at least two distinct energy depositions in a TPC's LXe. What characterises a gamma-X from other types of MS events is a non-detection of the $S2$ signal associated with one of these scatters. This feature allows gamma-X events to bypass the typical MS filtering process outlined in Chapter 3.

Gamma-X events are problematic for dark matter searches because of their ability to mimic a WIMP. This occurs when the energy deposition of the gamma-X event is reconstructed in the nuclear recoil (NR) band. But gamma-X events are MS by nature. As such, ER and NR events cannot be distinguished from one another with the method described in Section 3.1.6, which assumes events are all SS. As explained in Section 3.1.6, events are identified as either ER or NR based on their $S2/S1$ ratio. Since the scintillation photons from all energy depositions of a MS event will be detected as a single $S1$, MS identification relies on separated $S2$ signals. But in the case of a gamma-X, there is only one $S2$ signal. Thus the event is observed as a SS with a lowered $S2/S1$ ratio, potentially pushing it into the NR band. When this occurs in the energy range of a WIMP search, we are presented with a signal indistinguishable from a WIMP.

The production of a gamma-X event is contingent on a MS "losing" at least one of its $S2$ signals. This can happen if an energy deposition occurs in a region of the TPC where the

electric field lines do not drift ionised electrons towards the phase boundary. This means that ionisation electrons created at the interaction site never drift towards the gas phase, are never extracted, never undergo the photoluminescence, and hence never create the S_2 signal. For Run04 of the LUX experiment, this area of lost S_2 signal was confined to the reverse field region (RFR). The RFR is a small area between the lower PMT array and the cathode. As will be shown in Section 6.5.4 this means the major source of gamma-X events is from radioactive impurities in detector components near this RFR.

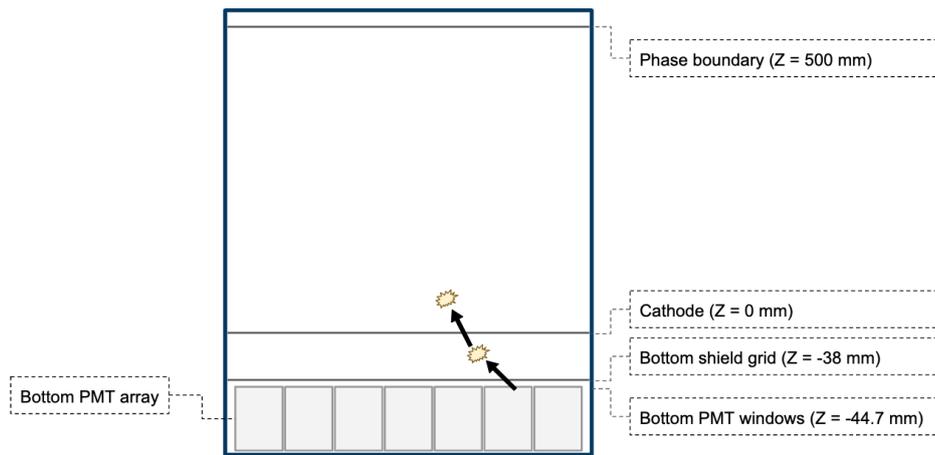


Figure 6.1 A schematic of a gamma-X event. This example illustrates a gamma-ray originating from the bottom PMT array (gamma-ray displayed as a bold black arrow, and the PMT array as grey blocks). The first scatter (represented in yellow) is below the cathode (therefore has its ionisation signal lost), and the second is in the fiducial region. For positional reference the height above the cathode is listed for various detector components in the LUX TPC.

To illustrate this background, a schematic of a gamma-X event originating from the bottom PMT array is shown in Figure 6.1. This image shows an emitted gamma-ray making two energy depositions: one below the cathode, and one above. This diagram is purely for the purpose of giving a basic illustration of a gamma-X event. The diagram is not to scale, nor is it intended to be representative of the PMT array.

As a point of nomenclature, gamma-X events fall under a category of backgrounds known as Multiple Scatter Single Ionisation (MSSI) events which were introduced in Chapter 4. The other kind of MSSI is known as a neutron-X, the difference between the two being that a gamma-X is thought to be caused by a gamma-ray, whereas a neutron-X by a neutron. Although neutron-X events tend to have an even lower S_2/S_1 ratio than gamma-X events, in practice there is no way to be certain whether an MSSI is a gamma-X or neutron-X. As such, gamma-X is often used within the LUX collaboration as a blanket term for all MSSI events.

The LZ collaboration differs on this point and specifically reserves the term gamma-X for the subset of MSSIs which are caused by a gamma-ray.

A difference between gamma-X (or MSSI) events in LUX and LZ is their origin. In LUX the only region where the electric field does not transport ionised electrons to the gas phase is in the RFR. However, in LZ field lines near field shaping rings (FSR) can point towards the wall in a few small regions. These localised pockets can lead to the production of “MSSI wall events” as they are known. Due to the charge build-up on the PTFE which occurred during LUX Run04, the source of these MSSI wall events was eliminated. As such MSSI wall events are not considered by any of the gamma-X cuts described in this thesis - since all of these cuts pertain exclusively to LUX. This will be a critical consideration to any future researchers hoping to extend this work to LZ.

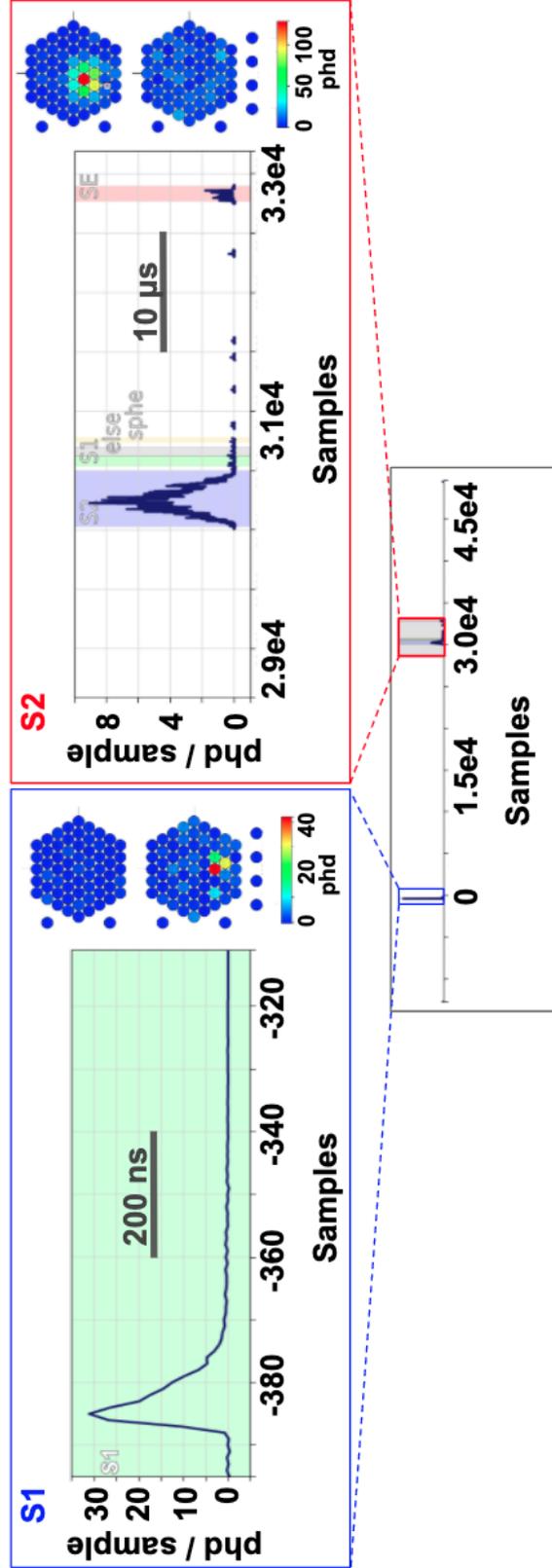


Figure 6.2 An example of a suspected gamma-X event in the Run04 data, showing the phd for S1 and S2 on the Y axes, and time in Samples on the X (1 Sample = 10 ns). The overall event summary can be seen in the grey box, located in the lower centre of the figure. The upper left box in blue is a close up of the S1 signal. The red box on the upper right is a close up of the S2 signal. The tight clustering of the S1 signal in the bottom PMT array and long drift time are both typical features of a gamma-X event.

6.2.2 Properties

In order to classify an event as a gamma-X, it is necessary to identify observables which differ from SS events. An example of a suspected gamma-X event observed in LUX Run04 data is displayed in Figure 6.2. Taken from Visualux (a visualisation tool for event waveforms) [113], this figure brings attention to features in the signal which would be typical of a gamma-X event.

The first identifying feature is the depth of the event, with the energy deposition appearing very close to the cathode. This is consistent with the MFP of a gamma-ray in LXe being quite short (Figure 4.5). For the backgrounds which produce high energy gamma-rays near the RFR, the MFP extends up to² 7.8 cm. As such the expected position of the detected S2 signal would be quite close to the RFR.

The second feature to note is the S1 photo-hit-pattern in the PMTs. Specifically that photons are primarily detected in the bottom PMT array, and are tightly clustered into a few PMTs. This is expected for a gamma-X event since one of the depositions would have had to occur below the cathode, with the other occurring a short distance above. The initial scatter then would see the S1 signal largely contained by a few PMTs in the bottom array. The subsequent S1, occurring above the cathode, would be more spread out. But since it is the combined S1 that is measured, we still expect a high proportion of this overall signal to be tightly clustered. It is these two features that we would expect to see in a gamma-X event, and thus the features which are typically used to distinguish gamma-X events from SS events.

6.2.3 Sources

This section will discuss potential sources of gamma-X events. This discussion will facilitate an understanding of which background sources need to be focused on in the full MC production of simulated gamma-X events.

As explained in Chapter 3, LUX's external shielding prevents all photons originating outside of the detector from reaching the fiducialised LXe. This leaves only radiogenic impurities in LUX's own internal components, and LXe contaminants as background sources capable of creating a gamma-X event. For the most part, these backgrounds are ER events³. These potential gamma-X sources can be categorised based on the detector volume they inhabit.

²7.8 cm for a 1460.82 keV photon emitted by a ⁴⁰K decay

³Notably alpha particles are present too, however they typically scatter in an energy range far beyond the region of interest for a WIMP search.

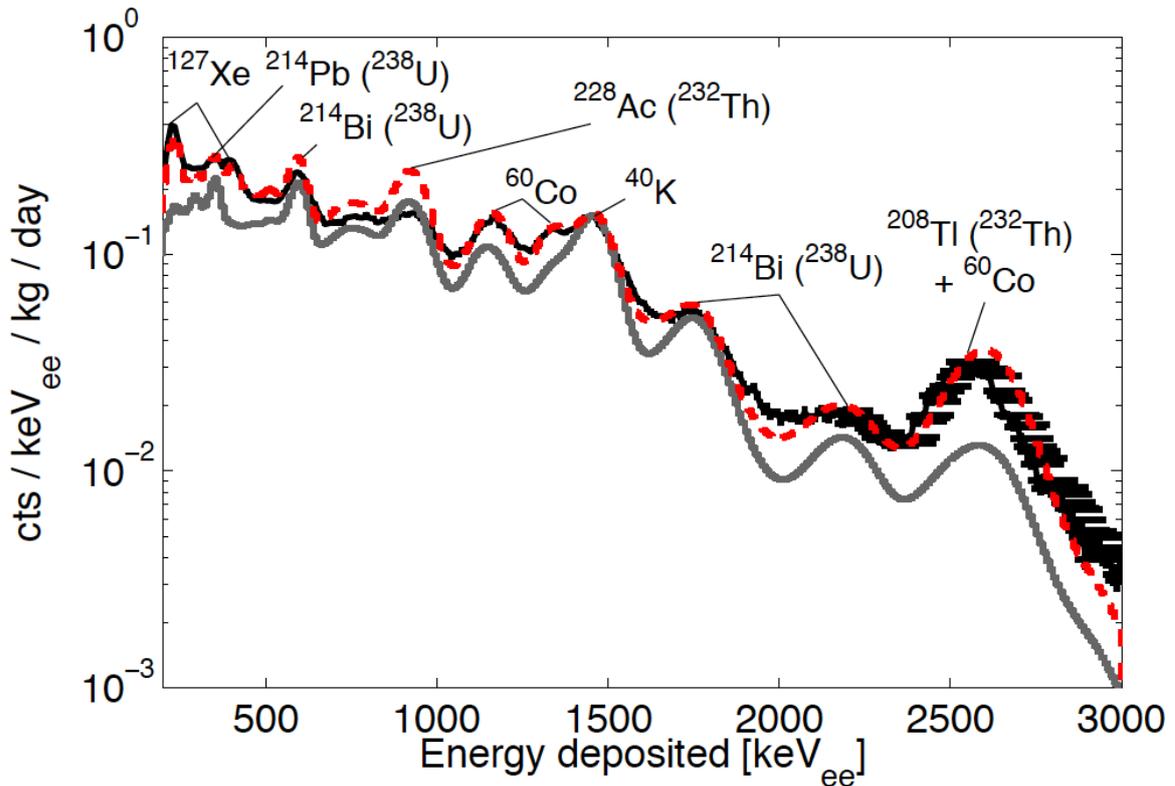


Figure 6.3 Measured gamma-ray spectrum in LUX during Run03 (black), with labels on peaks to identify sources [94]. This analysis used a 225 kg fiducial volume definition by not applying a radial cut and removing the top and bottom 2 cm of the drift region. Data includes both SS and MS events. Event energies are reconstructed from the combination of S1 and S2 signals. Horizontal error bars are shown, representing systematic uncertainties in energy reconstruction for high-energy events. Two simulation spectra are shown for comparison. A spectrum based on positive counting measurements alone is shown with a solid grey line. The spectrum with best-fit scaling for ^{238}U , ^{232}Th , ^{40}K , and ^{60}Co decays, with independent rates in top, bottom, and side regions of the detector, is shown as red dashed. Fitting was performed for energies above 500 keV_{ee} . Energies below 500 keV_{ee} were found to have the continued agreement between gamma-ray spectra and measured data. This spectrum has a lower bound at 200 keV_{ee} .

Broadly speaking these volumes are: the LXe, detector surfaces (e.g. TPC PTFE panels), detector components (e.g. PMT arrays), and the detector vessel (e.g. Ti ICV). The gamma-ray emission spectra of many of these sources can be seen in Figure 6.3. This figure shows the gamma-ray background spectrum from LUX Run03 [94], consisting of many isotopes which will be discussed in the following paragraphs.

Gamma-rays originating from sources within the bottom PMT array are a particular focus for this chapter. This is because the PMTs, and the structure which houses them, contain several radiogenic isotopes. The bottom PMT array is of particular concern due to its proximity to the RFR. This will be explored further in Section 6.5.4, where simulations indicated this was the primary source of gamma-X events for LUX.

The majority of gamma-rays emanating from radiogenic sources within the PMT arrays have an energy range of ~ 100 keV to a few MeV. The origin of these events primarily consists of naturally occurring isotopes present in the stainless steel structure of the PMT array, and the housing of the PMTs themselves. These sources include: ^{238}U , ^{232}Th and ^{40}K . The cosmogenic activation of metallic components in the PMTs (prior to LUX being moved underground), also leads to additional gamma-rays from the long lived ^{60}Co ($T_{1/2} \sim 5$ years).

In fact, ^{238}U , ^{232}Th and ^{40}K , are present in all detector materials, and ^{60}Co in any Cu or stainless steel components (for instance the field shaping rings). The potential for these isotopes to contribute to the gamma-X abundance is explored on a volume by volume basis in Section 6.5.4.

ER events can also be generated by the decay of radiogenic impurities within the Xe itself. Unlike backgrounds from detector components, these isotopes are distributed homogeneously within the LXe; having been mixed by convection. As such, a fiducial cut will not be effective in reducing their significance. Of the radiogenic isotopes present in the LXe are several activated products such as: ^{125}Xe , ^{127}Xe , ^{133}Xe , ^{125}I , $^{129\text{m}}\text{Xe}$ and, $^{131\text{m}}\text{Xe}$. However, each of these sources has a short enough half life (ranging from 17 hours for ^{125}Xe to 60 days for ^{125}I), that their abundance from cosmogenic activation will have decayed below a detectable level before the beginning of Run04. These isotopes can also be created by the DD calibration process. However time is taken between DD calibrations and Run04 to ensure these isotopes do not contribute to the background⁴. Additional contaminants from the Xe production process are also present in the LXe. These include: ^{85}Kr and ^{39}Ar . These sources persist in

⁴The higher rate of production of $^{129\text{m}}\text{Xe}$ and $^{131\text{m}}\text{Xe}$ means that they still need to be considered.

small amounts in LUX's Xe, but are reduced below a detectable level by the Xe filtration procedure described in [111].

A larger concern to the production of gamma-X events is the presence of the long lived radon isotopes ^{222}Rn and ^{220}Rn which emanates from detector components into the LXe. The decay products included in these series contains several high energy gamma-rays which can create gamma-X events. Their presence in the LXe means that these decay products can be produced very close to, or within the RFR, increasing their likelihood to create a gamma-X event.

6.3 History of gamma-X cuts in LUX

6.3.1 Background

Prior to the Run04 EFT WIMP search, two gamma-X cuts were developed by members of the LUX collaboration, which will be described in this section. One was developed for the standard WIMP search, and the other for double electron capture analysis. In addition, a third cut will be described which was used by the Run04 EFT search. This final cut was not designed for gamma-X removal, but the parameter space it operated in meant it was highly effective in the endeavour. The purpose of this section is to both: explain the cuts which will later be used to benchmark a new gamma-X cut against, and to introduce the reader to some of the considerations which go into gamma-X cut development. Note that at the time that these cuts were developed no dataset of real or simulated gamma-X events existed. As such they have not been previously assessed in their capacity to remove said events and were judged only on their calibration data efficiency.

6.3.2 WIMP search gamma-X cut

This cut was developed by Brian Lenardo, and operates in the parameter space of S1, and S1 cluster size (CS_{S1}) [114]. This CS_{S1} parameter can be considered as the weighted average radius of S1 light in the bottom PMT array as defined by [114]:

$$CS_{S1} = \frac{\sum_{i=0}^n \alpha_i \sqrt{(x_i - \bar{x})^2 + (y_i - \bar{y})^2}}{\sum_{i=0}^n \alpha_i} \quad (6.1)$$

where x_i and y_i describe the location of each PMT_i in the bottom array, α_i is the number S1 photons detected (phd) seen in each PMT, n is the number of PMTs in the bottom array, and

\bar{x} and \bar{y} are weighted means as defined by:

$$\bar{x} = \frac{\sum_{i=0}^n \alpha_i \cdot x_i}{\sum_{i=0}^n \alpha_i} \quad (6.2)$$

$$\bar{y} = \frac{\sum_{i=0}^n \alpha_i \cdot y_i}{\sum_{i=0}^n \alpha_i} \quad (6.3)$$

This cut is based on the observation that sub-cathode events will be more tightly clustered in the bottom PMT array. As explained in Section 6.2.2, this is thought to arise due to LUX's geometry. Since gamma-X events are assumed to be originating from the bottom PMT array this ought to be a useful parameter for gamma-X identification.

In order to specify a cut in this parameter space, tritium calibration data was used. Tritium ($T_{1/2} = 12.3$ years) is a single beta emitting source that produces ER events with a mean energy of 5.6 keV, with 64.2% of them occurring between 1 and 8 keV, and a maximum energy of ~ 20 keV. As such it is employed to characterise the ER response of dual-phase Xe TPCs. The decision to base the cut on tritium data was due to tritium's lower probability of causing a MS compared to other calibration sources. This allowed the cut to be tuned for maximal data acceptance. As illustrated in Figure 6.4 the cut was defined by selecting 20 $S1$ bins in the tritium data and specifying an 80% acceptance point based on the CS_{S1} of each bin. An exponential curve was then fit to these points which defines the cut. The specific curve shown on Figure 6.4 defines this specific gamma-X cut which was used in the Run04 WIMP search. The same methodology was applied to the Run03 data with a slightly different result. The tritium events chosen were all events occurring deep within the fiducial region ($r < 5$ cm) and deep in the detector, near the cathode⁵ ($200 \mu\text{s} < \text{drift time} < 300 \mu\text{s}$). When this cut was applied to the full tritium dataset, it had a total acceptance of 93.3% within the fiducial region and 70.3% below the fiducial region. When applied to the $^{83\text{m}}\text{Kr}$ data the total acceptance within the fiducial went up to 99.2%, and up to 80.0% below the fiducial cut [114].

6.3.3 Double electron capture gamma-X cut

Two-neutrino double electron capture is a decay mode of an atomic nucleus, which was recently observed for the first time in ^{124}Xe [115]. It is a rare process with an expected half-life higher than the age of the universe by several orders of magnitude. The precise value of this half-life is an important observable for nuclear structure models, and its measurement

⁵The drift time for electrons from the cathode is $\sim 310 \mu\text{s}$

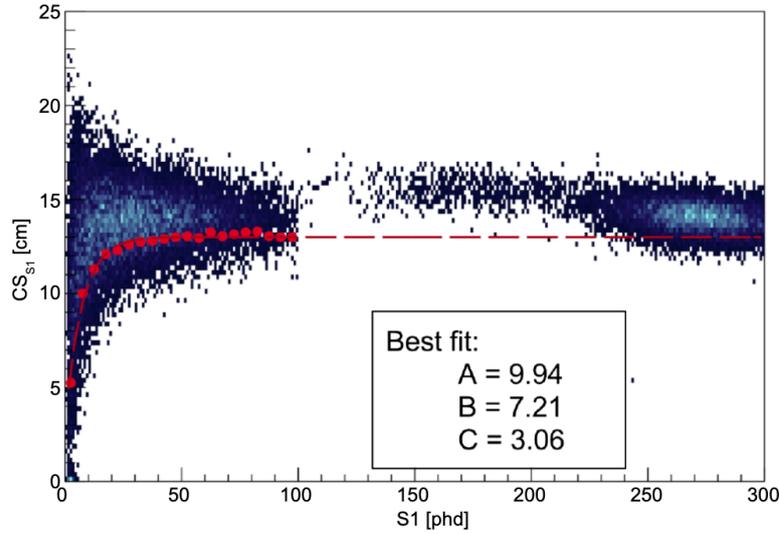


Figure 6.4 Definition of the CS_{S1} gamma-X cut used by LUX Run04 WIMP search [114]. The histogram shows the distribution of cluster sizes in data from a tritium calibration at low energies and a ^{83m}Kr injection at higher energies. The 20 80% acceptance level points are also given (red dot). The dashed red line is an empirical fit to these points which defines the cut. The curve takes the form $y = A(1 - e^{-\frac{x}{B}}) + C$. Events falling below this line will be rejected as gamma-X-like.

represents a meaningful step in the effort to discover the precise neutrino mass. Given the sensitivity of LUX to such an observation, an analysis attempt was made using the LUX Run04 data. In support of this study, an independent gamma-X cut was developed by Alex Lindote [116].

This cut was built in the parameter space of $S1$ and $S1$ Max Peak Area Fraction ($MPAF_{S1}$), where $MPAF_{S1}$ is defined as:

$$MPAF_{S1} = \frac{MPA_{S1}}{S1} \quad (6.4)$$

where MPA_{S1} is the amount of $S1$ light which is detected in the single brightest PMT (PMT_{max}). The cut was developed in a similar fashion to the standard WIMP search gamma-X cut, in that calibration events were binned in MPA_{S1} vs $S1$ space as shown in Figure 6.5. However, there were several key differences. First, ^{14}C events were used instead of tritium. While both sources are beta emitters, ^{14}C decays are more energetic, producing higher energy responses in the TPC⁶ extending out to 156 keV [92]. Second, bins were specified as having a width of 50 phd with 99% data acceptance. Finally, each bin was turned into a cut with a linear interpolation between acceptance points rather than with an empirical curve. Notably,

⁶The double electron capture analysis has an ROI of 58.9 keV to 69.7 keV.

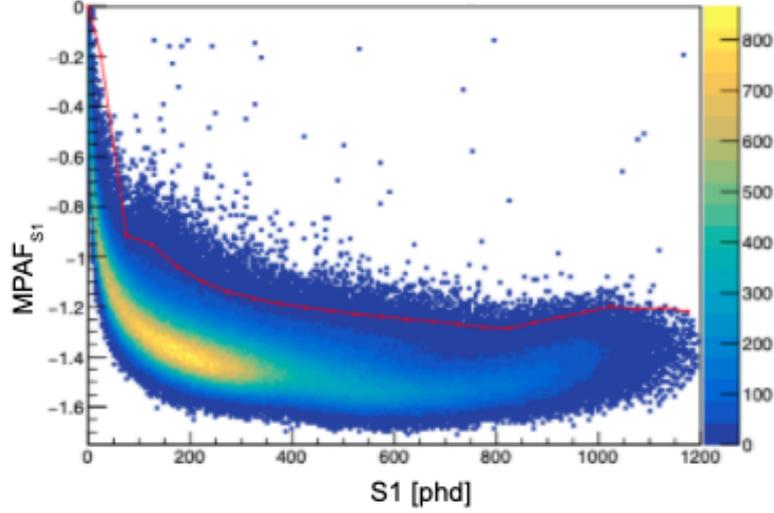


Figure 6.5 Gamma-X cut developed for the LUX double electron capture analysis. The heat map shows ^{14}C calibration data in $(S1, MPA_{S1})$ space. The red line overlaying this data is the cut in question, which was developed with a linear interpolation of 99% data acceptance bins in $S1$. Bins are defined by a width of 50 phd. Events above the red line are classified as gamma-X-like by this cut and removed.

the full suite of cuts used for the double electron capture analysis includes a rejection of all NR events. As such, this gamma-X cut is not concerned with NR acceptance.

6.3.4 EFT search leakage cut

The MPA_{S1} leakage cut was designed by Shaun Alsum for the removal of events where a PMT dark count has been counted as part of the $S1$ signal. These “leakage” events derive their name from the tendency of these events to have a suppressed $S2/S1$ ratio causing them to “leak” from the ER band into the NR band.

Although this cut primarily operates in MPA_{S1} vs. $S1$ space:

$$MPA_{S1} > M_b(S1) + D(S1) \quad (6.5)$$

the functions M_b and D are parameterised into 24 different bins of Z from 8 cm to 50 cm:

$$M_b(S1) = \begin{cases} \Sigma a_i S1^i & S1 < 29 \\ \Sigma c_i S1^i & S1 > 29 \end{cases} \quad (6.6)$$

$$D(S1) = \begin{cases} \Sigma d_i S1^i & S1 < 35 \\ \Sigma e_i S1^i & 35 < S1 < 110 \\ \Sigma f_i S1^i & S1 > 110 \end{cases} \quad (6.7)$$

where the coefficients a_i , c_i , d_i , e_i , and f_i are defined based on which *bin* they occur in:

$$bin(Z) = \frac{Z - 8}{1.75} \quad (6.8)$$

where the precise values of the coefficients are defined in [117].

Strictly speaking, this cut was not for the removal of gamma-X events. However, due to the parameter space it operates in there is reason to think it will make for an effective gamma-X cut. As such it will be assessed along with the others in order to help establish a benchmark for a new gamma-X cut.

6.3.5 Similarities, differences, and expected performance

Each of the MPA_{S1} leakage cut, the $MPAF_{S1}$ gamma-X cut, and the CS_{S1} gamma-X cut rely on energy depositions below the cathode having a higher proportion of $S1$ photons detected in a tighter radius on the bottom PMT array. For the cuts which rely on MPA_{S1} however, this is measured by looking at the ratio of $S1$ light in PMT_{max} , with the overall $S1$ signal. Intuitively, this cut relies on the $S1$ photons from the initial energy deposition of a gamma-X event being disproportionately detected by a single PMT in the bottom array. This will presumably arise naturally as a result of the scatter occurring so close to the PMT in question. Nevertheless, there is no apriori reason to believe a gamma-X cut based on MPA_{S1} will be any more or less effective than one based on CS_{S1} . Subsequently, interest exists in testing these cuts on gamma-X-like and SS-like datasets.

Although the rationale behind all of these cuts is similar, given the rather different implementation there is no guarantee they will behave similarly. As such, interest exists as to which cut is more effective; particularly for benchmarking the performance of a new gamma-X cut. Making such a decision based on their performance on simulated gamma-X events, and Run04 calibration data will set a standard for a new gamma-X cut. This usage is explored in Chapter 7.

Nevertheless, since gamma-X and SS events cannot be distinguished with certainty when looking at real data, a solely data-driven approach for gamma-X cut validation is not possi-

ble. This leaves accurate simulations as an essential step in benchmarking cut performance. However, since all of these cuts are based on parameters related to precise S1 photo-hit-pattern, the simulation effort is complicated. This is because the normal simulation package of LUX_BACCARAT and LUX_LLAMA does not determine these parameters. Any parameter requiring a full simulation of S1 light will suffer from this drawback since neither LUX_BACCARAT nor LUX_LLAMA simulates full photon paths by default. This feature is typically excluded, to improve simulation run-time. The result is the failure to provide GEANT4 all photon information to propagate full photon tracks. Therefore parameters like CS_{S1} and MPA_{S1} are simply not calculated, and the necessary photonic information is not available. The best which LUX_LLAMA can offer is the photon yield of energy deposition sites after event clustering has occurred. How this information is handled to determine S1 photo-hit-patterns on the PMT arrays is described in Section 6.5.1.

6.4 Run04 EFT data analysis cuts

In order to assess the gamma-X cuts described in Section 6.3 various data selections were made from the LUX Run04 data. The data selections used relied on cuts developed by several LUX analysts for the Run04 EFT search [117]. These cuts include some of what was presented in Section 3.7, but in many cases build upon it. In this section, these cuts will be described. Where cuts are not described in detail, a full description will be made available in [95]. Similarly to the leakage cut described in Section 6.3.4 many of these cuts use more than two parameters, which makes a simple illustration difficult. However, in cases where a cut operates in a 2-dimensional parameter space, or where a bespoke parameter can be created for the purpose of illustration, a plot is displayed in Figure 6.6. In this figure, each cut has been individually applied to a combination of tritium and ^{14}C data. This data was chosen since both sources have a well-understood recoil spectrum, are homogeneously distributed in the LXe, and combine to span the full energy range of the EFT search [82]. It should be noted that these cuts were only what was being used by the LUX EFT analysis as of February 2020. However, as of the moment this was written, cuts are still under development. As such, what is presented here should not be taken as being representative of the data selection eventually used in the final LUX EFT analysis [95], only as what cuts were used as a basis to inform data selections for LUX analysis in this thesis.

The first cut we will describe is the ROI cut. This cut is applied to both S1, and S2. It was defined as $S1 < 300$ phd, and S2 between 200 phd to 36000 phd. This cut is visualised in Figure 6.6a. The lower bound $S1 > 2$ phd which can be seen in this figure represents the 2

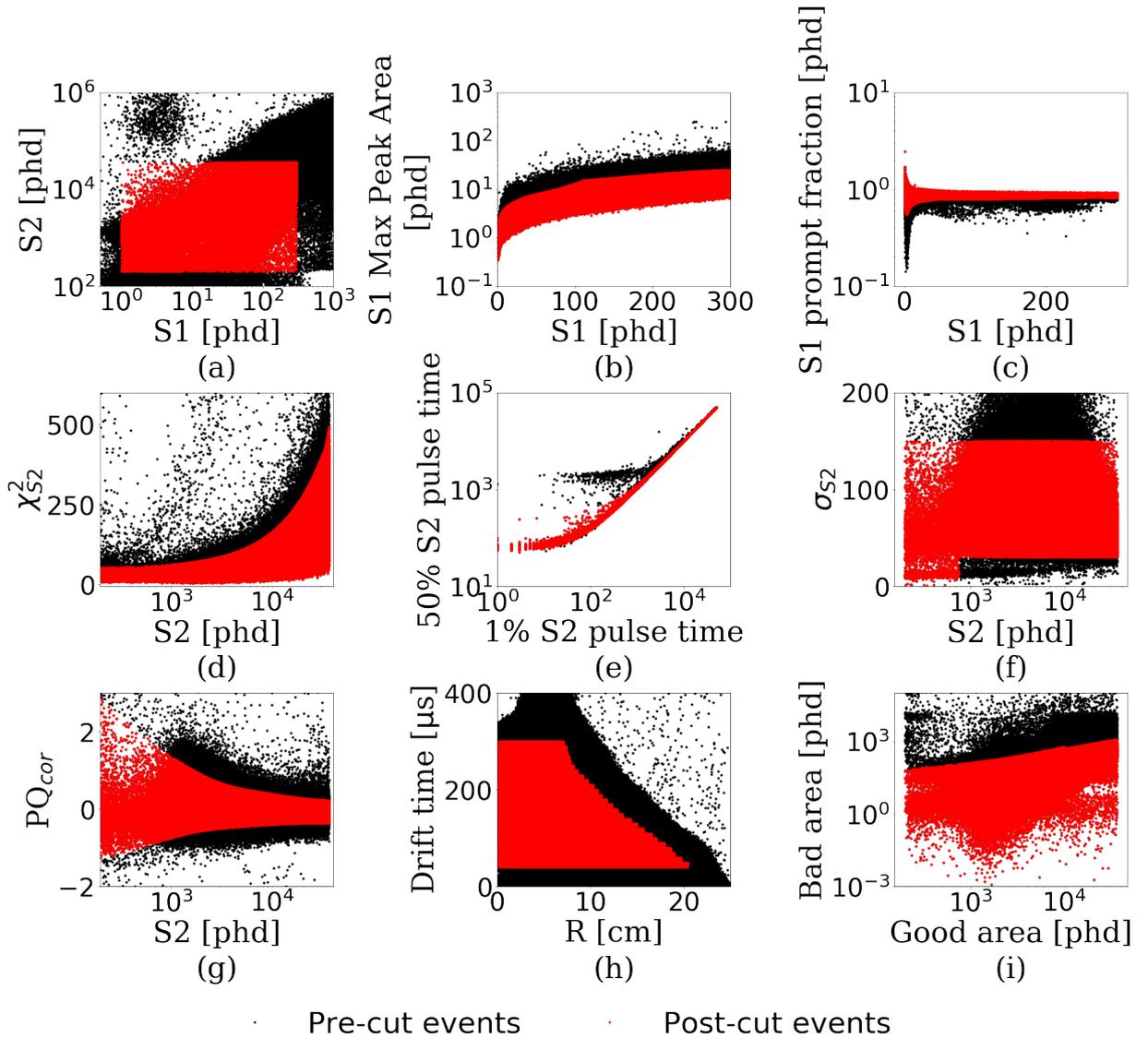


Figure 6.6 Various EFT data analysis cuts applied to a combination of tritium and ^{14}C data⁷. Events that pass a given cut are shown in red, removed events are in black. (a) shows the ROI cut, (b) shows the MPA_{S1} cut, (c) shows the $S1$ prompt fraction cut, (d) shows the $S2$ χ^2 cut, (e) shows the $S2$ pulse width cut, (f) shows the $S2$ σ cut, (g) shows the $S2$ shape cut, (h) shows the fiducial cut, and (i) shows the bad area cut. Excluded from this figure is the $S1$ partner cut.

fold coincidence which is required for signal detection. Another cut that is placed on the $S1$ signal is known as the $S1$ partner cut. This cut removes events which occur when only PMTs on the same readout channel observe an $S1$. The 2 fold coincidence and the $S1$ partner cut are needed to remove events where PMT noise has been classified as an $S1$ which occurs in conjunction with an unpaired $S2$ signal.

To eliminate gas events where the $S1$ and $S2$ are both recorded as a single $S1$ event, an $S1$ prompt fraction cut is applied. This cut examines the $S1$ pulse and removes any event which has a low proportion of photons appearing in the first 10% of the full $S1$ duration. This cut is illustrated in Figure 6.6c.

A cut used to help remove MS events which are separated in R but not in Z was also implemented. The cut works by looking at the χ^2 value of the position reconstruction algorithm. If the value is too high the event is removed. This cut is visualised in Figure 6.6d.

Another cut which aims to remove MS events which are misclassified as SS is the $S2$ pulse width cut. This cut uses a bespoke parameter called PW , which is the difference between the time for 1% of the $S2$ pulse to be detected, and the time for 50% of the $S2$ pulse to be detected. If PW falls outside the band shown in Figure 6.6e then the event is excluded.

A third cut used to remove events which appear to have multiple $S2$ pulses classified as a single pulse is the Gaussian Sigma cut. This cut functions by examining the shape of the $S2$ pulse, which ought to follow a Gaussian distribution under normal circumstances. If the standard deviation of this distribution (σ_{S2}) is outside of the range indicated in Figure 6.6f, then it is excluded.

The final cut concerned with MS events misclassified as SS is the $S2$ shape cut. This cut uses another bespoke parameter called PQ , which is defined as:

$$PQ = \frac{PW}{\sigma_{S2}} \quad (6.9)$$

This parameter PQ is used to define another bespoke parameter called PQ_{DTM} , via a function which depends on drift time (DT) and is available in [117]. The difference between PQ and PQ_{DTM} ($PQ_{cor} = PQ - PQ_{DTM}$) is then used to define the cut which can be seen in Figure 6.6g.

A fiducial cut is also used to reduce the rate of wall, and other background events orig-

inating from beyond the LXe. It removes all events which have a reconstructed position within 3 cm of the wall, and events outside of $DT \in [40 \mu\text{s}, 300 \mu\text{s}]$. This cut is visualised in Figure 6.6h. This figure also clearly demonstrates the effect of field non-uniformities described in Section 3.2.3.

The final cut which was implemented is the so-called Bad Area cut. This cut removes events with an overabundance of pulses which are not clearly identified as the $S1$ or $S2$. These events need to be removed as they are more likely to misidentified an anomalous pulse (e.g. an electron train) as an $S1$ or $S2$. This cut relies on two parameters: good area ($S1 + S2$), and bad area (full event - $S1 - S2$). Events are removed if the ratio of good area to bad area is too low. This selection is visualised in Figure 6.6i.

The only LUX Run04 data analysis cut which was excluded from this description is the MPA_{S1} leakage cut. This cut operates in the 3D parameter space of, $S1$, MPA_{S1} , and Z , and is designed to remove events which originate from outside the TPC. The reason this cut is not described in detail here is due to its relevance to gamma-X cut benchmarking. As such it is already described in detail in Section 6.3.4.

It should be noted that this is not the final list of cuts used by the EFT analysis, only what were the confirmed cuts as of February 2020. At the moment this was written other cuts were being developed that are not included in this list (e.g. the gamma-X cut amongst others). However, since the main purpose of the analysis conducted for LUX in this thesis is the development of a gamma-X cut for the Run04 EFT search, it was necessary to stop adding cuts to this suite so that a final analysis of the gamma-X cut could be produced. As such, any dataset with these cuts applied should not be considered the final dataset used by the LUX Run04 EFT analysis.

6.5 Gamma-X event simulation

6.5.1 Simulation process

This section will describe the full MC process used to simulate gamma-X events. The process relied upon LUX's standard simulation chain of LUX_BACCARAT and LUX_LLAMA, as illustrated in Figure 6.7. However, full photon paths resulting from the $S1$ light were required, which is not a standard output from these simulations. Rather than using the

full light transport supported by GEANT4 to simulate paths, this stage was replaced by a ray-tracing package known as NEST_light.

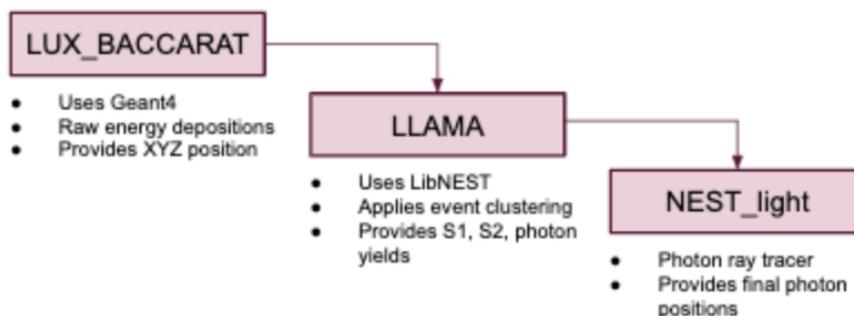


Figure 6.7 Simulation chain used for producing gamma-X events.

The description of this simulation process begins with the LUX_BACCARAT package described in Section 3.5.2 [90]. LUX_BACCARAT is the LUX version of BACCARAT as described in Chapter 3. The only difference is the detector geometry which is called. At the LUX_BACCARAT stage of simulations, the primary outputs of interest were: the XYZ position of energy depositions, and recoil energy. A useful feature of BACCARAT is its ability to simulate any background source in any detector volume. But for the simulation of gamma-X events only ^{60}Co , ^{40}K , ^{238}U , and ^{232}Th decays from the bottom PMT array were used. As will be shown in Section 6.5.4, the choice to focus on these sources was based on their likelihood to produce gamma-X events. Two reasons exist for this increased likelihood with respect to other sources. First, is that all of these sources produce a high energy gamma-ray, capable of scattering multiple times [93]. The second, is due to the position of the bottom PMT array with respect to the cathode. This results in upward moving gamma-rays emitted from the bottom PMT tending to have their first scatter in the (RFR).

BACCARAT boasts several features which simplified this analysis. First, its component-centric approach to simulations allowed for easy distribution of radioactive sources throughout detector components. It also simplifies the decay position, timing, and magnitude determination. However, certain detector specific parameters such as $S1$, and $S2$ were required that BACCARAT does not automatically output. This is due to the manner in which NEST is integrated, which does not calculate the precise photon and electron yields for energy depositions. One package developed by the LUX collaboration to handle this is known as LUX_LLAMA. LUX_LLAMA is the LUX version of LLAMA and is used by the entire collaboration to ensure a standardised analysis approach by collaboration members.

LUX_LLAMA is used to determine the clustered energy deposition magnitudes and locations based on LUX_BACCARAT outputs and integrates them with NEST. In doing so it is able to output several parameters of interest for gamma-X simulation studies. Specifically, these parameters were: Z , $S1$, $S2$, and photon yield. It does not however give the parameters of CS_{S1} , or MPA_{S1} . To determine these requires a full simulation of all scintillation photons generated at the interaction site. Such a simulation is computationally heavy if following the standard step-wise approach adopted by GEANT4. Subsequently to generate a statistically significant number of events is not feasible with LUX_BACCARAT, and LUX_LLAMA alone. Fortunately, a solution exists by means of a package known as NEST_light.

NEST_light is a ray-tracing package based on the NEST code. NEST_light calculates photon paths by choosing an initial random direction for the photon and then propagates it in a straight line. Photon paths only change direction when reflecting off of detector boundaries, or refracting through the phase boundary. This method proves to be significantly quicker whilst maintaining consistency with data and full MC results [118].

By following this full simulation chain, the position of photons reaching the PMT arrays was available. Combining this with the PMT geometry from LUX_BACCARAT, calculation of simulated CS_{S1} and MPA_{S1} parameters was facilitated. This process began by binning photons by the PMT which they had reached. From there photon counts for individual PMTs could be added into a combined $S1$, and the PMT_{max} could be identified⁸. This provided the last piece of information required to calculate the features CS_{S1} , MPA_{S1} used in Eq. 6.1, and Eq. 6.4.

6.5.2 Simulation tuning

To ensure this simulation chain faithfully reproduced real data, ^{83m}Kr events were simulated and compared to real Run04 ^{83m}Kr events. The focus of this comparison was on the three features used by the gamma-X cuts described in Section 6.3: $S1$, CS_{S1} , and $MPAF_{S1}$. It was observed that the agreement between simulations and data could be improved by tuning the output simulations to match the data. Once complete, this tuning was applied to simulations of another calibration source and compared for validation.

Simulations were tuned by first dividing events into 20 equal bins based on their Z position. This is represented in Figure 6.8, with Run04 data shown as a red-black heat map,

⁸Note that PMT efficiency is not taken into account. This should not be an issue though since the final parameters of interest CS_{S1} and $MPAF_{S1}$ are independent from PMT efficiency assuming individual PMTs have a similar efficiency to one another.

and the mean of each simulated bin in turquoise, along with a 1σ band. The figure is divided into six plots, displayed in a 3×2 matrix. Each row deals with a different feature, and the columns correspond to untuned, and tuned simulations respectively. The data on the left and right plots of each row is the same. But in the left column, the simulated bins and bands are built from untuned simulations. Whereas on the right, the bins and bands have been tuned to the data.

Simulations were adjusted on a bin by bin basis. In each bin, the mean of the simulations was compared with the mean of data. Tuning was then done by multiplying each simulated event by the ratio between the data and simulated means. For some features, this tuning method was slightly modified. For instance, to improve the fit for $S1$, simulations were modified by subtracting the difference of means for each event instead of multiplying by a ratio of means.

The goal of this effort was to tune simulations so that they better represented the data for $^{83\text{m}}\text{Kr}$. Based on the results shown in Figure 6.8 simulations have been improved to better reflect reality. The reason for this disagreement between simulations and data is beyond the scope of this thesis. However, one possibility is due to the way parameters are extracted from `NEST_light`, whereby individual PMT efficiencies are not taken into account (only the average). Another is an early implementation of an Agglomerative Hierarchical Clustering algorithm which LLAMA was using at the time being in a beta stage of development.

6.5.3 Simulation chain validation

Once a method for tuning simulations based on a comparison with Run04 $^{83\text{m}}\text{Kr}$ was established, simulations had to be validated with a new dataset. For this tritium calibration data was used. As such tritium events were simulated via the same method described in Section 6.5.1.

The comparison of these simulated results after tuning to data can be seen in Figure 6.9. This figure contains six plots comparing simulated $S1$, CS_{S1} , and $MPAF_{S1}$ to data for tritium. Each of these comparisons is performed over two dimensions for the three parameters of interest. To facilitate the comparison of data and simulated heat maps in these 2-dimensional parameter spaces, two types of plot are used. The first is a heat map of data events (red-black), overlaid with a contour plot of the simulated event heat map (light blue). The second is the opposite; a heat map of simulated events (blue-black), overlaid with a contour plot of the data event heat map (red).

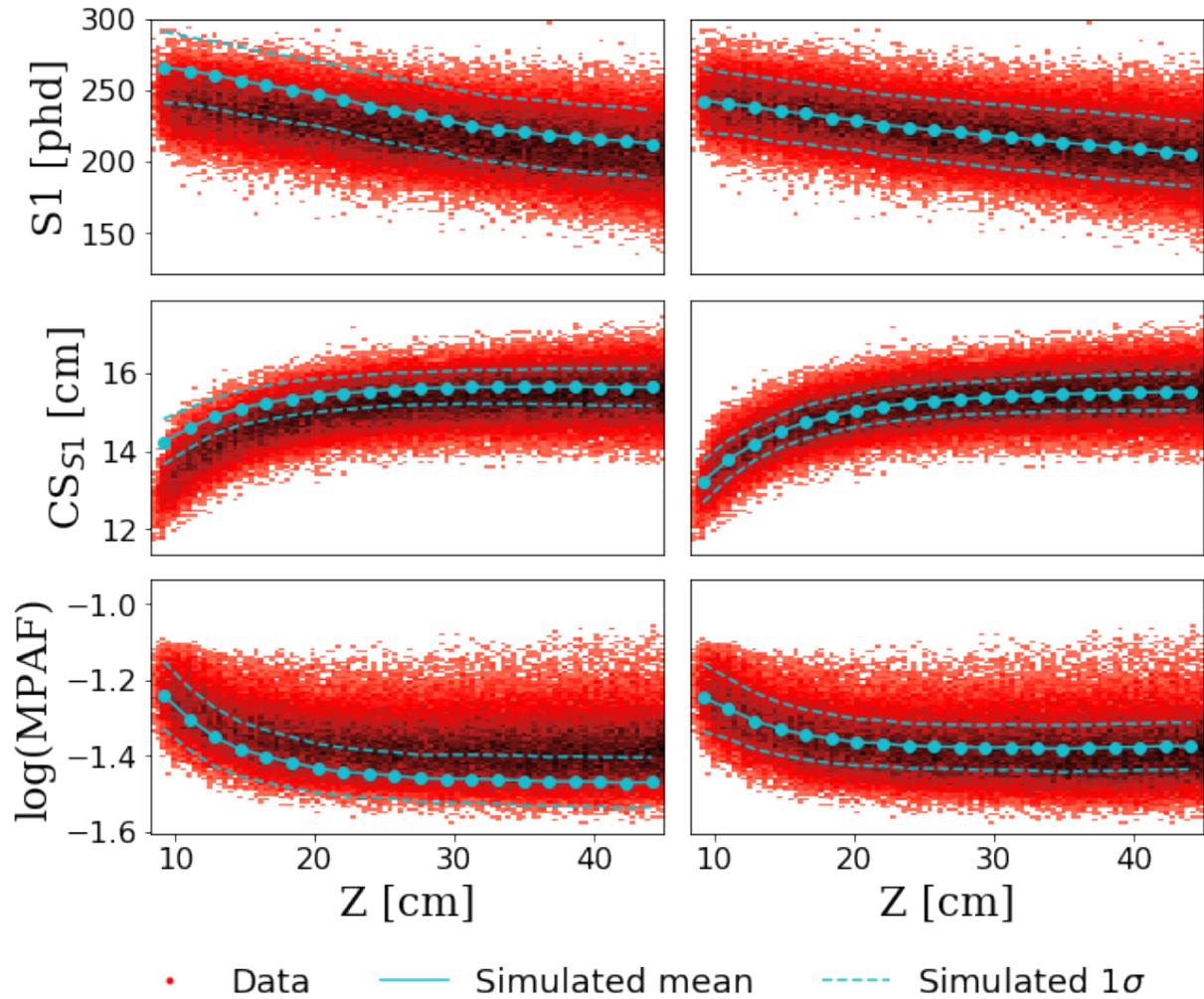


Figure 6.8 Tuning of raw simulations based on comparison with Run04 $^{83\text{m}}\text{Kr}$ data. For each variable listed on the vertical axis, data and simulations were compared in 20 different Z bins. The mean of each simulated bin is shown in turquoise, along with a 1σ band. Plots on the left show the simulated bands before tuning overlaid on data. Plots on the right show simulated bands after tuning.

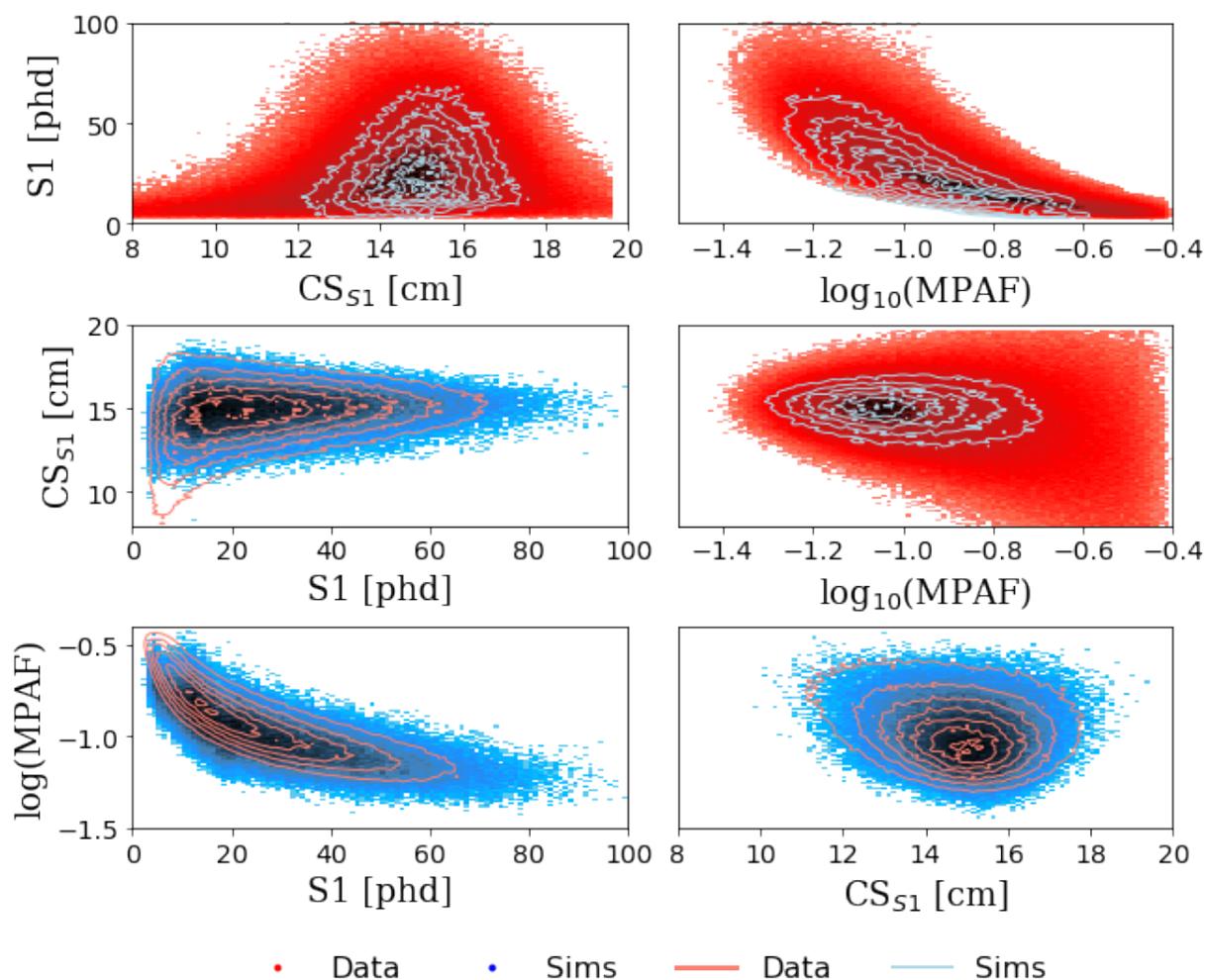


Figure 6.9 A comparison of simulated and real $^{83\text{m}}\text{Kr}$ data for LUX Run04. To ease visual comparison all plots display one dataset as a heatmap, with the other represented as a contour plot overlaid on top. Regardless of whether data or simulations are represented by the heatmap - simulations are always shown in blue, and data in red. The top left plot displays $S1$ vs CS_{S1} , the top right displays $S1$ vs $\log(\text{MPAF})$, and the middle right plot displays CS_{S1} vs $\log(\text{MPAF})$. The remaining three plots show the same datasets, but with the axes flipped, and data as a contour plot laid over the simulations.

Notably simulated events with $S1 < \sim 10$ phd still do not fully match data. This is highlighted in the $S1$ vs CS_{S1} plot of Figure 6.9. There we see the contours of data having a much wider CS_{S1} range for low $S1$ events than what simulations can reproduce. This wider range for $MPAF_{S1}$ and CS_{S1} observed in data for low $S1$ events is likely due to PMT dark count which can enhance the phd count in the PMT_{\max} . This would result in a decreased CS_{S1} , and larger $MPAF_{S1}$. This effect would not be reproduced by simulations that assume ideal PMT behaviour. Such an effect would still be present in events with a larger $S1$. However, the larger statistics would minimise its impact.

The goal of this effort was to assess whether the simulations faithfully reproduced data. Specifically for the parameters used by the gamma-X cuts described in Section 6.3: $S1$, CS_{S1} , and $MPAF_{S1}$. Based on the results shown in Figure 6.9 this appears to be the case. Although low $S1$ events may be offset due to dark PMT counts not being incorporated into the simulation.

6.5.4 Source selection for gamma-X simulations

The production of a gamma-X event is a rare process. As such producing a large set of simulated gamma-X events will require an extremely large number of simulations. In order to streamline this process, it is worthwhile determining which background sources produce gamma-X events at the highest rate. In this section, we determine the contribution of different radiogenic sources to the gamma-X abundance. To achieve this, the simulated results of the Run04 ER background model were analysed. These events represented the ER background seen in the LUX detector and allows for the identification of both the background source and volume of origin for any event. Within these simulations, events which could potentially become a gamma-X were selected. Based on the volume with the largest rate of potential gamma-X events, a choice was made of which volume to use in the production of gamma-X simulations.

The Run04 ER background simulations were generated via a similar process described in Section 6.5.1. This involved using the standard packages of LUX_BACCARAT and LUX_LLAMA. A crucial difference between the Run04 simulations and the aforementioned process is that `NEST_light` was not applied in this case. The reason for this is that this study was interested only in energy deposit positions and energies, and not with $S1$ light hit patterns in the PMTs.

Potential gamma-X events were selected by examining each simulated event to see if it had energy depositions in both the RFR, and the FFR. A potential gamma-X event is selected based on the clustered MC truth XYZ position of the energy deposition. To be considered a potential gamma-X, the event must:

- Have one energy deposition cluster above the cathode which produces an $S1$ and an $S2$
- Have at least one energy deposition cluster below the cathode which produces an $S1$ signal

No other data quality cuts were applied to this selection. As such the values obtained would not be useful in determining an expected gamma-X rate. This will inform us of which volumes to focus simulation efforts on when attempting to maximise the fraction of gamma-X events which are produced.

The results of this analysis can be seen in Table 6.1, which shows the number of gamma-X events produced after a specified number of simulated events. Sources such as ^{238}U and ^{232}Th include all progeny and subsequent decays. As clearly demonstrated here, the choice to focus on radiogenic sources in the bottom PMT array is supported. The bottom PMT array backgrounds of ^{40}K , ^{60}Co , ^{238}U , and ^{232}Th are the most probable sources to produce gamma-X events. When these sources decay in the bottom PMT array the fraction of gamma-X events produced is $\sim 10^{-5}$.

The next most prodigious volume for potential gamma-X production is the walls. This consists of various radiogenic sources located in field shaping rings and the PTFE lining of the TPC. The majority of these sources however are too far from the cathode to be capable of producing a gamma-X event. Given the length of this volume, it is likely that the rate of wall sources close to the cathode would make a larger contribution to a gamma-X population if examined in isolation. Furthermore, the activity rate in the walls was found to be quite low by post decommissioning radioassay measurements. The PTFE was the component closest to the LXe, and of the isotopes in the PTFE the most abundant were: ^{40}K at 0.122 mBq/kg, ^{232}Th at 0.028 mBq/kg, and ^{238}U at 0.021 mBq/kg; which corresponds to $\sim 1 \times 10^3$ total decays. These activity levels were far below what were observed in the PMTs, which had: ^{40}K at 66 mBq/PMT, ^{232}Th at 2.7 mBq/PMT, ^{238}U at 9.5 mBq/PMT, and ^{60}Co 2.6 MBq/PMT; which corresponded to $\sim 1 \times 10^8$ total decays. Given this paired with the low chance of each decay causing a gamma-X (as indicated in Table 6.1), this volume does not represent the most suitable candidate to maximise gamma-X production.

Immediately behind “the walls”, is the detector component dubbed as “the vessel”. This volume consists of the OCV and ICV. After the walls, this volume is the next most likely to cause a gamma-X event. Although this analysis indicates that these sources are capable of leaving energy depositions in the LXe, this does not occur frequently enough to be a serious consideration. This was likely due to their distance of these volumes from the LXe, and the density of the walls in between. This means that in spite of the fact that the activity rates of the vessel were on the same order of magnitude as the PMTs, fiducialisation is expected to remove the bulk of these events from becoming a background concern.

After the vessel, the volume with the next highest fraction of potential gamma-X production is the LXe. Sources in this volume also will not make a substantial contribution to the gamma-X abundance above what is already produced by the bottom PMT array due to their relatively low activity rates which are on the order of $\mu\text{Bq/kg}$; resulting in $\sim 1 \times 10^4$ decays. As such it too can be disregarded as only making a minor contribution to the overall gamma-X background once considering the chance of these events to produce a gamma-X event.

The final volume to be examined was the top PMT array. Due to the distance between this volume and the cathode, gamma-rays emitted from radiogenic sources in the top PMT array are not expected to produce energy depositions in the RFR. Nevertheless, this volume was assessed due to the simplicity of extending the analysis on the existing simulations. Unsurprisingly, no events were seen to have an energy cluster in the RFR. Thus no events were capable of being a gamma-X.

The goal of this analysis was to determine which sources to focus on when trying to simulate gamma-X events in order to produce a realistic spectrum of simulated gamma-X events, in the shortest possible time frame. The result of this analysis supports the decision to focus on the bottom PMT backgrounds of ^{40}K , ^{60}Co , ^{238}U , and ^{232}Th for in order to produce a realistic gamma-X population with the maximum number of simulated gamma-X events in the minimal number of computational hours.

6.5.5 Gamma-X simulations from the bottom PMT array

Gamma-X events were simulated by producing as many ^{40}K , ^{60}Co , ^{238}U , and ^{232}Th decays as possible in the bottom PMT array. The result can be observed in Table 6.2. As shown a total of 6808 gamma-X events were produced from 4.15×10^8 decays. Gamma-X events were selected by looking for events which had multiple *S1* producing clusters above and

Table 6.1 Run04 ER background model analysed for potential gamma-X events.

Detector volume	Source	Total simulated events	No. potential gamma-X	Potential gamma-X fraction
LXe (FFR+RFR)	^{238}U	20617862	13	6.31×10^{-7}
	^{232}Th	460500000	355	7.71×10^{-7}
	$^{129\text{m}}\text{Xe}$	31518255	9	2.86×10^{-7}
	$^{131\text{m}}\text{Xe}$	18192334	1	5.50×10^{-8}
Walls (PTFE+FSRs)	^{238}U	55135056	105	1.90×10^{-6}
	^{232}Th	32528649	98	3.01×10^{-6}
	^{40}K	1063069	7	6.58×10^{-6}
	^{60}Co	3345999	23	6.87×10^{-6}
Top PMT array	^{238}U	752286	0	0.00
	^{232}Th	826399	0	0.00
	^{40}K	332127	0	0.00
	^{60}Co	6880459	0	0.00
Bottom PMT array	^{238}U	2228197	37	1.64×10^{-5}
	^{232}Th	6567089	108	1.65×10^{-5}
	^{40}K	1166700	21	1.80×10^{-5}
	^{60}Co	24118746	537	2.23×10^{-5}
Vessel	^{238}U	3054730	3	9.82×10^{-7}
	^{232}Th	3221490	2	6.21×10^{-7}
	^{40}K	1704990	1	5.87×10^{-7}

below the cathode, but only one $S2$ producing cluster above the cathode.

These simulations were used to assess the three cuts described in Section 6.3. The result of this analysis is described in Chapter 8. However, before this gamma-X cut benchmarking could commence a final simulation validation had to be performed on the distance between subsequent energy deposition clusters produced by the simulations. This final validation is discussed in Section 6.5.6.

Table 6.2 Gamma-X simulation results. Four background sources from the bottom PMT array (left column) were simulated. The decay rate per PMT is shown in the second column, the total number of events in the third, the number of days this represents in the fourth, and the final column gives the total number of gamma-X events produced.

	Activity [mBq/PMT]	Events simulated	Days simulated	Total simulated gamma-X events
^{40}K	66	1.21×10^8	170	2070
^{238}U	2.6	1.97×10^8	1970	2973
^{60}Co	2.7	4.71×10^7	3000	964
^{232}Th	9.5	5.00×10^7	1750	801

6.5.6 Comparison of multiple scatter simulations with data

Before proceeding with a search to find distinguishing features of a gamma-X event, these simulations had to be assessed as to whether they faithfully reproduce the vertical distance between energy deposition clusters. If simulations do not seem capable of correctly reproducing this feature, simulations may require further tuning.

Since it was not possible to directly compare simulated and real gamma-X events, a new category of MS was selected - the near-miss-gamma-X event (a.k.a. nearGX). The nearGX event was defined as a MS with one $S1$, and two $S2$ signals (a.k.a. 1 $S1$ 2 $S2$ MS) which occurs within 3.8 cm of the cathode⁹. These types of MS are of interest to gamma-X studies, as they are the closest events to a gamma-X which a dual-phase TPC is capable of identifying. These events are likely caused by gamma-rays emitted from the bottom PMT array which scatter twice in the FFR. The choice of 3.8 cm was made as this is the same distance as the space between the bottom shield grid, and the cathode. As such, they can be thought of as the kind of event which would have been a gamma-X if they had only occurred exactly 3.8 cm lower.

⁹3.8 cm for the lowest scatter, the higher scatter was free to have any value for Z .

This study consisted of two datasets: nearGX events from the LUX Run04 data, and simulated nearGX events from the bottom PMT background simulations described in Section 6.5.5.

Run04 nearGX data was acquired by applying all of the data analysis cuts described in Section 6.4 after removing all but the 1S1 2S2 data. Data selection involved applying the S1 cuts to the S1 signal as usual, but the S2 cuts were only applied to the upper S2 signal (i.e. the S2 furthest from the cathode). Furthermore, fiducialisation was based on the position of the upper S2, and the good area was extended to include both S2 pulses. As can be seen in Figure 6.10, this left 14 nearGX events at 3.8 cm above the cathode after all cuts.

The selection of the simulated nearGX events was performed on the bottom PMT background simulations described in Section 6.5.5. For consistency, the same cuts which were applied to the data were also applied to the simulations.

The total number of real nearGX events can be seen in Figure 6.10a, and a comparison of the Z separation of the simulated and real nearGX events can be seen in Figure 6.10b. Figure 6.10b shows six lines, each of which represents the mean vertical separation in S2 deposits for different nearGX selections. The four simulated decays of ^{40}K , ^{60}Co , ^{238}U , and ^{232}Th are shown separately in light blue, orange, light green, and red dashed lines respectively. These simulated results are scaled by their relative activity in LUX Run04 (based on post decommissioning radio assays of detector components) and combined. The combined result is shown in green, with a shaded region around it representing the 1σ band. The final line represents events from the Run04 data and is shown in blue. It too has a shaded 1σ region.

A comparison of the mean Z separation for simulations and data shows agreement within 1σ . As indicated by Figure 6.10, as the number of data events increases this agreement appears to strengthen. While the argument could be made that setting the nearGX selection threshold of 3.8 cm to a lower value would make events more gamma-X-like, such a decision for the purposes of validation would be undermined by the lower number of events available in the data. As shown in Figure 6.10a, below a threshold of 2.4 cm there are less than 10 events. As such Figure 6.10b does not consider comparisons with a selection threshold below this level.

For the time being it appears data and simulations are in close enough agreement that further tuning is not required to ensure they are capable of creating gamma-X events. However, this conclusion is limited by low statistics. As an aside, it is worth noting that future

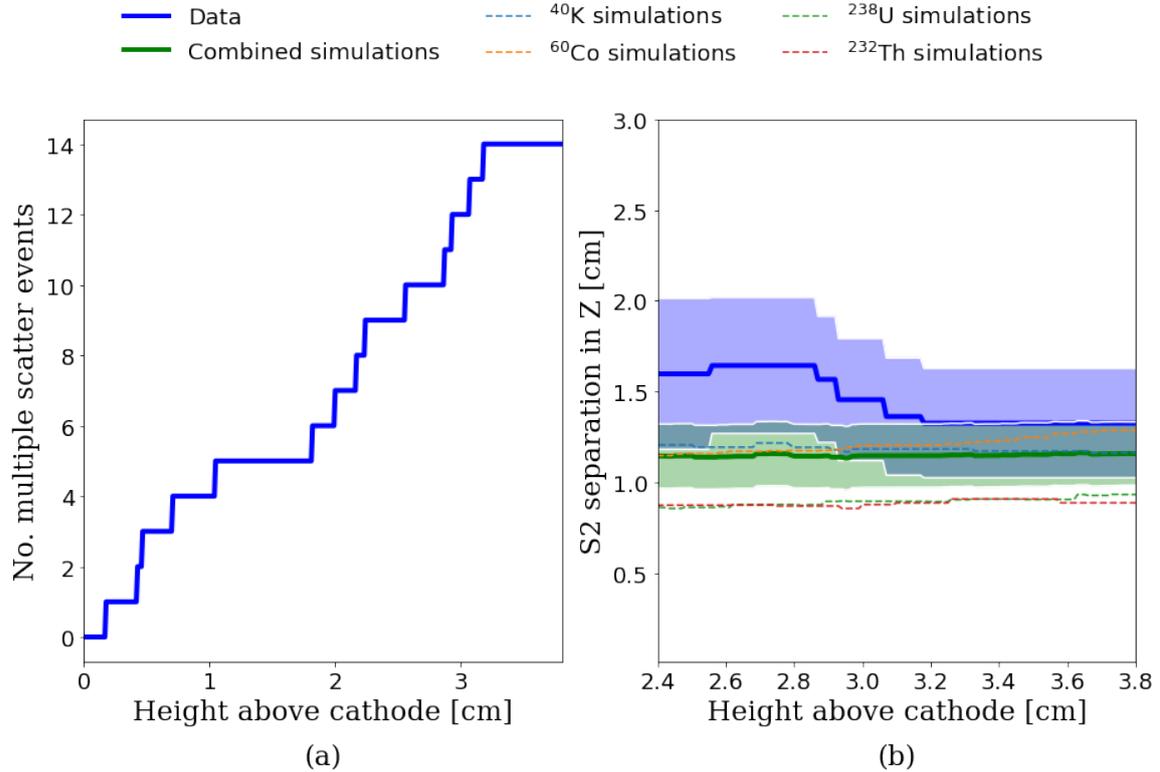


Figure 6.10 Simulated and real nearGX events. (a) the number of MS events which occur within 3.8 cm of the cathode, (b) average $S2$ separation in Z as a function of the height above the cathode for data and simulations.

dual-phase Xe TPCs such as LZ expect to see more gamma-X events on account of their larger volume. As such, it would be prudent to repeat this study on LZ once data is available.

6.6 Gamma-X feature search

Once gamma-X events had been simulated exploration of their feature space could be conducted. This was performed to search for differences between gamma-X and SS events which might be used to inform a cut.

Of all the features that an event possesses six stood out as having potential for adding information to a gamma-X search. These are the three features used in the gamma-X cuts of Section 6.3 ($S1$, CS_{S1} , and $MPAF_{S1}$), in addition to $S2$, Z and Top-Bottom (TB) ratio¹⁰. These features are displayed in Figures 6.11, 6.12, and 6.13. These figures show each parameter

¹⁰TB ratio is the fraction of $S1$ detected in the top PMT divided by the fraction in the bottom

for three independent datasets. In red is the Run04 $^{83\text{m}}\text{Kr}$ data, in blue is tritium calibration data, and in yellow are the simulated gamma-X events from Section 6.5.5. Also included in Figure 6.11b and 6.11d is the 1σ NR band in magenta. The plots on the diagonal of Figures 6.11 and 6.13 are regular 1 dimensional histograms. For all other plots, $^{83\text{m}}\text{Kr}$ is represented as a red-black heat map, tritium as a blue-black heat map, and the simulated gamma-X events as a yellow scatter plot. The simulated gamma-X events were not displayed as a heat map on account of only 410 events falling within the ROI. These plots were made in order to determine if a 1 dimensional or 2-dimensional parameter space might be capable of identifying gamma-X events by looking for any separations between the SS and gamma-X distributions.

All the Run04 EFT cuts presented in Section 6.4 were applied in making these datasets, with two exceptions. These exceptions were the MPA_{S1} leakage cut, and the lower Z fiducial cut. The MPA_{S1} leakage cut was excluded so that the cut could be compared with other gamma-X cuts. The fiducial cut was loosened in order to increase the number of simulated gamma-X events present in the analysis.

Table 6.3 shows the confusion matrix for these six parameters. A confusion matrix shows the cross-correlation between variables. In this table, each entry shows the correlation coefficient (R) between two variables (hence the value of 1 appearing on the diagonal, and symmetry in these values around this diagonal). This matrix was generated based on the correlation of these parameters in the simulated gamma-X events discussed in Section 6.5.5, and also a set of simulated SS events from the same source.

The idea behind this matrix is to reduce redundancy by removing strongly correlated variables. Specifically, this means if $R > 0.95$ or $R < -0.95$, for any pair of parameters then one of them should be removed from consideration as being used in a cut. If parameters are poorly correlated, but also appear to distinguish well between gamma-X and SS events then they have a better chance of positively contributing to a multidimensional gamma-X cut. The matrix in Table 6.3 contains no strongly correlated parameters¹¹. This makes the 6-dimensional parameter space which is described by these parameters appropriate for developing a gamma-X cut.

By examining Figures 6.11, 6.12, and 6.13 several parameter spaces in which a cut on gamma-X events might be developed become evident. Some distinctions between gamma-X

¹¹Note that the reader should not expect to see predicted strengths in this table, only whether two parameters correlate with one another.

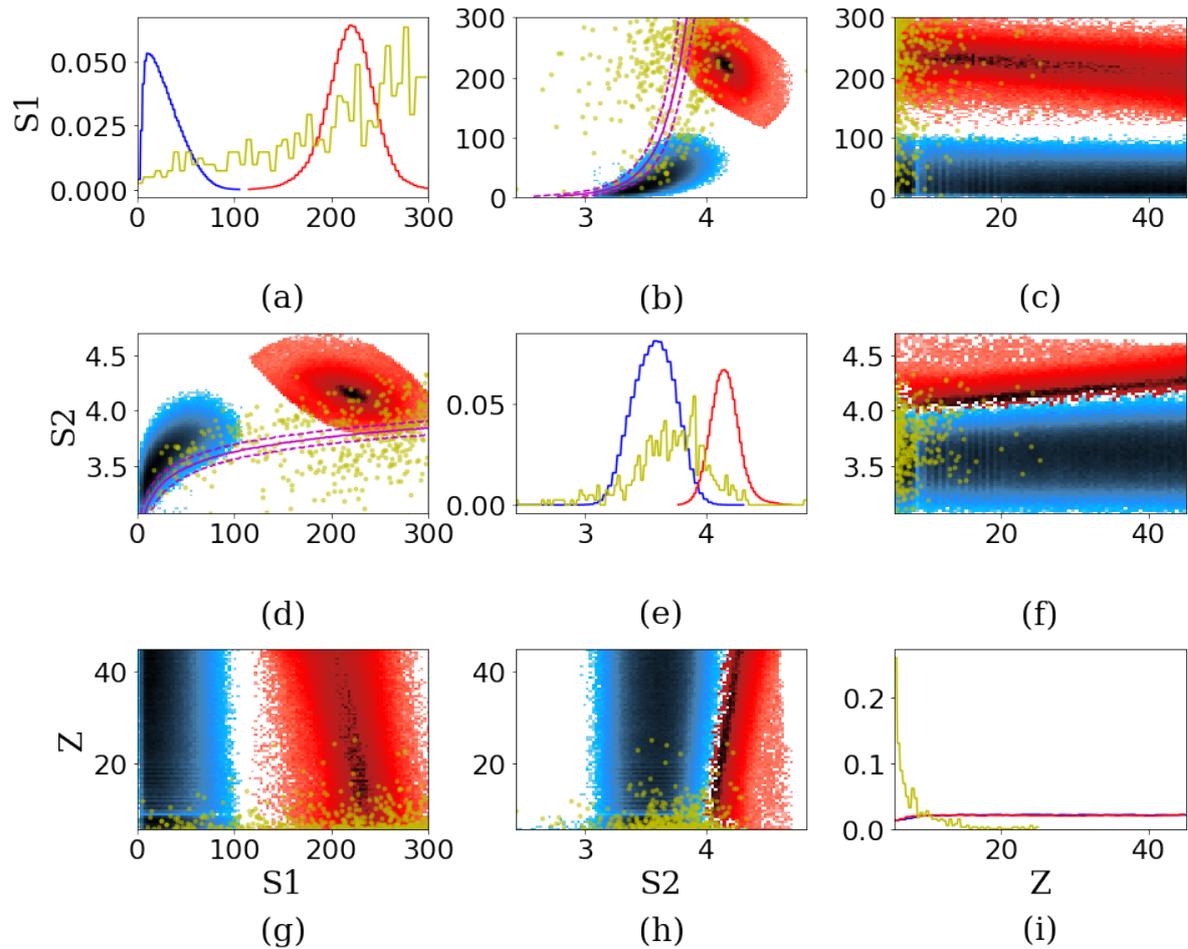


Figure 6.11 Potential parameters for a gamma-X cut. This figure examines the parameter spaces of $S1$, $S2$, and Z by plotting three distributions side by side: ^{83m}Kr data as a red-black heat map, tritium as a blue-black heat map, and the simulated gamma-X events as a yellow scatter plot.

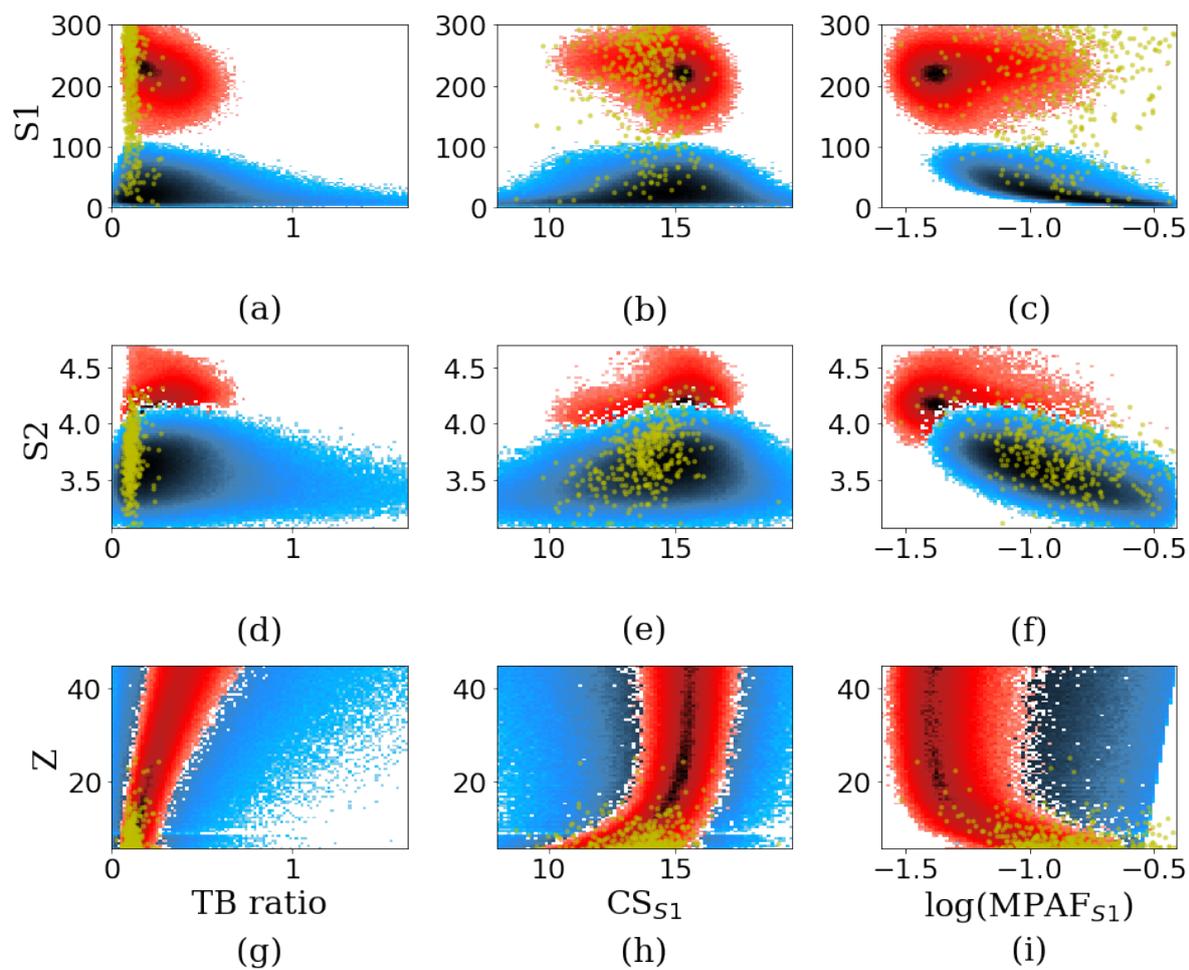


Figure 6.12 Potential parameters for a gamma-X cut. This figure examines some of the parameter spaces of $S1$, $S2$, and Z , TB ratio, CS_{S1} and $MPAF_{S1}$ by plotting three distributions side by side: $^{83\text{m}}\text{Kr}$ data as a red-black heat map, tritium as a blue-black heat map, and the simulated gamma-X events as a yellow scatter plot.

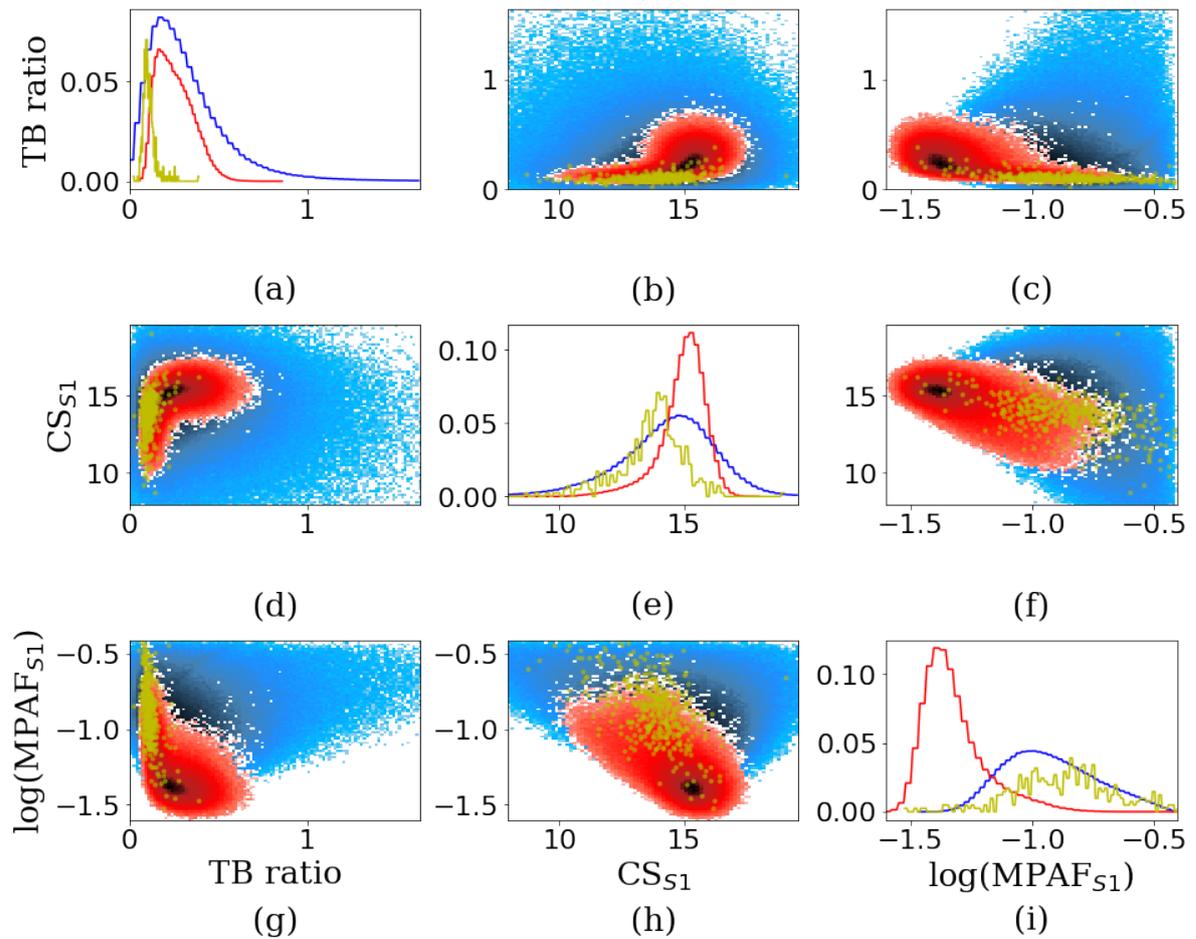


Figure 6.13 Potential parameters for a gamma-X cut. This figure examines some of the parameter spaces of TB ratio, CS_{S1} and $MPAF_{S1}$ by plotting three distributions side by side: ^{83m}Kr data as a red-black heat map, tritium as a blue-black heat map, and the simulated gamma-X events as a yellow scatter plot.

Table 6.3 Confusion matrix for features with the potential for gamma-X identification.

	S1	S2	TB ratio	Z	CS_{S1}	log(MPAF_{S1})
S1	1	0.588	0.027	0.147	-0.124	-0.617
S2	0.588	1	0.165	0.293	-0.001	-0.473
TB ratio	0.027	0.165	1	0.710	0.268	-0.345
Z	0.147	0.293	0.710	1	0.344	-0.477
CS_{S1}	-0.124	-0.001	0.268	0.344	1	-0.278
log(MPAF_{S1})	-0.617	-0.473	-0.345	-0.477	-0.278	1

and SS events are visible in 1-dimensional parameter spaces, whereas others only seem to emerge in 2 dimensions. Beginning with 1-D plots, the first separation in gamma-X and SS events that can be seen is apparent in Figure 6.11i, where gamma-X events are clearly being detected quite close to the cathode. This is consistent with the expectation set out in Section 6.2. For a similar reason Figure 6.13a shows gamma-X events to have a consistently low TB ratio, since TB ratio is sometimes taken as a proxy of depth. However, the TB ratio depends on the depth of the *S1* sites, whereas *Z* depends on the *S2* site of the scatter above the cathode. As such TB ratio is likely to be more valuable to a multidimensional cut. However, since *Z* and TB ratio do not have a strong correlation they will both be considered.

When looking at the 2-dimensional parameter spaces of Figures 6.11, 6.12, and 6.13 additional regions for cut development stand out. The two obvious ones are *S1* vs. *CS_{S1}* from Figure 6.12b, and *S1* vs. *MPAF_{S1}* from Figure 6.12c. Each of these appears to have some level of separation between gamma-X and SS events. Another one which may be useful is *S2* vs. *MPAF_{S1}* from Figure 6.12f. By eye, this one is clearly the weaker gamma-X predictor, but it may contribute to gamma-X identification in a multidimensional parameter space.

One pair of parameters which appears to have a strong distinction between SS and gamma-X events is the *S1* vs. *S2* space from Figure 6.11d. However, this is not desirable as this could lead to a rejection of NR events. Nevertheless, the use of ^{83m}Kr and tritium for these plots highlights the need for NR data when developing a gamma-X cut. As such care will need to be taken to ensure the gamma-X cut does not remove a disproportionate fraction of NR events.

6.7 Conclusion

This chapter described how gamma-X simulations were generated and validated. It then went on to compare these to calibration data in various parameter spaces to determine which event

features might be useful in the development of a gamma-X cut. Based on this outcome, a new gamma-X cut was developed in a 6-dimensional parameter space, one dimension for each feature that showed potential in developing a gamma-X cut in. This new gamma-X cut will be discussed in Chapter 7. These same gamma-X simulations were used to help determine the effectiveness of this new gamma-X cut, along with the cuts described in Section 6.3. This was done in order to determine which gamma-X cut ought to be used by the Run04 EFT search and is discussed in Chapter 8.

Chapter 7

Identification of gamma-X events with a BDT in LUX

7.1 Aim

The aim of this chapter is the development of a gamma-X cut which maintains at least a 95% data acceptance¹, whilst maximising the number of potential gamma-X events which are removed. Given the result of Chapter 6, a 6-dimensional parameter space would be used to develop this cut. To facilitate this a machine learning (ML) algorithm for binary classification known as a Boosted Decision Tree (BDT) was employed.

This chapter will begin by explaining how the BDT of XGBoost (a python ML library) functions in Section 7.2. Following this will be a description of the training events used to develop the BDT in Section 7.3 This will be followed by a discussion of BDT hyperparameter (user-specified variables which define the BDT structure) selection in Section 7.4. Finally, Section 7.5 will discuss the use of the BDT's interpretation of calibration data in the development of a gamma-X cut.

7.2 The XGBoost Boosted Decision Tree

7.2.1 Overview

XGBoost (**eXtreme Gradient Boosting**) is a type of gradient boosted decision tree (BDT), used for binary classification. Capable of making decisions based on any number of input

¹where data acceptance is a proxy for genuine SS event acceptance.

parameters (a.k.a. features), XGBoost is intended to classify large and complicated datasets². The algorithm has a simple operating procedure which begins with feeding the algorithm a set of training data, with each event labelled as either signal (1), or background (0). XGBoost then trains an algorithm on this training data. Once training concludes the classifier can handle input testing data, for which a prediction is made between 0 and 1 indicating how background-like or signal-like each event appears.

Understanding the XGBoost algorithm is best done piecemeal on account of it being made up of many small steps rather than a few complex ones. For this reason, this section will introduce XGBoost by explaining its training algorithm sequentially. Section 7.2.2 will begin by describing the basic building blocks of a BDT. This will include the initial prediction, binary decision nodes, and decision trees. In Section 7.2.3 the iterative process of building an individual decision tree from a series of binary decision nodes will be described. At this stage data is split at each node with the use of the parameters *Similarity Score* and *Gain*, which will also be defined in Section 7.2.3. In Section 7.2.4 the process of tree “pruning” will be described. This involves decision trees having ineffective decision nodes cutaway based on each node’s *Gain* and a *tree complexity parameter* (γ), which are also described in Section 7.2.3. Section 7.2.5 will describe how the output of a decision tree nudges the existing probability score towards the true value, by a small factor ϵ (a.k.a. the Learning Rate). This is based on an output value which is calculated at each terminal node (a.k.a. leaf). Finally in Section 7.2.6 the conditions by which this iterative sequence is halted will be explained by describing what are known as stopping conditions.

7.2.2 BDT building blocks

The basic structure of this BDT can be thought of in this way:

- A BDT is made up of a series of decision trees.
- Each decision tree is made up of a sequence of binary decision nodes.
- The first decision tree depends on a user-specified initial prediction.

To better illustrate this, the first decision tree of the gamma-X cut presented in this chapter is shown in Figure 7.1. This tree is made up of a sequence of decision nodes. The classification

²Notably other ML libraries exist which perform the same task. A common alternative is SKlearn. However, when this was tested head to head against XGBoost early in the BDT development XGBoost was found to be significantly quicker. This finding is consistent with recent results from Kaggle, and is a sentiment echoed in the ML community [119].

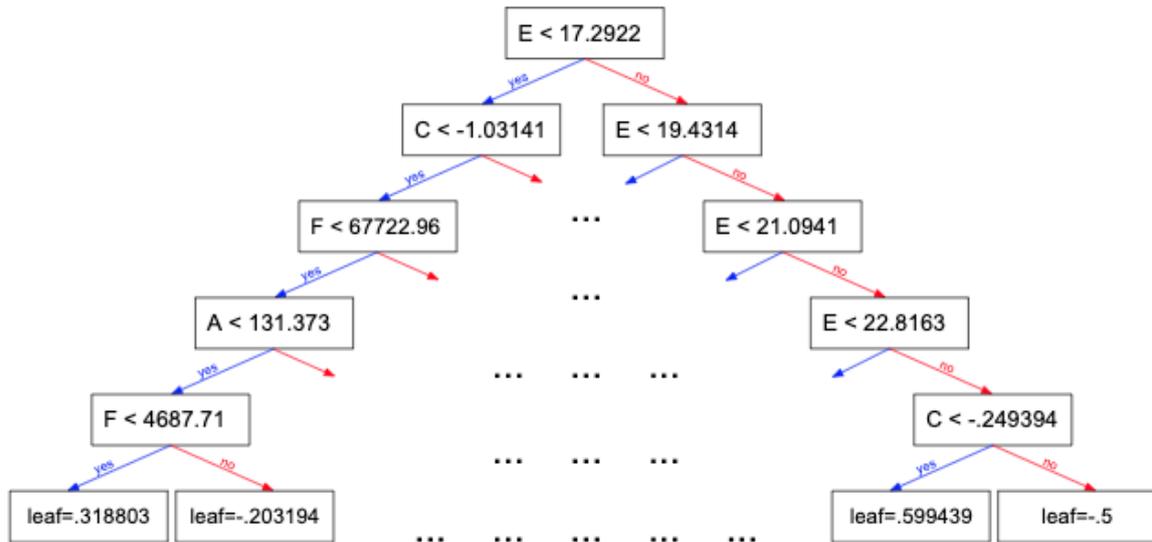


Figure 7.1 Root tree from the BDT used for gamma-X classification. The tree has been simplified by removing most internal nodes (as indicated by the suspension periods). The names variables *A-F* are not relevant to the structure of this tree but are provided in Section 7.2.3.

process of a decision tree begins at an initial decision node, which is known as the root node. As this procedure unfolds, each node is recursively split into two branches, with each branch feeding into a new node³. This pattern continues until a stopping condition is reached. Several decision trees similar to this one combine together to become a BDT.

The number of trees (N) is an example of a hyperparameter; meaning it is specified by the user before training begins. For the gamma-X cut presented in this chapter, this value is 221. This means the output of 221 decision trees is combined to generate this BDT's prediction. The optimisation of this value will be shown in Section 7.4.

Another hyperparameter is the maximum tree depth (D), which for this model is 5. This can be clearly seen in Figure 7.1, where the tree terminates after 5 levels of binary decision nodes. This value was also found to be optimal for this BDT, as will be shown in Section 7.4.

The BDT training sequence outlined in Section 7.2.1 only describes how XGBoost modifies an existing prediction, not how a prediction is initially generated. The reason for this is that the initial prediction is also a hyperparameter known as the global bias (I_0). As a default the value 0.5 is used, which is the optimal value for a set of training data which is 50% signal and 50% background. For highly imbalanced data a more meaningful I_0 may improve BDT

³The node which is split will sometimes be referred to as a parent node, and the two new nodes as the left and right child nodes.

performance. But in theory, so long as there are enough steps to train (N), the starting base score should not affect the result. This effect of various I_0 choices will be shown in Section 7.4.6.

7.2.3 Growing a decision tree

The first step in growing a tree involves feeding all of the training data into the root node. Under ideal circumstances, training data would be real data where an events nature (i.e. signal or background) was known. Since this is not possible for gamma-X events simulations will be used. The training data consists of two types of events SS (s_i), and gamma-X (b_i). Each event in s_i and b_i is described by a set of variables (\bar{x}_i), where:

$$\bar{x}_i = \{A_i, B_i, C_i, D_i, E_i, F_i\} \quad (7.1)$$

The variables A - F in equation 7.1 have been shortened for brevity, their true names are:

- A : $S1$ - the scintillation signal [phd]
- B : CS_{S1} - the cluster size of $S1$ light in the bottom PMT array [cm]
- C : $\log_{10}(MPAF_{S1})$ - the max peak area fraction of $S1$ light
- D : TB ratio - the ratio of $S1$ light in the top PMT array with respect to the bottom
- E : Z - the depth of the $S2$ [cm]
- F : $S2$ - the ionisation signal of an event [phd]

A full description of these variables, and the rationale for their selection was given in Chapter 6. But for now there is little utility in understanding these variables, as BDTs are agnostic to the data they process.

The XGBoost algorithm begins by feeding \bar{x}_i into the root node, where it then considers how to best split the data. To do this the node first sorts all events k times. Where k is the number of variables in \bar{x}_i (which is 6 in this case). For our BDT this will result in six ordered lists; one for each feature in \bar{x}_i :

- $A^{s1} < A^{b22} < \dots < A^{b19} < A^{s3}$
- $B^{s3} < B^{s12} < \dots < B^{b81} < B^{b21}$
- ...

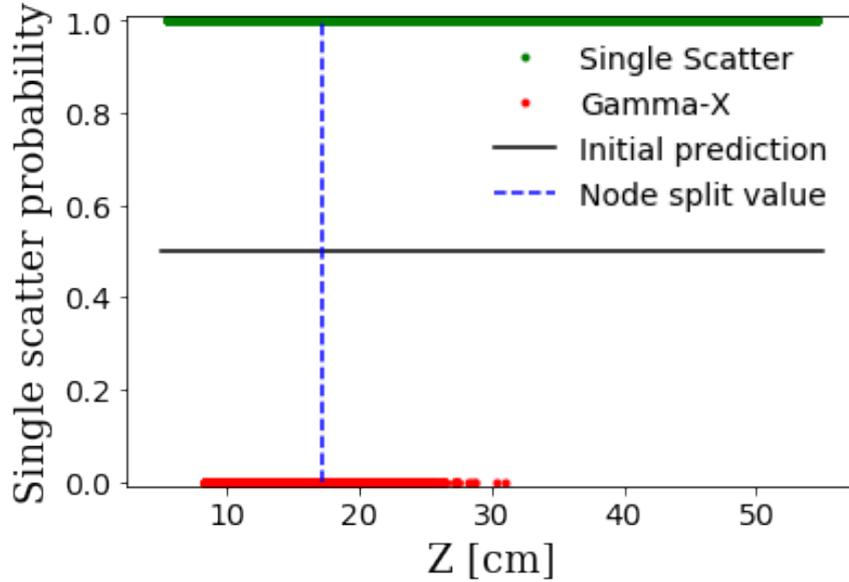


Figure 7.2 Ordering and splitting of training data in the root node. All events (SS and gamma-X) are ordered based on a particular variable; Z in this case. The node then chooses a splitting value which maximises the *Gain* given by Eq. 7.3.

$$\bullet F^{b13} < F^{b99} < \dots < F^{b72} < F^{s98}$$

Note: the symbol “<” indicates the nature of this ordering, and the superscript indicates an event number, and whether it is signal or background. Events are ordered for each variable A - F based on their magnitude with respect to variables A - F .

The node then determines all possible splitting values in each of these hierarchies by calculating the adjacent values (i.e. the value half-way between F^{b13} and F^{b99} is an adjacent value). One instance of this is shown in Figure 7.2, where the training data has been ordered with respect to variable Z . This figure shows the true value of all training events on the Y-axis and the Z on the X-axis. Note that all true values are either 0 (if the event is gamma-X) or 1 (if the event is a SS). This is because the availability of MC truth information eliminates any uncertainty towards the true nature of simulated events. The black line represents the initial prediction, which is the same for all events at the root node. The node then calculates the residual of each event (true value - initial prediction) in order to calculate the Similarity Score (*Sim.Score*):

$$Sim.Score = \frac{(\sum_{i=1}^{i=n} Residuals_i)^2}{\sum_{i=1}^{i=n} (p_i)(1 - p_i) + \lambda} \quad (7.2)$$

where the previous probability (in this case the initial prediction) is given by p_i , and λ is the regularisation parameter. The regularisation parameter will be discussed in Section 7.2.4, where it takes a critical role, but in short it is a user-specified factor used to increase model conservatism. For this model $\lambda = 1$. The selection of this value will be discussed in Section 7.4.6.

The *Sim.Score* is required in order to calculate *Gain*. *Gain* can be thought of as the information gained by making a particular node split. This is the parameter which is maximised in order to determine the splitting value of the node and is given by:

$$Gain = Sim.Score_{LEFT\ CHILD\ NODE} + Sim.Score_{RIGHT\ CHILD\ NODE} - Sim.Score_{PARENT\ NODE} \quad (7.3)$$

Given that the *Gain* depends on the *Sim.Score* of the two output nodes it also depends on both: a defined value to split the node, and a feature upon which to make that split (Z in the case of Figure 7.2). In Figure 7.2, this cut is represented by the dashed blue line. It was chosen by the BDT algorithm because it was the value which maximised the *Gain* on this node.

Note that the *Sim.Score* sums the residuals before squaring. This makes it a measure of how alike events in the summation are. For instance, consider a node split value chosen so that it includes an equal number of gamma-X and SS events. If $I_0 = 0.5$ (thus $p_i = 0.5$ for all events), the summation will be zero. But if more gamma-X events are added to the sample the sum of residuals will decrease, the limit of this sum as the number of gamma-X events approaches ∞ will be $-\infty$. Likewise for adding more SS events to the sample the limit of the sum will increase, and approach $+\infty$ as the number of SS events approaches ∞ .

The process of *Gain* maximisation is run over all variables $A - F$. In doing so the BDT is able to decide which variable best splits the data so as to best increase the purity of subsequent nodes. This means ensuring the highest possible proportion of SS events in one child, and the highest gamma-X in the other. There are no other rules about variable selection, meaning the same variable can be used repeatedly on subsequent nodes if need be.

This process iterates on each child node, and the children of those nodes until either the maximum tree depth (D) is reached, or another stopping condition is triggered. This will be further discussed in Section 7.2.6. When a node cannot be split by the rules of a BDT it is considered a leaf. At a leaf, an output value is returned and used to determine a new probability of the event occurring. The output of several leaves can be seen in Figure 7.1. The output value is similar to the *Sim.Score* from Eq. 7.2, except that the sum of residuals are

not squared, and it is multiplied by another hyperparameter called the Learning Rate, which will be discussed in Section 7.2.5. An intuitive way of interpreting these values is to see leaves with more negative values containing more gamma-X events, and leaves with more positive values containing more SS events. This quantity is a key ingredient in determining the contribution that a leaf makes to the final probability prediction of an event as we will soon discuss. The manner in which this new probability is calculated is where the individual decision tree ends and Boosting takes over. Since Boosting relies upon an understanding of pruning, a description of this process will continue only in Section 7.2.5, after pruning is described in Section 7.2.4.

A final feature of decision trees to note is the difference between the root tree which was presented and subsequent trees. In the initial tree, all events have the same default predicted probability. However, this will change as we move on to the second tree. Not visible in Figure 7.1 is the presence of 32 different leaf nodes. This means events will have 32 different initial predictions to choose from going into the second tree; depending on which leaf they ended up in after the first decision tree. This complexity only increases after the second tree, which also had 32 leaves. Meaning there will be $32^{32} = 1.45 \times 10^{48}$ different initial predictions going into the third tree. It is this iterative process which leads to so many different scores despite each tree only consisting of 5 layers at most.

7.2.4 Decision tree pruning

Once a tree is grown, pruning takes place. Tree pruning is a technique to reduce the size of decision trees by removing nodes that provide little classification power. The reduction of tree complexity aims to improve predictive accuracy by limiting over-fitting. XGBoost employs several methods to enforce pruning, some of which occur top-down, and others which occur from the bottom up. Some of these top-down approaches such as having a pre-set maximum tree depth have already been described in Section 7.2.3. This section will focus on the bottom-up processes of applying a minimum threshold: to node size using the parameter of *Cover*, and to node *Gain* using the *tree complexity parameter* (γ).

In order to reduce overtraining, XGBoost has a minimum number of events allowed at each node. This depends on a parameter known as *Cover*, and is given by:

$$Cover = \sum_{i=1}^{i=n} (p_i)(1 - p_i) \quad (7.4)$$

where p_i once again is the probability assigned to the event by the previous decision tree (or the initial prediction for the first tree). As can be seen, *Cover* depends on the number of events in a node. As a default value, a $Cover > 1$ is required in order for the node not to be pruned. As an indicator of how many events this will contain, one can see that if all events had a probability of 0.5 the node would have to have at least 5 events to avoid pruning. Increasing this threshold can be done as a measure to avoid overtraining. As will be shown in Section 7.4, overtraining did not become a concern for the model presented in this chapter. As such, the threshold for *Cover* was left at the default level (i.e. 1).

The tree complexity parameter is a user-defined constant used to ensure each node makes a positive contribution to the information gained. A tree will be pruned if:

$$Gain < \gamma \quad (7.5)$$

where *Gain* is defined by Eq. 7.3. By default $\gamma = 0$. As will be shown in Section 7.4.6, increasing this value did not significantly reduce over-fitting but did slightly harm the model's predictive power.

One can also see how the regularisation parameter (λ) introduced in Section 7.2.3 is influential in pruning. Since a higher value of λ will give a lower *Gain*. In other words, higher values of λ decrease the tree's sensitivity to outlying events by pruning them, and including them in nodes with other events. In Section 7.4.6 the effect of this parameter on the gamma-X cut will be shown to not affect overtraining, but slightly hamper the model's classification power.

7.2.5 Boosting

Generally boosting refers to the process of combining several weaker predictors into a new more stable one to reduce the error rate. XGBoost achieves this by building decision trees in a sequential manner. This causes subsequent trees to better classify events that are not well classified by the previous trees. This process will be introduced now, as proper explanation depends on the understanding of decision tree building and pruning which were described in the preceding sections.

Once a tree is grown and pruned it can be used to modify the existing probability of each event in the direction of either 0 or 1 depending on whether the event is a gamma-X or SS at the MC truth level. This process begins with an initial probability, which is then iterated on

by each subsequent decision tree. Each iteration is designed to nudge the probability in the direction of the true value.

The first step in this process sees the conversion of the previous probability into a logarithm of the odds, given by:

$$\log(odds) = \log\left(\frac{p_i}{1 - p_i}\right) \quad (7.6)$$

Now comes the minor contribution from the new tree. The new $\log(odds)$ is given by:

$$\log(odds)_{NEW} = \log(odds)_{OLD} + \varepsilon \cdot (Tree\ Output) \quad (7.7)$$

Tree Output is just the *Sim.Score* from the leaf in question. Learning Rate is given by ε and is set to 0.3, the default value.

The Learning Rate controls how much each new tree influences the final prediction by increasing or decreasing its weight by a constant factor. This is another hyperparameter. The selection of 0.3 for this parameter is described in Section 7.4.6.

This new $\log(odds)$ is finally converted back into a probability with the use of the logistic function:

$$Probability = \frac{\exp(\log(odds))}{1 + \exp(\log(odds))} \quad (7.8)$$

For each new tree, this process repeats.

In this way, each new tree is built from the residuals of the previous iteration. When XGBoost functions correctly, each iteration should see the residuals decreasing as the probability moves closer to the true value.

7.2.6 Decision tree stopping conditions

During decision tree building the recursive splitting of a node needs to terminate at some point. Several conditions can lead to this termination. The node which is stopped on is called a leaf, and its *Sim.Score* contributes to the output for the tree. Such stopping conditions include:

- Having reached perfect separation.
- Insufficient information gain by further splitting. As described in Section 7.2.4 this is governed by the *Gain* and γ parameters. If a model seems to be under-fitting this

hyperparameter should be looked at more closely. This is because a weak split in one node may still lead to a powerful split in a subsequent node.

- Having reached a minimum sample size in a node. This requirement ensures a minimum level of statistical significance in each decision node. As discussed in Section 7.2.4, this is handled by the *Cover* parameter.
- Having reached a maximum tree depth (D) as specified by the user. This allows for the optimisation of computational resources.

In addition to the reasoning given for each stopping condition above, another reason for these to be implemented is to prevent overtraining. This is where the decision trees become overly sensitive to minor differences that exist between training and testing samples⁴.

7.3 BDT training event selection

7.3.1 Rationale

Ideally, a BDT is trained on a large set of data, where the binary classification is known. This trained BDT is then used to classify a set of unknown events. For the case of gamma-X classification, this is not possible since no large set of known gamma-X events has ever been verifiably observed. If this were even possible, it would mean gamma-X identification had already been achieved - potentially even rendering this study moot. In order to best handle this lack of identifiably true gamma-X events in data, the BDT was trained on simulated events⁵.

This approach brings with it a different set of concerns surrounding the type of overtraining which can occur as a result of imperfections in simulations. However, it does not prevent the development of a new gamma-X cut which performs better than its predecessors based on predetermined benchmarks (these benchmarks will be discussed in Chapter 8). To achieve this, and mitigate the concerns brought on by training a BDT on simulations, several data sources were called upon.

This section will begin with a discussion about the simulations used to train the BDT known as the “Toy Model”. This will include descriptions of how these simulations were

⁴These differences are statistical in nature.

⁵As will be discussed later data was used to test the BDT, but this only consisted of SS events from calibration data and MS events near the cathode so was inappropriate for training.

produced, and the dataset used to inform them. This will be followed by Section 7.3.3 which gives a description of overtraining - the phenomenon whereby a classifier over-fits due to signal/background differences specific to either the training sample (training sample overtraining) or imperfections in the simulations (data/MC overtraining). As will be discussed these two different types of overtraining need to be handled in different ways. The former involved the division of simulations into training and testing sample sets to ensure consistent BDT response to formerly unseen simulations. The latter involved optimising the structure of the BDT based on its classification of calibration data (known SS events) and MS events near the cathode. Although this section will introduce the concept of overtraining, and the methods employed to mitigate it, the implementation of this will be discussed later in Section 7.4. Section 7.3.4 will discuss how simulations are used as part of measuring classification power and the level of overtraining of a BDT. Finally, in Section 7.3.5 we will discuss the concept of BDT stability, and how its measurement influenced the number of simulated events used to train the BDT.

7.3.2 Toy Model simulations

Due to the low statistics in the gamma-X events produced via full MC chain simulations (see Section 6.5.5), new simulations were created to train the BDT. The new simulated gamma-X events were performed by Greg Rischbieter and were known as the “Toy Model” [120]. These new simulations could be produced significantly quicker, and being data-driven would not suffer from the inconsistencies seen in Section 6.5.2. The production of these simulations followed a four-step process.

Step one began by selecting a set of 1S1 2S2 MS events from LUX Run04 data which best mimicked the gamma-X population. These events were known as nearGX events and were introduced in Section 6.5.6. However, the data selection cuts for these events varied slightly. The first difference was that the lower S2 only had to be within 4 cm of the cathode. The radial fiducial cut for these nearGX events was also loosened to include all events up to 1 cm from the wall. A final selection cut was applied to the nearGX sample to ensure the S2 sites from the same MS event were at least 4 cm apart in Z. This 4 cm distance ensured that if an event was translated vertically so that the initial scatter of the MS event occurred beneath the cathode, the upper scatter would still be above. After these exclusions, 17 events remained.

Step two addressed the necessity of increased statistics for BDT training. This was solved by simulating events with NEST. These simulations were all 1S1 2S2 MS events in the same

detector region as the nearGX selection. For each event, this gave: both XYZ positions of energy depositions, both $S1$ signals, and both $S2$ signals, and both photon-yields from scintillation sites.

Step three involved comparing these simulations to the data selection. This was done in the following feature spaces:

- $S2$ vs. $S1$ space.
- MPA_{S1} vs. $S1$ space.
- CS_{S1} vs. $S1$ space.
- TB ratio vs. $S1$ space.

These feature spaces can be seen in Figure 7.3, with nearGX data represented as open circles and simulations as scatter dots. In each one a correction factor was applied if the simulated spectrum did not match the data [120].

Finally, each simulated event in this model was then translated down 4 cm, to become a simulated gamma-X event (a.k.a. ToyGX). The field values at the modified site were then used to recalculate the $S1$, $S2$, and photon yields for each energy deposition site. These photon yields were then used by NEST_light to determine CS_{S1} , MPA_{S1} , and TB ratio in the same method described in Section 6.5.1. Finally, each of these parameters was tuned based on its $S1$ signal as mentioned in the previous paragraph.

Toy Model events proved to be much more useful for BDT training than the full MC simulations described in Chapter 6. There were two main reasons for this. First, they could produce SS and gamma-X events at a significantly faster rate. This is because the standard simulation chain required the generation of many events before applying gamma-X/SS selection cuts, and the data quality cuts to see what remained. Since the Toy Model was producing events based on this selection, all simulated events were usable by the BDT. Secondly, NEST_light was integrated more smoothly into the Toy Model event production. This meant that the $S1$ features: CS_{S1} , TB ratio, and MPA_{S1} could be calculated more quickly.

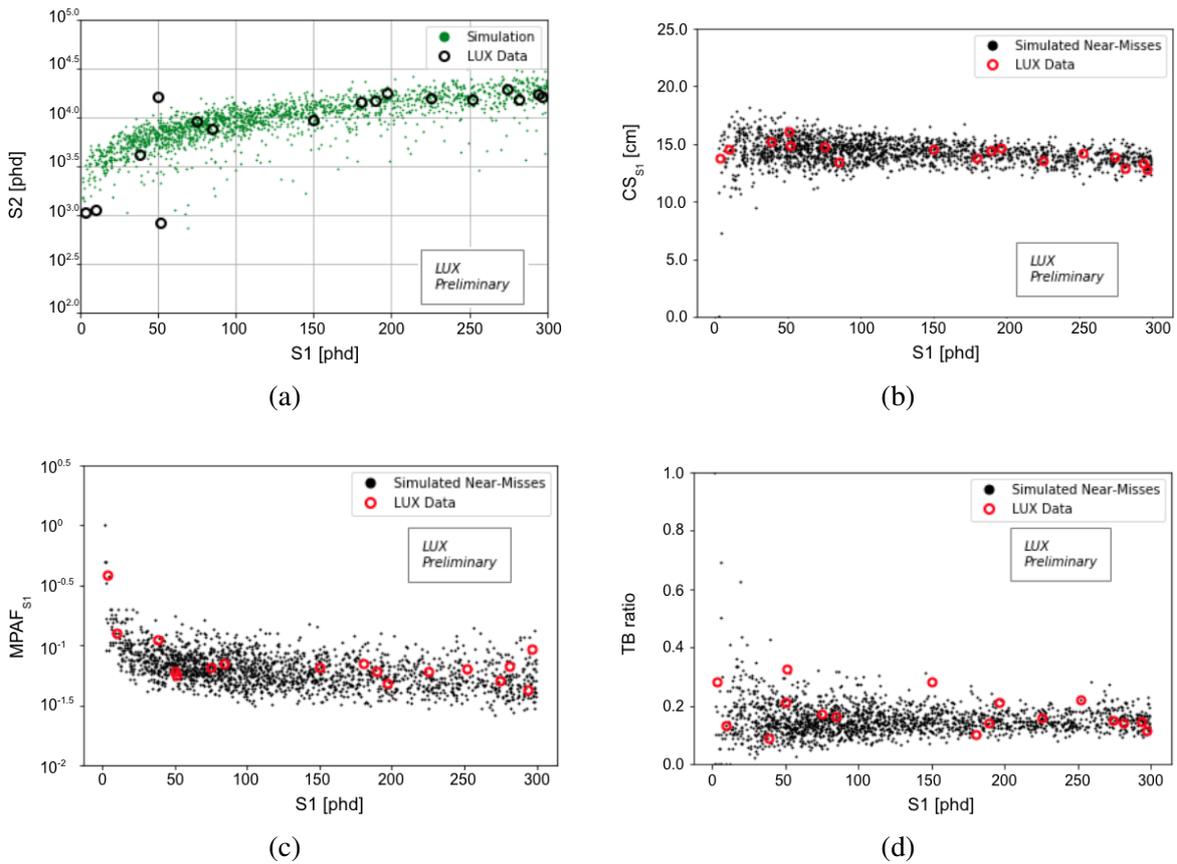


Figure 7.3 Toy Model validation plot taken from [120]. This figure displays nearGX data (open circles) plotted alongside events simulated with NEST. In (a) this is displayed in $S2$ vs $S1$ space, (b) displays the events in CS_{S1} vs $S1$ space, (c) displays the events in $MPAF_{S1}$ vs $S1$ space, and (d) displays the events is TB ratio vs $S1$ space.

7.3.3 Overtraining mitigation

When building a decision tree one initially begins with an arbitrarily long depth and no minimum number of events per node. For the events in this tree's training sample, training would continue until there were only a few events in each leaf, and perfect signal/background separation had been obtained. But if this BDT was then put to the task of classifying events from an independent testing sample drawn from the same underlying distribution, the same level of success might not be achieved. This situation is called *overtraining*.

More precisely, this kind of overtraining is known as training sample overtraining and can appear in any type of classifier. It occurs when a classifier makes decisions based on signal/background differences that are specific to the particular training sample used. Correct BDT training requires that a classifier models signal/background differences that are present in one or more separate testing samples as well as in the unknown events which will ultimately be classified.

In a BDT training sample overtraining can be detected by comparing the score distributions for the training and testing samples for both signal and background. One way to do this is by means of the Kolmogorov-Smirnov test (KS test) [121].

The KS test provides a p value equal to the statistical probability that two samples are drawn from the same distribution. In Section 7.4.2 several p values are compared in order to optimise hyperparameter selection. It is preferable to maximise the p value since the smaller the p , the greater the chance of overtraining. Since the training and testing samples will never be identical, a very small degree of overtraining is unavoidable. This means a p equal to 1 is never achievable. As will be seen in Section 7.4, during hyperparameter tuning a significant number of hyperparameter combinations were found to give a $p > 0.999$. As such this was chosen as a minimum threshold for this analysis. However, this is not a general rule. Since overtraining is unavoidable and specific to the model, the tolerable level of overtraining will vary. But if one is looking for a rule of thumb, if $p < 0.95$, or if the separation between training and test data is visibly poor (see the Classifier Output Distribution in Section 7.3.4), then the classifier should be reworked.

This approach does not address a different kind of overtraining which arises from differences between simulations and data. This is known as data/Monte Carlo (MC) overtraining and only occurs in classifiers built upon simulations and then applied to data (or vice versa). Since MC can never reproduce real data perfectly, not all signal/background differences in the training

sample are really due to the difference between a gamma-X and a SS event. Some differences are inevitably due to undocumented imperfections in simulation. If the training stopping criteria are too relaxed, then BDTs can make cuts that separate SS data events from simulated SS.

Since the standard overtraining check described above does not reveal data/MC overtraining, it is much harder to identify its presence or effect. To minimise the risk of this occurring during hyperparameter optimisation, BDT testing was run over two independent datasets. These datasets should not be confused with the so-called test sample which will be introduced in Section 7.3.4. Unlike the test sample, the datasets referred to here are real data, not simulations; and as such, have not been generated via the same process as the training data.

First was tritium calibration data injected in December 2013 prior to Run04 [82]. This consisted of 1.7×10^5 fiducial events after applying all Run04 LUX EFT data analysis cuts (described in Section 6.4), which was further reduced to 2.6×10^4 by applying the stricter fiducial cut of $Z \in [5 \text{ cm}, 12.5 \text{ cm}]$ and $R < 15 \text{ cm}$. This stricter cut was employed to further reduce the presence of MS events originating from detector surfaces which may have passed MS tagging⁶. The Z range was also limited to 12.5 cm to ensure SS events are not being removed just because of their depth. In Section 7.4.4 this dataset was used to test the BDT for a range of hyperparameter combinations, in order to choose a combination which maximised SS identification in the data.

Assessing gamma-X classification efficiency based on the BDTs response to real data was more complicated due to the lack of positively identified gamma-X events. To handle this an alternative dataset was turned to - a variation of the nearGX events described in Section 7.3.2. This dataset also consisted of 1S1 2S2 events occurring within 4 cm of the cathode, and with all LUX Run04 EFT cuts applied (described in Section 6.4). However, the requirement to have subsequent S2 sites separated by 4 cm was removed.

Although these nearGX events are not gamma-X events themselves, they are the closest event which can be identified in the data. Because of this, they are expected to have similar features to gamma-X events. Thus, an effective gamma-X cut ought to remove a higher proportion of these than it does of SS events.

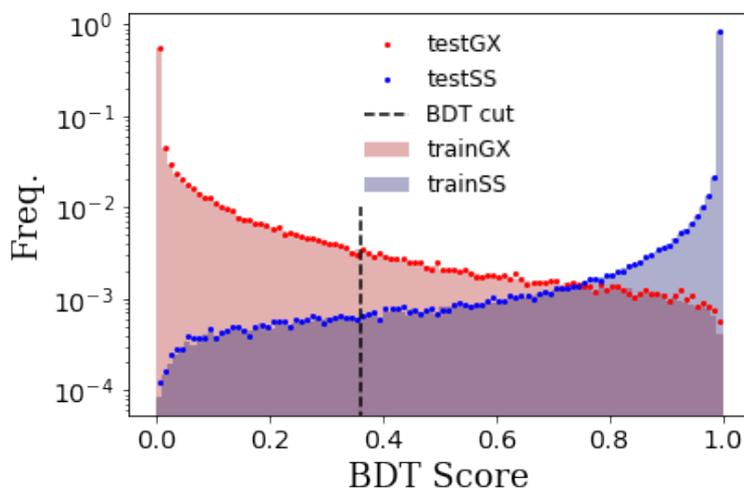


Figure 7.4 The Classifier Output Distribution used to assess BDT performance. Training events are given by the histograms, test events by the dots. Red indicates the event is gamma-X, blue is for SS. The dotted black line indicates the classification threshold. All events below this line are categorised as gamma-X. The BDT is trained to give events which appear more single scatter-like a score closer to 1, and events more gamma-X-like a score closer to 0.

7.3.4 BDT performance

BDTs are employed to make binary classifications on events in a dataset. As such any focus given to measuring and minimising overtraining should also be met with the measurement and maximisation of classification power. A useful figure for comparing BDT performance on simulated training and test samples is known as Classifier Output Distribution (COD) and can be seen in Figure 7.4. The X-axis of this figure shows the BDT score, and the Y-axis shows the frequency of events. In the COD gamma-X and SS events are differentiated to show if there is any inconsistency between background and signal classification. A distinction is also made between training and test events to ensure consistent behaviour on events that it has never been exposed to before. A classification threshold is also illustrated on the curve. This threshold determines what BDT score will correspond to a gamma-X classification. In this case, the value is 0.36, the evaluation of this parameter is described in Section 7.5.

Once events are classified the BDT prediction can be compared with MC truth. This allows events to be categorised into 4 groups:

- True positive (TP) for a correctly identified SS events.

⁶Although tritium will only produce SS events, MS events will still be present due to backgrounds originating from detector surfaces.

- False positive (FP) for gamma-X events identified as a SS events.
- True negative (TN) for correctly identified gamma-X events.
- False negative (FN) for SS identified as gamma-X events.

With this categorisation a classification score can be calculated to measure the quality of a BDTs prediction. This score is defined as:

$$Class.Score = \frac{TP + TN - FP - FN}{TP + TN + FP + FN} \quad (7.9)$$

This classification score will be used in Section 7.4 as a measure of the BDTs accuracy in classifying simulated events.

Agreement between training and test samples in Figure 7.4 indicates that the model is not overfitted to the training sample. This is a kind of overtraining that occurs when a classifier is tuned to the point that random fluctuations in the training data have a measurable effect on final classification. Since the model behaves consistently between training and test samples (the latter of which the BDT has not been exposed to before) the model is said to not be overtrained. To quantitatively measure overtraining a p value derived from a KS test, which was run between the training and test distributions of both the signal and background. This leads to two p values being produced: one for the gamma-X KS test ($p = 0.9921$), and one for the SS KS test ($p = 0.9995$) [121]. The gamma-X and SS p values tended to be quite similar, but to ensure that the measure of overtraining used was representative of both types of events, an average of the two was taken.

Since this averaged KS test metric is based purely on simulations it does not account for overtraining effects stemming from differences between simulations and data. This type of overtraining cannot be measured or addressed with simulations alone. Even though hyperparameter tuning may enhance the effect, it cannot address the root cause.

7.3.5 BDT stability

When splitting simulated events into training and testing samples as described in Section 7.3.4 the need for a large number of events to draw upon quickly becomes evident. If there are not enough events to draw from, the capability for outlying events included in the training sample to sway the classification is enhanced. When this occurs, the BDT may not necessarily appear to be overtrained by looking at the COD, but when the BDT is retrained based on a different random selection of events it may not classify the same events in the same way. Due to

the inconsistency of a BDT's classification ability under these conditions, the name for this phenomenon is *BDT instability*.

To reiterate, BDT stability refers to the ability of the classifier to produce the same result when altering the random allocation of events to the training and test samples. If the number of simulated events is high enough, changing this will have a small effect on BDT performance. The best solution for an unstable BDT is to include more simulations in the training and test samples. This comes at the cost of BDT creation time though. As such BDT stability needs to be measured.

To ensure the BDT used in Figure 7.4 was stable the following steps were followed. First, a mean of 10 different BDTs classification scores (from Eq. 7.9) was calculated. Each BDT in the set was identical, except for the random allocation of events to training and test samples which changed. The classification score can be simply expressed as the rate of correctly identified events minus the rate of incorrectly identified events. This definition is expanded upon in Section 7.3.4.

The mean (μ) of these BDT scores for the various number of simulated events can be seen in the upper plot in Figure 7.5. A stable BDT would produce a consistent prediction for the same events regardless of change to the training data. This variation is particularly notable in the error bars, which is the standard deviation (σ) of the 10 BDT classification scores plotted in each point. The most likely explanation for any variation in this BDT classification score is that the training distributions between BDTs did not match.

The mean value of repeated BDT classification scores as the number of simulations increases is seen to converge on ~ 0.919 . Much in the same way σ is also seen to converge towards zero as the number of simulations increases. For this reason, the σ was taken as a measure of BDT stability such that smaller values indicate predictions are becoming more consistent.

Notably, the classification score is only sensitive to the total number of events being classified as gamma-X, not the effect of the BDT on individual events. However, if the same events are being tagged as gamma-X, and the test sample has the same ratio of signal to background, σ will approach zero. As such, a large σ is certainly due to differences in which particular events are being classified as gamma-X.

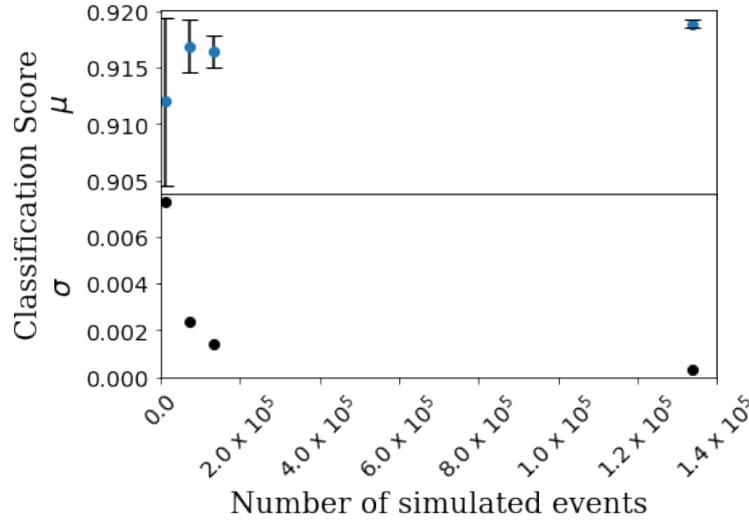


Figure 7.5 BDT performance on a varying number of simulated events. The top figure shows how the average of 10 BDTs classification efficiencies varied as the number of simulated events increased. The error-bar represents 1σ on the variation of the mean represented in each bin. The lower figure shows the standard deviation of the upper plot. The number of simulations increased until this σ was below 0.001.

In order to determine the minimum number of simulations required for the BDT to behave stably, the number of simulations was continually increased until σ fell below 0.001. As can be seen from Figure 7.5, this occurred after 1335911 events were simulated. A stricter value could have been required; however, this did not seem to significantly affect the Run04 data which was being tested.

Of these simulations $\frac{1}{3}$ were ER SS events, $\frac{1}{3}$ NR SS events, and $\frac{1}{3}$ were gamma-X. Ideally, BDTs operate on 50% signal and 50% background. However, since the BDT was using $S1$ and $S2$ as cutting parameters keeping a balance of ER and NR events in the SS population was prudent. This helped ensure the BDT did not favour $S1$ and $S2$ in such a way as to remove NR events instead of gamma-X events.

7.4 Hyperparameter selection

7.4.1 Purpose of hyperparameter optimisation

A BDT requires several user-defined parameters known as hyperparameters. Careful hyperparameter selection can improve a BDTs classification power, and reduce overtraining.

For the BDT presented in this chapter, this means testing several different hyperparameter combinations. Ideally, this is achieved by generating BDTs with all possible combinations of hyperparameters, then choosing the one which gives the best performance. However, this approach is computationally expensive. Subsequently, an alternative approach has been taken, whereby hyperparameters were optimised individually. This involves looking for an optimal value of one hyperparameter, with all others held constant. The optimisation was based on a BDTs classification score, and KS test p value, both of which were introduced in Section 7.3.4. Once a hyperparameter was optimised it too was held constant before moving onto the next hyperparameter for tuning.

This process began with the simultaneous optimisation of N and D (the number of trees and the maximum tree depth), which were thought to have a most dramatic influence on the outcome of the final BDT. The simultaneous approach to this optimisation stemmed from a cross-correlation whereby altering of either N or D seemed drastically affect the optimal value of the other. This optimisation was handled in unison by means of a grid search, which will be described in Sections 7.4.2, 7.4.3, and 7.4.4. Following this was the standard 1 dimensional optimisation of the Learning Rate, which is discussed in Section 7.4.5. Finally, all other hyperparameters were examined, as described in Section 7.4.6.

7.4.2 Coarse N and D simulations grid search

Both hyperparameters N and D influence the classification ability of our BDT. Increasing either of these leads to more final leaves in a BDT. The benefit of this is that the BDT can better categorise an event. However, too many leaves can potentially lead to overfitting. Subsequently, optimisation of N and D was done in 2 dimensions by employing a grid search.

To begin this optimisation the potential number of N and D hyperparameter combinations needed to be restricted. This was achieved by running a coarse grid search. This involved training a BDT with specific values of N and D , and determining its classification score and p value, before moving to a new pair of N and D values. This grid search was described as coarse since not every possible combination of N and D was examined. For this search, N was restricted between 50 and 200 and only measured at steps of 50. D was restricted to between 5 and 20 and only measured in steps of 5. This yielded 36 different classification scores and p values for each combination of N and D .

The result of this coarse grid search can be seen in Figure 7.6. As can be seen, the larger the number of leaves the better the classification score is, with a marked improvement in

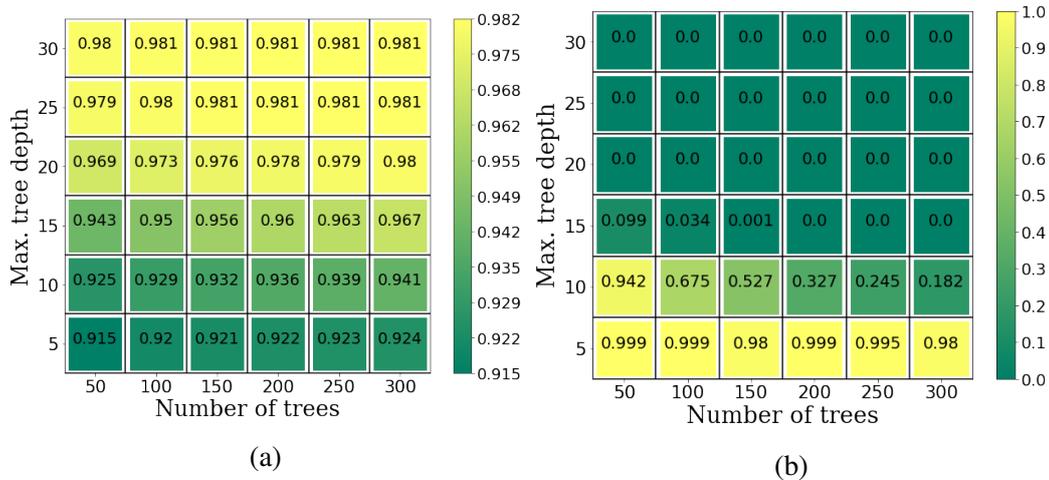


Figure 7.6 Figure (a) shows the variation of the classification score with the number of trees and the tree depth. Figure (b) shows how these two hyperparameters affect the p value measuring overtraining. This grid search was performed over a coarse range of N ranging from 50 - 200, at steps of 50, and D ranging from 5 - 20 with steps of 5.

classification scores in grids from the top right corner of Figure 7.6a. However, this comes at the direct trade-off with overtraining, as indicated by the decline of p values in grids from the upper right corner of Figure 7.6b.

Based on the available classification scores and p values for various N and D combinations thresholds were set. Only grids with a classification score greater than 0.92 and p values greater than 0.999 would be considered. This yielded a reduced range for N of 75 to 225, while D was⁷ 3 to 8.

The result of this analysis was used to reduce the (N, D) space in which to search for an optimal value. In this fine grid search, all N and D combinations would be examined.

7.4.3 Fine N and D simulations grid search

With a reduced (N, D) space to search in all combinations of these two hyperparameters could be examined. The result of this can be seen in Figure 7.7, which shows the classification score (Figure 7.7a), and p value (Figure 7.7b) for all (N, D) combinations. Given the range in classification scores and p values available the number of (N, D) combinations was reduced

⁷The point of this analysis was to reduce the field of possible (N, D) combinations. This is why although grids corresponding to $N = 75$, $N = 225$, $D = 3$, and $D = 8$ are not shown on this figure, the fine grid search was extended to include them.

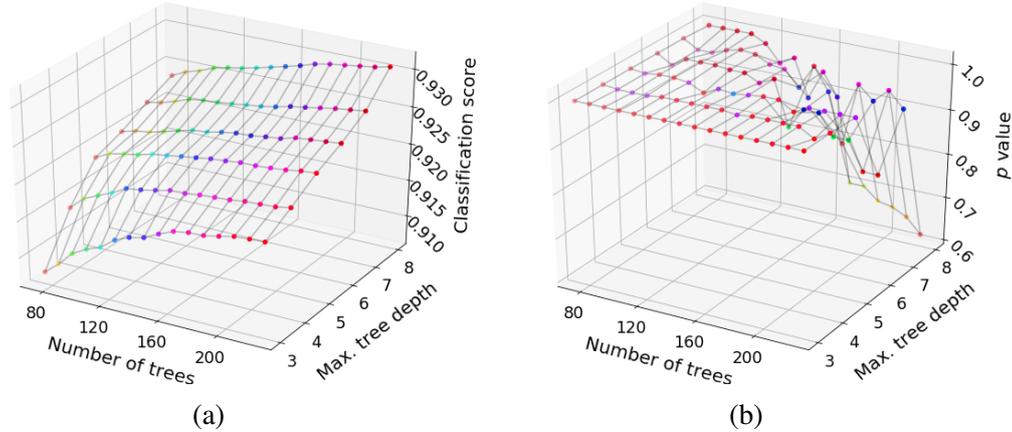


Figure 7.7 (a) shows the variation of the classification score with the number of trees and tree depth. (b) shows how these two hyperparameters affect the p value measuring overtraining. This grid search was performed for all (N, D) combinations in the N range of 75 to 225, and the D range of 3 to 8. For visual clarity of the trend not all values for N are displayed.

to 100 by setting minimum thresholds on both of these metrics. Specifically, the BDT needed to have a classification score > 0.916 , and a p value > 0.997 . These thresholds defined a minimum BDT performance on simulated events. However, to ensure that this did not introduce an unnecessary level of overtraining on simulated events, the final 100 BDTs were tested on calibration data to make the final selection for N and D .

7.4.4 N and D selection

An unforeseen variance between simulation and data is a practical certainty, but also unpredictable in how it will manifest. For this reason, the BDT cannot be solely tuned on simulations, data must also contribute to minimise this impact. This is why a minimum threshold for the BDTs performance on simulations was loosened to allow many versions of the model to pass. These finalists were then tested against data to make the final (N, D) selection. To make this selection two tests were run. First, the signal efficiency against tritium data was examined. Second, the rejection rate of nearGX events was examined.

Tritium data was selected because it cannot cause a MS event. Similar to the validation to determine the CS_{S1} gamma-X cut [114], strict fiducial boundaries were set so as to minimise wall events, and true gamma-X event contamination (as described in Section 7.3.3). The BDT was set to classify SS/GX events from this dataset, and the efficiency was recorded. The nearGX event test looked at 1S1 2S2 MS events from the Run04 data, where the lower

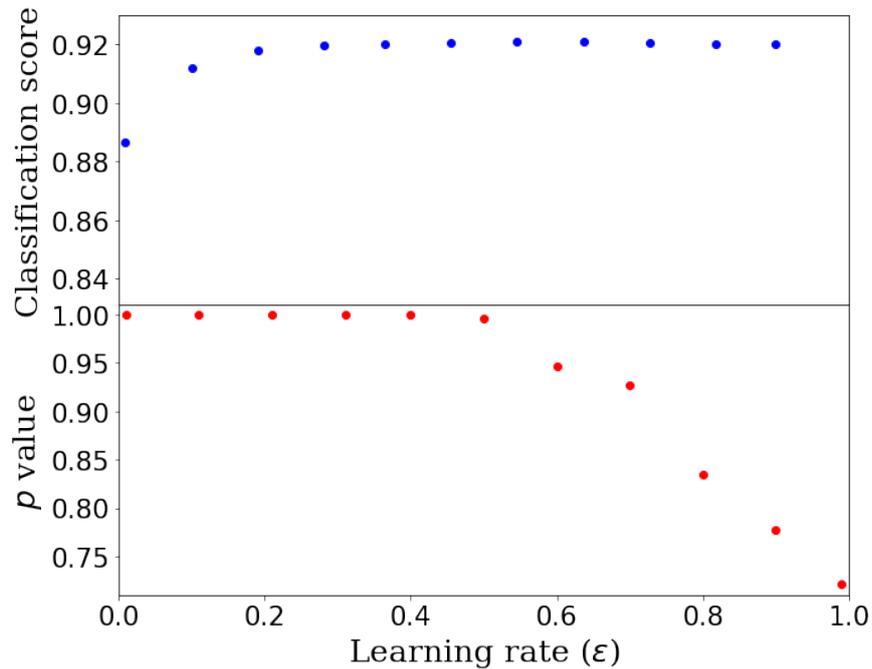


Figure 7.8 This figure shows the effect of changing ϵ on the classification score (upper plot), and the p value measuring overtraining (lower plot).

S_2 was within 4 cm of the cathode - as described in Section 7.3.3.

The 100 (N , D) combinations that worked best on simulations were all put through this test. Of these, only 25 had a tritium efficiency $> 85\%$ and a nearGX rejection rate $> 40\%$. The best performing of these was (N , D) = (221, 5) with a tritium efficiency of 88.2%, and a nearGX rejection rate of 42.9%.

7.4.5 Learning rate

Various values for the learning rate (ϵ) were also examined in order to find a value which optimised classification power without overtraining. To achieve this, several values for ϵ were used between 0 to 1. The result of this can be seen in Figure 7.8, which includes the effect of this tuning on both classification score and p value.

As can be seen in the upper plot of Figure 7.8, the classification score starts high; with even the lowest value of ϵ returning a classification score > 0.88 . This is because the values for N and D which were used were already optimised, leaving little room for improvement. The improvement available to the BDT by increasing the learning rate clearly suffers from a

steeply diminishing return, with the classification score rapidly approaches an asymptote at ~ 0.92 . Above $\varepsilon = 0.2$, improvement in the classification score is essentially negligible. The effect of increasing ε on overtraining is also measured in the lower plot of Figure 7.8. This plot shows the p value, which appears to be flat from $\varepsilon = 0$ up until about $\varepsilon = 0.4$. At this point signs of overtraining begin to emerge in a swift drop in p value.

This gave a range of appropriate values for ε between 0.2 and 0.4. To maximise the classification score while minimising the risk of overtraining the values of 0.2, and 0.4 were split giving $\varepsilon = 0.3$. This represented a fair trade-off between the two considerations, as it is above the $\varepsilon = 0.2$ threshold where classification score is essentially maximised, but also below the $\varepsilon = 0.4$ point where the KS test becomes sensitive to overtraining.

7.4.6 Other hyperparameters and final hyperparameter selection

XGBoost has a long list of hyperparameters that can be tuned. Just like N , D , and ε which have already been discussed, each hyperparameter serves to either make the model more or less conservative. As a result of the tuning of the aforementioned hyperparameters, additional hyperparameter tuning yields a diminishing return on performance.

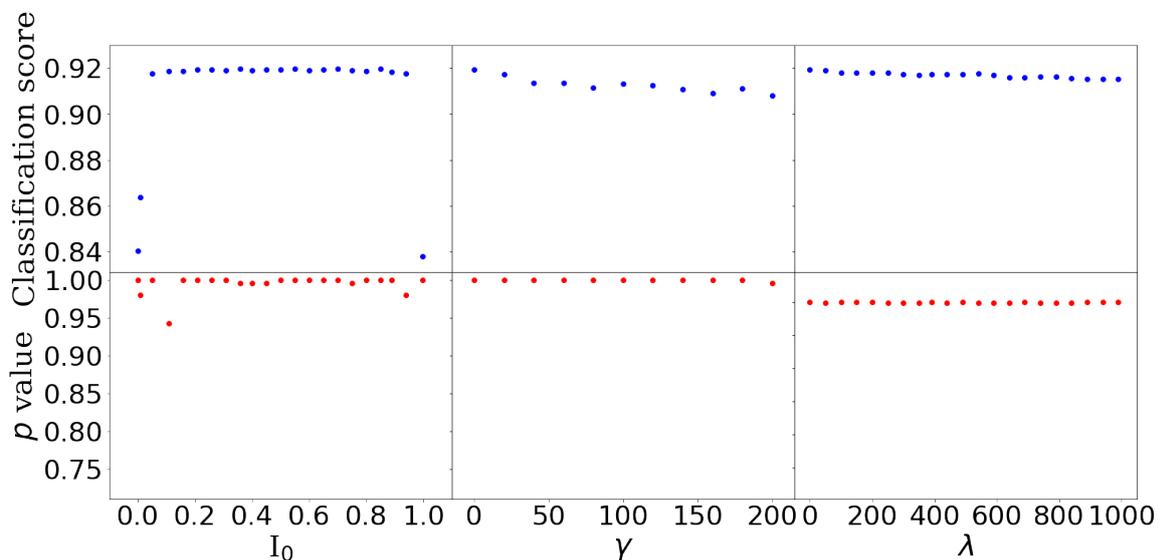


Figure 7.9 The effect of altering the hyper parameters I_0 (left), γ (centre), and λ (right) on classification score (upper), and the p value measuring overtraining (lower).

This point is illustrated in Figure 7.9 where three other hyperparameters were altered to see the effect on the BDT's performance. The three hyperparameters here are global bias (I_0), the

tree complexity parameter γ , and the regression parameter λ . For each hyperparameter, the effect of altering it is measured in terms of classification score and p value.

On the left of Figure 7.9 I_0 is adjusted. I_0 is the initial prediction for all events before the effect of the first tree is applied and is bound between $[0, 1]$. XGBoost sets this to 0.5 by default, as this is considered the optimal number for a training dataset which is 50% signal, and 50% background. However, for a sufficiently large N , this ought to not have a significant impact. This was confirmed in Figure 7.9, where aside from reduced performance when I_0 was particularly close to either 0 or 1, altering this parameter did not have a significant impact. Similar effects were seen for γ (Figure 7.9 centre) and λ (Figure 7.9 right). This figure shows that deviation from the default values of 0 and 1 respectively barely affect performance. If anything there was a slight negative trend. As a result of this examination, I_0 , γ , and λ were left at their default levels of 0.5, 0, and 1 respectively.

The final values chosen for the BDT hyperparameters discussed in this chapter are shown in Table 7.1. With a finalisation of the hyperparameters, the BDT was capable of returning a BDT score for any event. This score was based on how gamma-X-like the BDT considered an event to be. BDT scores ranged from 0 to 1. If an event had a score closer to 1 it was considered more SS-like, and closer to 0 more gamma-X-like. However, to create a gamma-X cut with this information meant defining a BDT cut score. Strictly speaking, the BDT cut score is not a hyperparameter, as such its selection will be described in Section 7.5.

Table 7.1 Hyperparameter selection for the BDT

Hyperparameter	N	D	ϵ	I_0	γ	λ
Value	221	5	0.3	0.5	0	1

7.5 BDT cut development with tritium and ^{14}C data

The definition of the BDT cut value was a question of conservatism, which in this case was a trade-off between two main considerations. The first was to ensure the cut was conservative enough to remove any true gamma-X events. The second was to ensure the cut was not so conservative that too many SS events are accidentally removed by the cut. To meet these two considerations an efficiency threshold of 95% was set. With this value in mind, the BDT was tested on a selection of tritium and ^{14}C calibration data. The BDT cut value was then lowered until it met this requirement.

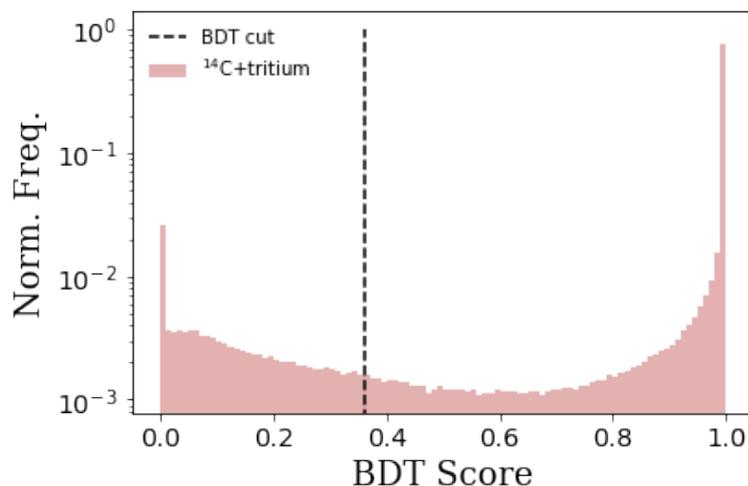


Figure 7.10 The BDT score for a combination of tritium and ^{14}C calibration data. This plot was used to inform the position of the BDT cut which is displayed as a dashed black line. A cut value of 0.36 was chosen as this was the threshold which gave a 95% acceptance.

The result of this can be seen in Figure 7.10. This figure shows the BDT scores of this calibration data in a histogram. It also includes the finalised BDT score of 0.36, which excluded 5% of calibration data. The data in this analysis had all LUX Run04 data analysis cuts from Section 6.4 applied, including the MPA_{S1} leakage cut.

Notably, two parameter spaces were excluded from consideration by the BDT gamma-X cut. The first was events above $Z = 20$ cm. A few events were still being classified as highly gamma-X-like by the BDT above this threshold. This was not considered realistic based on how far from the cathode they were occurring. The rationale for this is described in Section 6.2. The second set of events which were not considered by the gamma-X cut were events with an $S1 < 11$ phd. This was applied due to the difficulty experienced in reconstructing bottom PMT photo-hit-patterns when simulating events with such few photons. This difficulty can be clearly seen in Figure 6.9. Although this figure pertains to the full MC simulations, it extends to the Toy Model events (since the same method was used to determine CS_{S1} and $MPAF_{S1}$). Although these exclusions will leave an analysis vulnerable to gamma-X events with a low $S1$, or which occur far away from the cathode, for reasons outlined in Section 6.2 this was considered reasonable.

None (or very small number) of the events in this sample are expected to be a gamma-X. So in order to give an idea of what kinds of SS events are being removed Figure 7.11 is

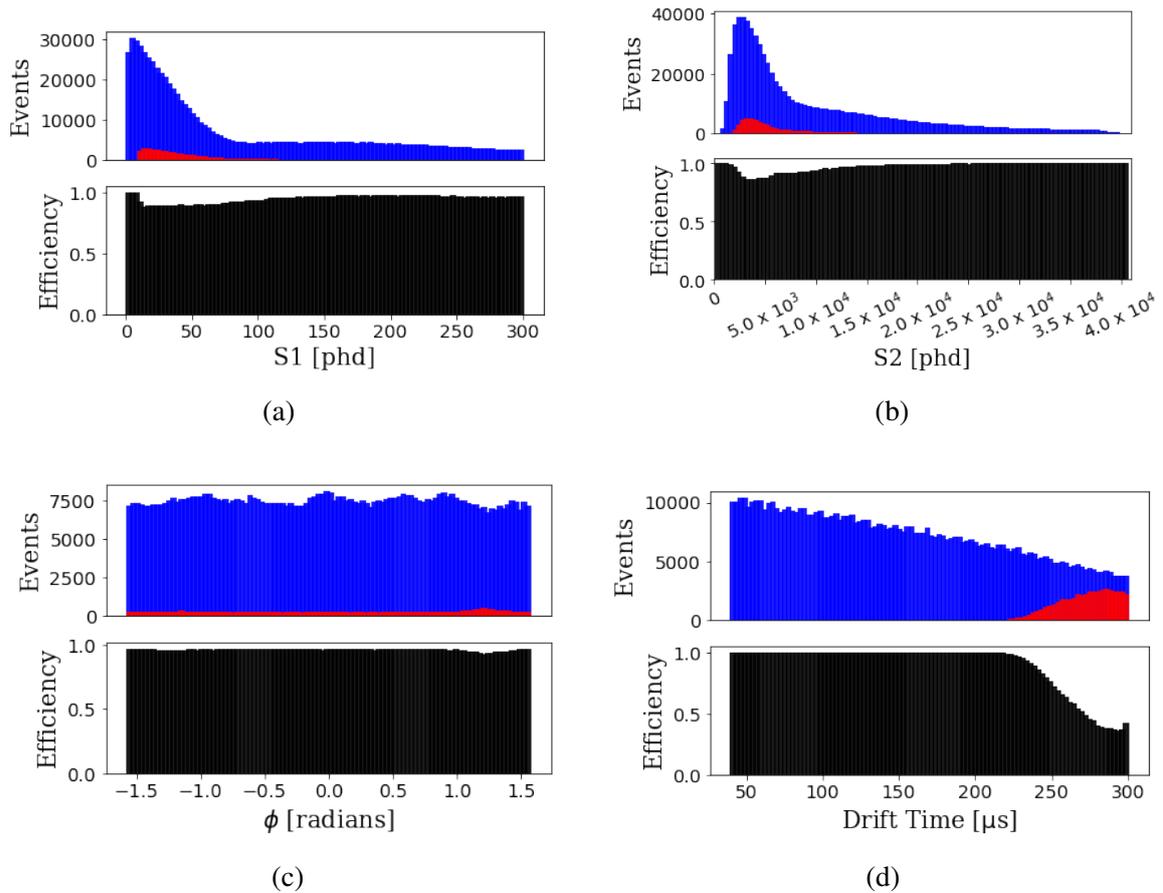


Figure 7.11 Efficiency curves for the BDT cut on a selection of tritium and ^{14}C calibration data. The upper plot in each figure shows all events in blue and removed events in red. The overall efficiency is shown below in black: (a) $S1$, (b) $S2$, (c) ϕ (azimuthal angle), and (d) drift time.

included. Both figures follow the same format. They consist of two plots, with each plot split in two. In the upper portion of these plots are two histograms - a red one overlaid on top of the blue one. The blue plot represents all events, and the red shows events that are removed by the cut. The lower half of each plot shows the overall efficiency of each bin (i.e. red divided by blue). Figures 7.11a and 7.11b show this for $S1$ and $S2$ respectively⁸; Figure 7.11c for ϕ ; and Figure 7.11d for drift time.

Figure 7.11a show a reasonably flat efficiency, albeit is a slight preference to remove events with a smaller $S1$ signal. This is likely due to how CS_{S1} and $MPAF_{S1}$ are simulated - in that smaller $S1$ signals are more likely to be observed by fewer PMTs. So long as $S1$ photo-hit-patterns are a major consideration in gamma-X identification this will be an unavoidable feature. A similar slight preference for removing low $S2$ events is seen in Figure 7.11b. This is likely due to low energy depositions having a low $S1$ and low $S2$ signals.

Figure 7.11c shows that the cut has no preference in removing events based on their ϕ . This is to be expected since neither R nor ϕ was interpreted by the BDT. Nor should they be, since there is no preference in this parameter space for the generation of gamma-X events. However, there is a preference in Z , and this is shown in Figure 7.11d where drift time acts as a proxy for depth. Here we see that the likelihood for an event to be removed by the cut increases steeply beyond a drift time of $\sim 220 \mu\text{s}$. This overzealous removal of events is indicative of how difficult it is to identify a gamma-X event. The only thing one can be confident of is that they will occur close to the cathode.

7.6 Conclusion

By employing the BDT from XGBoost, a gamma-X cut has been developed in a 6-dimensional parameter space. The event features which define this parameter space were identified in Chapter 6 as being the best for distinguishing between gamma-X and SS events. In order to maintain a 95% data acceptance the value which the BDT cut uses to make its binary classification was set to this threshold on calibration data. However, there are two parameter spaces in which events will not be considered gamma-X regardless of their BDT score. The first are events $Z > 20$ cm, as this is considered too far from the cathode to realistically be a gamma-X event. The second are events with $S1 < 11$ phd, since in this parameter space $S1$ light in the bottom PMT array is too likely to be tightly clustered regardless of whether the

⁸The low energy peaks of these distributions are from tritium beta decay, and the long flat tails are from ^{14}C decay

event is SS or not.

Notably absent from this analysis is a measurement of the effectiveness on NR calibration data. The reason for this is that NR calibration data, which comes from Deuterium-Deuterium (DD) neutron generator [122]. This is injected into the TPC via a source tube at the top of the TPC. This results in NR calibration data not being homogeneously distributed throughout the detector, with only a few events occurring near the cathode. As such the data does not provide any useful insight into the effectiveness of gamma-X cuts. However, this is not expected to be a problem since the simulated SS events used to train the BDT were 50% ER and 50% NR⁹.

Before this cut can be employed however, its performance must be compared to the previously developed gamma-X cuts described in Section 6.3. This benchmarking will be the subject of Chapter 8.

⁹So for the BDT to distinguish between ER and NR events would be surprising.

Chapter 8

Gamma-X cut performance on LUX data

This chapter will discuss the benchmarking of gamma-X cuts against several independent datasets. This is done in order to choose an appropriate gamma-X cut to add to the suite of data analysis cuts used by the Run04 EFT search (as seen in Section 6.4). The chapter will then go on to look at the performance of the chosen cut on the LUX Run04 EFT search data.

8.1 Cut benchmarking

This section will discuss the benchmarking of all gamma-X cuts on four independent datasets. These datasets are: the Run04 1S1 2S2 MS data near the bottom of the TPC (a.k.a. nearGX events), Run04 $^{83\text{m}}\text{Kr}$ data, tritium calibration data, and the simulated gamma-X events from Chapter 6. This analysis was used to decide which gamma-X cut will be used by the Run04 EFT search.

Figure 8.1 shows the rejection power of four different cuts on the Run04 1S1 2S2 MS data near the bottom of the TPC. The cuts in question were: the BDT gamma-X cut described in Chapter 7; and the three cuts described in Section 6.3 (two gamma-X cuts and the MPA_{S1} leakage cut). Selection of 1S1 2S2 events involved the application of all Run04 EFT search analysis cuts from Section 6.4, except for the fiducial cut. The selection of these events also varied from previous descriptions of nearGX cuts in that no restriction on the distance between S2 scatters was employed. The application of S2 cuts to MS data involved only applying them to the S2 furthest from the cathode (i.e. the scatter with the highest value of Z). In Figure 8.1a the X-axis gives the maximum height above the cathode for which events were tested, and the Y-axis shows how many events appear within this specified distance from the cathode. In this figure, the solid black line shows how many 1S1 2S2 MS events are selected as the maximum height above cathode defining these events changes, and the 4

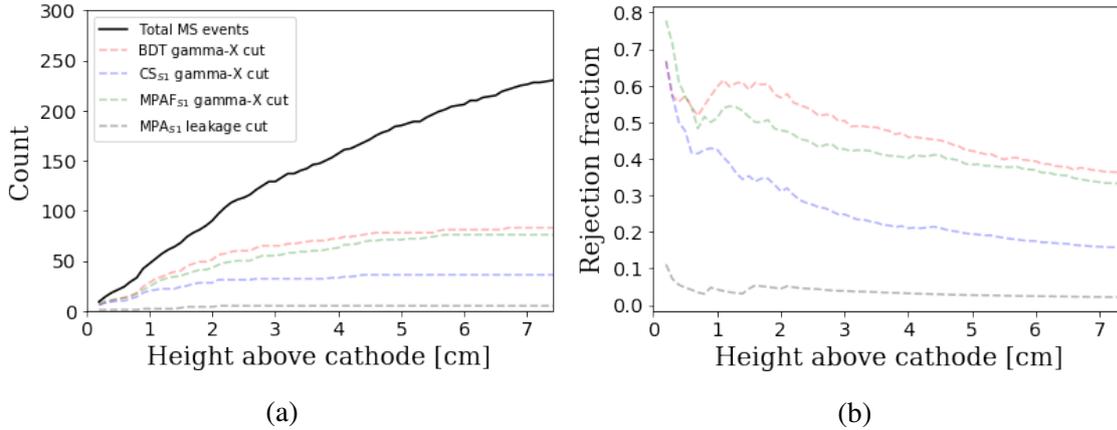


Figure 8.1 The rejection power of various cuts on 1S1 2S2 MS events near the cathode. In (a) the number of events removed by four different cuts is displayed for as the maximum height above the cathode required for nearGX selection is varied. The fraction of all events removed for each cut is displayed in (b). Cuts represented in this figure are: the BDT gamma-X cut (red dashed line), the CS_{S1} gamma-X cut (blue dashed line), the $MPAF_{S1}$ gamma-X cut (green dashed line), and the MPA_{S1} leakage cut (grey dashed line). The total number of events for each height above cathode selection criteria is also shown in (a) (thick black line). Note that the MPA_{S1} leakage cut is included in the data selection of each of the other cuts displayed.

coloured dashed lines show how many of these events are removed by each cut. Figure 8.1b uses the same X-axis, but the Y-axis now shows the fraction of events removed by each cut.

Notably, gamma-X cuts are not designed to remove MS events near the cathode. However, these events are expected to mimic gamma-X events in many ways. For instance, just like gamma-X events they are also expected to: occur close to the cathode (admittedly by definition), have S1-photo-hit-patterns in the bottom PMT array which are tightly clustered, and are thought to largely be created by the same gamma-rays which cause gamma-X events. As such, a good gamma-X cut which operates under the criteria outlined in Section 6.2 should remove a higher proportion of these events than SS events. In fact, a disproportionate removal of these MS events is a good indicator for a gamma-X cut. Based on this standard, the BDT gamma-X cut appears to be the best performing cut in Figure 8.1. At all heights, the BDT cut appears to outperform the others, with the $MPAF_{S1}$ gamma-X cut a close second.

The further away these MS events are from the cathode, the less gamma-X-like they become. Not only is this because the event is occurring higher up in the detector, but also because the S1-photo-hit-pattern will become more dispersed on the bottom PMT array. As such, it

is also considered a good indicator if the gamma-X cut becomes less effective on nearGX events if the maximum height above the cathode used to define these events is increased. This appears to be the case in Figure 8.1 for all gamma-X cuts, where the fraction of rejected events steadily decreases as the height above cathode increases from 1 cm to 7 cm. Due to the low number of events associated with setting such a shallow Z limit for nearGX selection, the rise in rejection fraction between 0.5 cm and 0.9 cm above the cathode shown by all cuts is not statistically significant. As illustrated with the solid black line in Figure 8.1a at this depth there are only between 20 to 40 1S1 2S2 MS events to test this cut on.

NearGX events were not the only dataset used to test these gamma-X cuts however. Illustrated in Figure 8.2 are four plots, each displaying a different cut applied to three other datasets: Run04 $^{83\text{m}}\text{Kr}$ data¹, tritium calibration data, and the full MC simulated gamma-X events introduced in Section 6.5. In Figures 8.2a, and 8.2b tritium and $^{83\text{m}}\text{Kr}$ are represented with blue-black and red-black heatmaps respectively. The simulated gamma-X events are displayed as yellow scatter points so as to more easily make out individual events for this smaller sample. The solid green line in Figure 8.2a shows the CS_{S1} gamma-X cut from Section 6.3.2 (events below the line are removed), and the dashed green line in Figure 8.2b shows the $MPAF_{S1}$ cut from Section 6.3.3 (events above the line are removed). Figure 8.2c combines both SS datasets of $^{83\text{m}}\text{Kr}$ and tritium into a single black scatter plot, with events removed by the MPA_{S1} leakage cut covered with green crosses. A scatter plot approach was taken to best highlight the effectiveness of the cut, since the cut shown relies on Z binning it does not strictly act in MPA_{S1} vs $S1$ space. Thus, it cannot be accurately represented with a single line. However, in this depiction a sense of which gamma-X event are going to be removed can be gathered. Figure 8.2d is the final benchmarking plot and shows the BDT score of each dataset. Once again $^{83\text{m}}\text{Kr}$ and tritium are shown in red and blue respectively, simulated gamma-X in yellow, and the dashed black line shows the BDT gamma-X cut. For each plot in Figure 8.2, the parameter space chosen to display each cut was selected to best demonstrate how the cut works.

Table 8.1 is included which shows the performance of each of the cuts from Figures 8.1 and 8.2. Each dataset from these figures is displayed in the four columns of Table 8.1. The first column is $^{83\text{m}}\text{Kr}$ data from Run04, the second is tritium calibration data, the third is the full MC simulated gamma-X events described in Section 6.5.5, and the last is the 1S1 2S2 MS data from Figure 8.1, which occurred within a 5 cm interval above the cathode. The rows of this table show the cut which is being tested. The first four rows show the MPA_{S1}

¹Only $^{83\text{m}}\text{Kr}$ events with one $S1$ (i.e. not separated) were selected for this analysis.

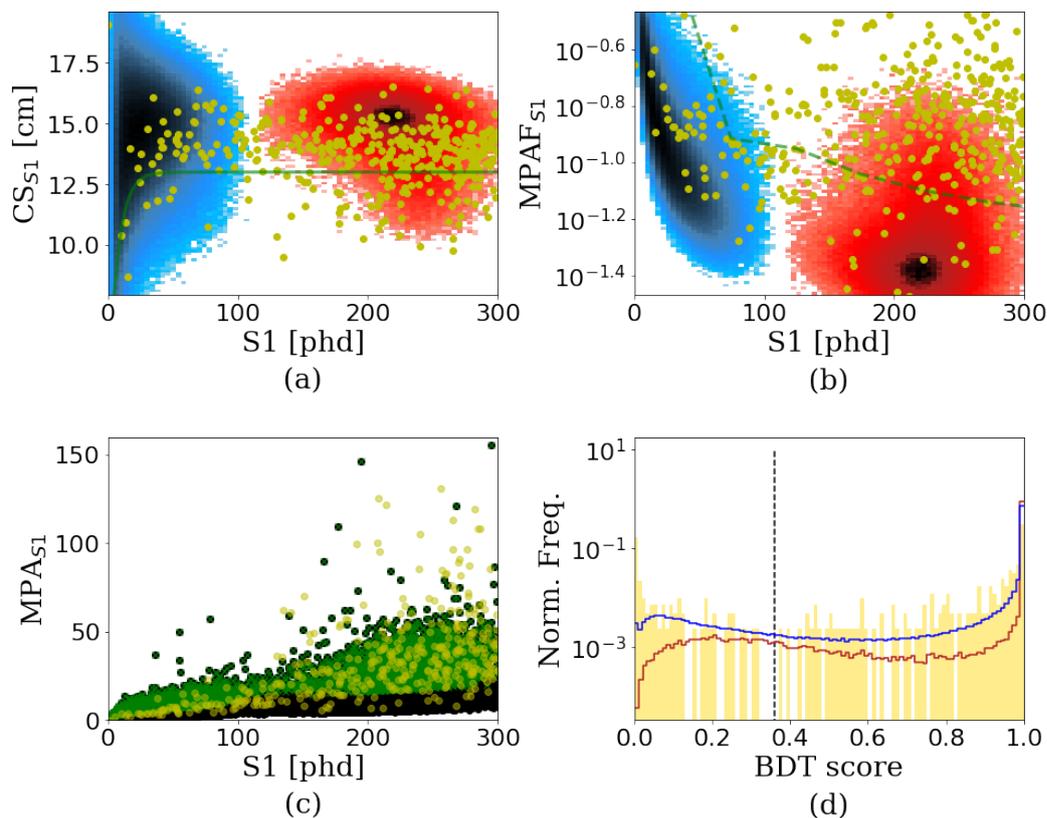


Figure 8.2 Gamma-X cut benchmarking plot. In all sub-figures, blue-black heat maps represent tritium data, red-black represents ^{83m}Kr , and yellow scatter points represent simulated gamma-X events. The cut in (a) is the CS_{S1} gamma-X cut (solid green line) and removed all events below the curve. The cut in (b) is the $MPAF_{S1}$ gamma-X cut (dashed green line) and removes all events above the curve. In (c) tritium and ^{83m}Kr data are combined into one scatter plot (black) and shows events removed by the MPA_{S1} leakage cut in green. In (d) the BDT score used to inform the BDT cut is shown for the same tritium (blue) and ^{83m}Kr (red) data, and simulated gamma-X events (yellow). Events below the dashed black line are removed by the BDT gamma-X cut, so long as they are in the range: $Z < 20$ cm, and $S1 > 11$ phd.

Table 8.1 The fraction of rejected events of each cut on various datasets. Near gamma-X events are the same as from Figure 8.1 where the maximum height above cathode is set to 5 cm. Each cut was measured on datasets after all data analysis cuts from Section 6.4 were applied, except for the leakage cut, and the lower fiducial cut (the ROI cut was also abandoned on the simulated gamma-X events).

Cut name	$^{83\text{m}}\text{Kr}$ data rejection fraction	Tritium data rejection fraction	Simulated γ -X rejection fraction	NearGX rejection fraction
Leakage cut	0.0 ($< 1 \times 10^{-5}$)	0.028 ± 0.007	0.70 ± 0.02	$0.03^{+0.01}_{-0.03}$
CS_{S1} γ -X cut	$0.01^{+0.03}_{-0.01}$	0.131 ± 0.003	0.43 ± 0.02	0.2 ± 0.2
$MPAF_{S1}$ γ -X cut	$0.01^{+0.03}_{-0.01}$	0.01 ± 0.01	0.78 ± 0.02	0.38 ± 0.09
BDT γ -X cut	0.04 ± 0.01	0.044 ± 0.005	0.87 ± 0.02	0.41 ± 0.09
CS_{S1} γ -X + leakage cuts	$0.01^{+0.03}_{-0.01}$	0.144 ± 0.003	0.83 ± 0.02	0.2 ± 0.2
$MPAF_{S1}$ γ -X + leakage cuts	$0.01^{+0.03}_{-0.01}$	0.029 ± 0.006	0.78 ± 0.02	0.38 ± 0.09
BDT γ -X + leakage cuts	0.04 ± 0.01	0.071 ± 0.004	0.91 ± 0.02	0.42 ± 0.09

leakage cut, and the three gamma-X cuts (CS_{S1} , $MPAF_{S1}$, and the BDT). However, since the leakage cut is going to be applied to the EFT data regardless, we are also interested in how each of these gamma-X cuts behaves in conjunction with the leakage cut (which was the best cut able to remove gamma-X events from the suite of cuts already used by the Run04 EFT analysis prior to February 2020). For this reason, the last three rows show these same gamma-X cuts, but with the MPA_{S1} leakage cut also applied.

For $^{83\text{m}}\text{Kr}$ and tritium data the expectation is that a good gamma-X cut will have a high acceptance (since these are SS events). However, for simulated gamma-X events and nearGX data, a high rejection fraction is ideal. When looking at the performance of these cuts without applying the leakage cut it appears that the $MPAF_{S1}$ gamma-X cut is a good choice. It is conservative on SS rejection, with the lowest rejection rate over both $^{83\text{m}}\text{Kr}$ data, and tritium data (excluding the leakage cut); and it performs well on both the full MC simulated gamma-X events, nearGX data. However, the consequence of the $MPAF_{S1}$ gamma-X cut operating in a similar parameter space to the MPA_{S1} leakage cut is that the $MPAF_{S1}$ cut adds little additional gamma-X rejection power above what is already given by the MPA_{S1} leakage cut. This is relevant to the Run04 EFT search since the MPA_{S1} leakage cut is going to be applied regardless of which gamma-X cut was to be chosen. For this reason, the performance of each cut in conjunction with the leakage cut is also given in Table 8.1. These additional rows indicate that the performance of the $MPAF_{S1}$ cut alone, or

in conjunction with the MPA_{S1} leakage cut is essentially the same. However, their performance on tritium data indicates that using both cuts together will increase the chance of mistakenly rejecting SS events.

The fact that the $MPAF_{S1}$ gamma-X cut works in a similar parameter space as the MPA_{S1} leakage cut hurts its performance in the context of the LUX Run04 EFT search. Any gamma-X events which remain after the leakage cut is applied will be harder for the $MPAF_{S1}$ cut to identify. The BDT cut on the other hand works in a multidimensional parameter space. This means other identifying features of a gamma-X event which might distinguish it from a SS can be harnessed for their exclusion. As such, the BDT gamma-X cut can be seen to have the highest rejection rate for both simulated gamma-X events, and nearGX data when also applying the leakage cut. Although it also has a comparably high SS event rejection fraction when looking at ^{83m}Kr or tritium data, these levels fall within the 5% level of tolerance described in Section 7.5.

This analysis concludes that the BDT gamma-X cut is the best choice for the LUX Run04 EFT search. However, it should be stressed that this conclusion is predicated on the application of the MPA_{S1} leakage cut, hence makes this conclusion specific to the Run04 EFT analysis. Without this, the $MPAF_{S1}$ gamma-X cut may very well be the better performing cut in a different context where a lower SS rejection rate is required, and a lower simulated gamma-X rejection rate is acceptable. This fact highlights the strength of the MPA_{S1} parameter in gamma-X classification.

8.2 Gamma-X cut performance on Run04 EFT data

Until this point, the discussion of gamma-X cut performance was restricted to calibration data or simulations. However, a crucial test of any data analysis cut is to measure its effect on the data under analysis. In this case, with a gamma-X cut required for the Run04 EFT search, this was the Run04 EFT data as it appeared in February 2020. This involved applying all data analysis cuts from Section 6.4 to this dataset. Notably, cut development continued beyond February 2020. As a result, this dataset will not have much bearing on the data which is to be presented in the final LUX EFT search paper [95].

Also available for analysis were a set of salt events from the Run04 WIMP search result [47]. Salt events are fake WIMPs which are injected into the data. Analysts are blinded to which events these are until after the analysis has been completed. The rationale behind the use of salt is to ensure that any data analysis cuts which are developed do not inadvertently

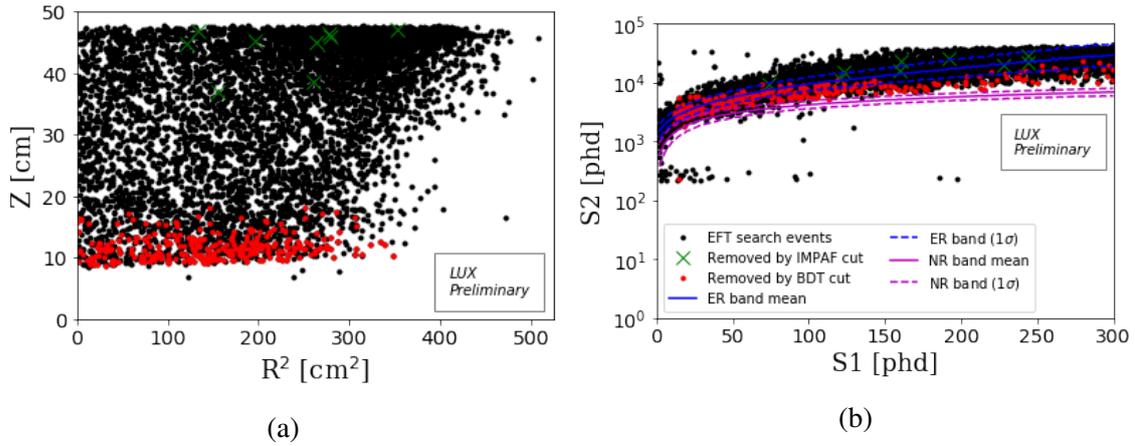


Figure 8.3 BDT cut effect on Run04 EFT search data circa February 2020. In (a) data is displayed in (R^2, Z) space, with removed events in red. In (b) the same data is displayed in $(S1, S2)$ space, with the 1σ NR band displayed in magenta, and the 1σ ER band in blue. For comparison, the events removed by the $MPAF_{S1}$ cut are also displayed with green crosses.

remove a potential signal. This final check of data analysis cuts on salt is known as unsalting. Unsalting for the Run04 WIMP search was performed for events up to $S1 < 50$ phd. This provided a sample of salt event upon which to do a final check of the BDT, CS_{S1} , and $MPAF_{S1}$ gamma-X cuts. For the salt, no data analysis cuts were applied.

The effect of the BDT cut on the Run04 EFT data as it stood in February 2020 can be seen in Figure 8.3. In Figure 8.3a, the R^2 vs. Z distribution of this data is represented in black with removed events shown in red. For comparison events that would be removed by the $MPAF_{S1}$ leakage cut are represented with green crosses. Figure 8.3b shows the same data in $S2$ vs. $S1$ space, alongside the 1σ NR and ER bands in magenta and blue respectively. The effect of the BDT gamma-X cut on salt events is shown in Figure 8.4. Figure 8.4a shows the BDT score of salt events in green, and the BDT cut value as a dashed red line. Also included are the BDT scores of all Run04 EFT search data (grey). Figure 8.4b shows the same salt events in $(S1, S2)$ space. In magenta is the 1σ NR band. Blue crosses represent salt events misidentified as gamma-X by the CS_{S1} cut, and the red crosses are salt events with a BDT score < 0.36 . However, it should be noted that these red crosses are not removed by the BDT cut on account of their $S1$ value. These events range from 1.8 phd to 5.5 phd, which is below the 11 phd threshold required to be considered a gamma-X by the BDT cut. No salt events were tagged as gamma-X by the $MPAF_{S1}$ cut.

The result of this analysis is shown in Table 8.2, which gives the acceptance rate after each of the three gamma-X cuts were applied for Run04 EFT search data, and salt events. The

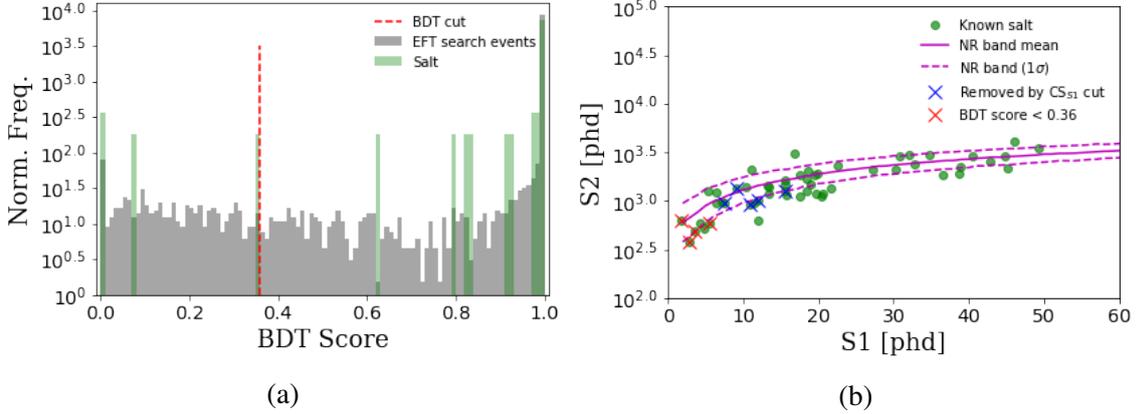


Figure 8.4 BDT cut effect on salt and Run04 EFT search data. In (a) the BDT score of all Run04 EFT search data is displayed in grey, along with salt events in green, and the BDT cut as a dashed red line (events below the line are removed). In (b) are salt events (green) in (S_1 , S_2) space. Blue crosses represent events which would be removed by the CS_{S_1} gamma-X cut. Red crosses show events with a BDT score < 0.36 . Note that these red events are not removed by the BDT cut on account of their low S_1 value.

Table 8.2 Various gamma-X cut performances on Run04 EFT data, and salt events.

	Data acceptance	Salt acceptance
CS_{S_1} cut	0.95 ± 0.02	0.9 ± 0.1
$MPAF_{S_1}$ cut	$1.00^{+0.0}_{-0.02}$	$1.0^{+0.0}_{-0.1}$
BDT cut	0.96 ± 0.02	$1.0^{+0.0}_{-0.1}$

MPA_{S_1} leakage cut does not feature in Table 8.2 since this cut was used in data selection. As a result, it would not remove any additional data. Its effect on salt events is beyond the scope of this thesis and will be described in the Run04 EFT search paper [95].

To determine whether a gamma-X cut is behaving well or not on these datasets we should consider how rarely we expect gamma-X events to occur. Given that 14 nearGX events that were seen in Section 6.5.6, it is reasonable to assume that a similar order of magnitude of true gamma-X events would be present in the data. As of the moment of writing, since nearly 6500 events were present in the Run04 EFT data after data analysis cuts², it is reasonable then to assume the vast majority of these are not gamma-X events. As such a high data acceptance is desirable.

Of the three gamma-X cuts assessed in Table 8.2 the $MPAF_{S_1}$ is by far the most conservative. However, as previously explained the high efficiency of this cut is likely due to the MPA_{S_1}

²EFT data analysis cuts are under development. The final event count is likely to have a different value.

leakage cut being applied in data selection, and having already removed the majority of events that would be excluded by this cut. Evidence of this can be seen in Figure 8.3b, where the events which the $MPAF_{S1}$ cut selects for removal are not only far from the cathode, but are also in the ER band. Not only are the features shown here extremely un-gamma-X-like, but also makes the cut redundant; since all events will be handled by the ER discrimination described in Section 3.1.6. This observed inability of the $MPAF_{S1}$ cut to add additional rejection power in the EFT search data, beyond what is already added by the MPA_{S1} leakage cut, echos the results from Table 8.1. So although the $MPAF_{S1}$ cut is more conservative on data, it is too conservative on gamma-X events to be a useful cut.

Trailing the performance of the $MPAF_{S1}$ cut are the BDT cut with a $96\% \pm 2\%$ acceptance rate, and the CS_{S1} gamma-X cut with an acceptance of $95\% \pm 2\%$. There is so little differentiating the performance of these cuts that the distinction is insignificant. However, the better performance of the BDT cut over the CS_{S1} cut manifests when comparing them based on the rejection of salt events. While the CS_{S1} cut excludes $\sim 10\%$ of these events, the BDT cut matches the performance of the $MPAF_{S1}$ cut with a perfect acceptance rate (where $S1 > 11$ phd and $Z < 20$ cm).

Notably, there are four salt events at $S1 < 6$ phd which have a low enough BDT score to be classified as a gamma-X. However, since the BDT cut does not operate below an $S1 < 11$ phd the cut does not apply. It is unsurprising that the BDT cut performs poorly in this region. Features which depend on $S1$ photo-hit-patterns in the lower PMTs become increasingly meaningless for low $S1$. For a BDT which is trained to remove gamma-X events, as opposed to the other cuts which were designed to maintain a given efficiency, the increased chance for the misidentification of a gamma-X is to be expected. The solution to this is not to apply either the CS_{S1} or the $MPAF_{S1}$ gamma-X cuts though; as these are also developed in parameter spaces which are blind to low $S1$ gamma-X events. As such they are more prone to misidentifying a SS as a gamma-X. In order to identify the extremely rare, and as yet unobserved, case of the low $S1$ gamma-X event a new parameter will need to be identified for its classification. This however does not detract from the effectiveness of existing parameters to classify higher energy gamma-X events.

8.3 Conclusion

A BDT cut was developed for the removal of potential gamma-X events in the Run04 EFT search data. Having been developed in a 6-dimensional parameter space, this cut was found

to outperform all previously developed gamma-X cuts. The performance of several gamma-X cuts (including the BDT cut) were tested on four independent test datasets (as seen in Table 8.1) before being applied to the EFT search data and known salt events (as seen in Table 8.2). The BDT cut appears to outperform the CS_{S1} cut in part due to the latter cut not considering more than two parameters, leading to an overzealous exclusion of salt events. The $MPAF_{S1}$ gamma-X cut on the other hand was outperformed due to a similar cut already existing in the suite of cuts used for the selection of the EFT data, reducing the capacity of the $MPAF_{S1}$ cut to contribute to the overall exclusion of gamma-X events. However, it should be noted that the performance of the $MPAF_{S1}$ gamma-X cut in the absence of this existing cut was quite close to the BDT cut, which speaks to the strength of the MPA_{S1} parameter as a classifier of gamma-X events. Nevertheless, due to the stronger performance of the BDT cut to identify simulated and near-miss gamma-X events, it was approved by the LUX Run04 EFT group to be included in their suite of data analysis cuts.

Chapter 9

Concluding remarks

This thesis describes the contribution of this author to methods for the identification and reduction of backgrounds; specifically for the two dual-phase xenon TPC based experiments of LUX and LZ. This work is introduced in Chapter 2 with an overview of the status of dark matter theory and experimental attempts to detect it, and contextualised in Chapter 3 with a description of the LUX and LZ experiments respectively.

With LZ due to begin its maiden commissioning run later in 2020 (pending any further COVID-19 related delays), it has not produced any data for the duration of this thesis. Nevertheless, contributions towards LZ were possible through the production and analysis of simulations; and the development of data analysis tools. Specifically, as was seen in Chapter 4, the calibration source $^{131\text{m}}\text{Xe}$ was simulated after writing a generator for its decay scheme in BACCARAT. With the simulations this provided, two mechanisms by which this source can produce a background were identified - PEDs, and MSSIs. For the PED backgrounds of the WIMP and EFT searches and the MSSI background for the WIMP search, the effectiveness in the fiducial cut saw all background events removed from the analysis. The null results this produced led to an upper limit on the estimated background of 6.88×10^{-4} for each of these analyses over 1000 days (assuming a 0.02 Hz average $^{131\text{m}}\text{Xe}$ decay rate). The MSSI background for the EFT search on the other hand did not have a null result and was found to have a background rate of $(8 \pm 2) \times 10^{-3}$ over 1000 days at 0.02 Hz $^{131\text{m}}\text{Xe}$ decay rate. This work on LZ was continued in Chapter 5, which describes the development of a module to simulate the path of electrons through the LXe of LZ. This module was then used to improve the realism of simulations, and help write an analysis package to correct the XYZ position of reconstructed events, based on the expected path of S2 electrons from ionisation sites to the phase boundary.

Whilst LUX has been decommissioned for quite some time now, there is still analysis to be performed on its data. For instance, the EFT analysis of Run04 data remains outstanding, which until recently was in part due to a need for an effective identification procedure for gamma-X events. As described in Chapter 6, simulations showed that these gamma-X events have the potential to originate from any detector surface containing a gamma-ray emitting background. However, it was shown that they were overwhelmingly likely to originate from ^{238}U , ^{232}Th , ^{40}K , and ^{60}Co sources in the bottom PMT array. Although previous analysts had made excellent ground in the removal of these events, the gamma-X cuts they had developed were not validated with full MC simulations. The production of these simulations proved valuable not just in the benchmarking of various gamma-X cut against one another, but also in the identification of parameters useful in distinguishing gamma-X from SS events. As described in Chapter 7, these simulations led to the development of a BDT capable of correctly identifying simulated gamma-X events at a rate of $91\% \pm 2\%$; much higher than previous gamma-X cuts. When this BDT was used to develop a cut for the LUX Run04 EFT analysis, as described in Chapter 8, it was also found to outperform previous gamma-X cuts in the removal of nearGX events at a rate of $41\% \pm 9\%$ and fall within the 5% SS event rejection rate of calibration data considered acceptable by the EFT analysis group.

LZ represents the second generation of dark matter searches and builds upon the experience of earlier experiments such as LUX. With multiple analyses planned for LZ's data, each based on a different theoretical model for dark matter, excitement for the potential of observing the interaction between dark matter and baryonic matter is high. In the event of a positive dark matter detection, thoroughly verified confirmation from collider, indirect, and other direct detection experiments will be required. This will allow the third generation of dark matter searches to probe its nature in greater detail. Alternatively, if LZ were to report another null result, the electro-weak parameter space for dark matter models such as the WIMP would be more tightly confined. This in itself may lead to the development of new theories governing the nature of dark matter. For over a quarter-century the dark matter community has eagerly awaited a verified signal of this kind. As we step into an age of unprecedented sensitivity to the potential of such a dark matter signal, all eyes fall upon LZ to see what we might just discover.

References

- [1] D. Akerib *et al.*, “The Large Underground Xenon (LUX) experiment,” *Nuclear Instruments and Methods in Physics Research*, vol. A704, p. 111, 2013.
- [2] B. Mount *et al.*, “LUX-ZEPLIN (LZ) Technical Design Report,” 3 2017.
- [3] H. S. Kragh, *Conceptions of Cosmos: From Myths to the Accelerating Universe: A History of Cosmology: From Myths to the Accelerating Universe: A History of Cosmology*, vol. 1 of *illustrated, reprint*. Oxford UK: OUP Oxford, 2006.
- [4] F. Zwicky, “The redshift of extragalactic nebulae,” *Helvetica Physica Acta*, vol. 6, pp. 110–127, 1933.
- [5] E. Hubble and M. L. Humason, “The Velocity-Distance Relation among Extra-Galactic Nebulae,” *The Astrophysical Journal*, vol. 74, pp. 43–80, 1931.
- [6] V. C. Rubin and W. K. Ford Jr., “Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions,” *The Astrophysical Journal*, vol. 159, pp. 379–408, 1970.
- [7] K. G. Begeman, A. H. Broeils, and R. H. Sanders, “Extended rotation curves of spiral galaxies: dark haloes and modified dynamics,” *Monthly Notices of the Royal Astronomical Society*, vol. 249.3, p. 523–537, 1991.
- [8] M. Milgrom, “A Modification of the Newtonian dynamics as a possible alternative to the hidden mass hypothesis,” *The Astrophysical Journal*, vol. 270, pp. 365–370, 1983.
- [9] A. Aguirre, J. Schaye, and E. Quataert, “Problems for MOND in Clusters and the Ly- α Forest,” *The Astrophysical Journal*, vol. 561, pp. 550–558, 2001.
- [10] F. Sánchez-Salcedo, E. Martínez-Gómez, V. Aguirre-Torres, and H. Hernández-Toledo, “Low-mass disc galaxies and the issue of stability: MOND versus dark matter,” *Monthly Notices of the Royal Astronomical Society*, vol. 462, p. 3918–3936, 2016.
- [11] F. Iocco, M. Pato, and G. Bertone, “Testing modified Newtonian dynamics in the Milky Way,” *Phys. Rev. D*, vol. 92, no. 8, p. 084046, 2015.
- [12] P. Muirhead, “Using gravitational lensing to measure stellar properties,” *SPIENewsroom*.
- [13] H. Hoekstra, M. Bartelmann, H. Dahle, *et al.*, “Masses of Galaxy Clusters from Gravitational Lensing,” *Space Sci Rev*, vol. 177, p. 75–118, 2013.

- [14] Y. Nagashima, *Beyond the standard model of elementary particle physics*. Weinheim, USA: Wiley-VCH, 2014.
- [15] M. Markevitch, “Chandra observation of the most interesting cluster in the universe,” *ESA Spec. Publ.*, vol. 604, p. 723, 2006.
- [16] D. Clowe, M. Bradac, A. H. Gonzalez, M. Markevitch, S. W. Randall, C. Jones, and D. Zaritsky, “A direct empirical proof of the existence of dark matter,” *The Astrophysical Journal Lett.*, vol. 648, pp. L109–L113, 2006.
- [17] Y. Akrami *et al.*, “Planck 2018 results: I. Overview and the cosmological legacy of Planck,” *arXiv*, vol. 1807.06205, p. 615, 2018.
- [18] S. Dodelson., *Modern Cosmology*. San Diego: Academic Press, 2003.
- [19] C. Baugh, “Correlation Function and Power Spectra in Cosmology,” *Encyclopedia of Astronomy Astrophysics*, 2007.
- [20] P. Tisserand *et al.*, “Limits on the Macho Content of the Galactic Halo from the EROS-2 Survey of the Magellanic Clouds,” *Astron. Astrophys.*, vol. 469, pp. 387–404, 2007.
- [21] S. D. M. White, C. S. Frenk, and M. Davis, “Clustering in a neutrino-dominated universe,” *The Astrophysical Journal*, vol. 274, p. 1–5, 1983.
- [22] C. Wu, E. Ambler, R. Hayward, D. Hoppes, and R. Hudson, “Experimental Test of Parity Conservation in β Decay,” *Phys. Rev.*, vol. 105, pp. 1413–1414, 1957.
- [23] A. Aguilar-Arevalo *et al.*, “Significant Excess of ElectronLike Events in the Mini-BooNE Short-Baseline Neutrino Experiment,” *Phys. Rev. Lett.*, vol. 121, no. 22, p. 221801, 2018.
- [24] A. Boyarsky, M. Drewes, T. Lasserre, S. Mertens, and O. Ruchayskiy, “Sterile neutrino dark matter,” *Progress in Particle and Nuclear Physics*, vol. 104, pp. 1 – 45, 2019.
- [25] G. Bertone and D. Hooper, “History of dark matter,” *Reviews of Modern Physics*, vol. 90, Oct 2018.
- [26] T. Flacke, D. W. Kang, K. Kong, G. Mohlabeng, and S. C. Park, “Electroweak Kaluza-Klein dark matter,” *JHEP*, vol. 4, p. 41, 2017.
- [27] E. W. Kolb, D. J. Chung, and A. Riotto, “WIMPzillas!,” *AIP Conf. Proc.*, vol. 484, no. 1, pp. 91–105, 1999.
- [28] T. Mannel, “Theory and phenomenology of CP violation,” *Nucl. Phys. B Proc. Suppl.*, vol. 167, pp. 115–119, 2007.
- [29] L. D. Duffy and K. van Bibber, “Axions as dark matter particles,” *New Journal of Physics*, vol. 11, p. 105008, oct 2009.
- [30] M. Tanabashi *et al.*, “Review of Particle Physics,” *Phys. Rev. D*, vol. 98.3, 2018.

- [31] M. Archidiacono, S. Hannestad, A. Mirizzi, G. Raffelt, and Y. Y. Y. Wong, “Axion hot dark matter bounds after Planck,” *J. Cosmol. Astropart. Phys.*, vol. 1310, 2013.
- [32] L. J. Rosenberg and K. A. van Bibber, “Searches for invisible axions,” *Phys. Rep.*, vol. 325, pp. 1–39, 2000.
- [33] M. W. Goodman and E. Witten, “Detectability of Certain Dark Matter Candidates,” *Phys. Rev. D*, vol. 31, p. 3059, 1985.
- [34] C. Savage, G. Gelmini, P. Gondolo, and K. Freese, “Compatibility of DAMA/LIBRA dark matter detection with other searches,” *JCAP*, vol. 04, p. 010, 2009.
- [35] J. I. Read, “The Local Dark Matter Density,” *J. Phys. G: Nuc. Phys.*, vol. 41, 2014.
- [36] A. G. A. Brown *et al.*, “GAIA Data Release 2. Summary of the contents and survey properties,” *Astronom. and Astrophys.*, vol. 616, p. A1, 2018.
- [37] A. M. Green, “Astrophysical uncertainties on the local dark matter distribution and direct detection experiments,” *arXiv*, vol. 1703.10102v, 2017.
- [38] M. C. Smith *et al.*, “The RAVE Survey: Constraining the Local Galactic Escape Speed,” *Mon. Not. R. Astron. Soc.*, vol. 379, p. 755–772, 2007.
- [39] A. H. Peter, V. Gluscevic, A. M. Green, B. J. Kavanagh, and S. K. Lee, “WIMP physics with ensembles of direct-detection experiments,” *Phys. Dark Univ.*, vol. 5–6, pp. 45–74, 2014.
- [40] J. Lewin and P. Smith, “Review of mathematics, numerical factors, and corrections for dark matter experiments based on elastic nuclear recoil,” *Astroparticle Physics*, vol. 6, no. 1, pp. 87 – 112, 1996.
- [41] E. Aprile *et al.*, “Dark Matter Search Results from a One Ton-Year Exposure of XENON1T,” *Phys. Rev. Lett.*, vol. 121, no. 11, p. 111302, 2018.
- [42] V. Chepel and H. Araújo, “Liquid noble gas detectors for low energy particle physics,” *Journal of Instrumentation*, vol. 8, pp. R04001–R04001, apr 2013.
- [43] R. Agnese *et al.*, “Projected Sensitivity of the SuperCDMS SNOLAB experiment,” *Phys. Rev. D*, vol. 95, no. 8, p. 082002, 2017.
- [44] F. Petricca *et al.*, “First results on low-mass dark matter from the CRESST-III experiment,” *J. Phys. Conf. Ser.*, vol. 1342, no. 1, p. 012076, 2020.
- [45] J. Angle *et al.*, “First Results from the XENON10 Dark Matter Experiment at the Gran Sasso National Laboratory,” *Phys. Rev. Lett.*, vol. 100, p. 021303, 2008.
- [46] L. D. Duffy and K. van Bibber, “Axions as dark matter particles,” *New Journal of Physics*, vol. 11, p. 105008, oct 2009.
- [47] D. Akerib *et al.*, “Results from a search for dark matter in the complete LUX exposure,” *Phys. Rev. Lett.*, vol. 118, no. 2, p. 021303, 2017.

- [48] D. Akerib *et al.*, “Projected WIMP sensitivity of the LUX-ZEPLIN dark matter experiment,” *Phys. Rev. D*, vol. 101, no. 5, p. 052002, 2020.
- [49] X. Cui *et al.*, “Dark Matter Results From 54-Ton-Day Exposure of PandaX-II Experiment,” *Phys. Rev. Lett.*, vol. 119, no. 18, p. 181302, 2017.
- [50] J. Aalbers *et al.*, “DARWIN: towards the ultimate dark matter detector,” *JCAP*, vol. 11, p. 017, 2016.
- [51] K. Abe *et al.*, “A direct dark matter search in XMASS-I,” *Phys. Lett. B*, vol. 789, pp. 45–53, 2019.
- [52] P. Agnes *et al.*, “Low-Mass Dark Matter Search with the DarkSide-50 Experiment,” *Phys. Rev. Lett.*, vol. 121, p. 081307, Aug 2018.
- [53] R. Ajaj *et al.*, “Search for dark matter with a 231-day exposure of liquid argon using DEAP-3600 at SNOLAB,” *Phys. Rev. D*, vol. 100, no. 2, p. 022004, 2019.
- [54] R. Agnese *et al.*, “New Results from the Search for Low-Mass Weakly Interacting Massive Particles with the CDMS Low Ionization Threshold Experiment,” *Phys. Rev. Lett.*, vol. 116, no. 7, p. 071301, 2016.
- [55] F. Ruppin, J. Billard, E. Figueroa-Feliciano, and L. Strigari, “Complementarity of dark matter detectors in light of the neutrino background,” *Phys. Rev. D*, vol. 90, p. 083510, Oct 2014.
- [56] R. Bernabei *et al.*, “DAMA/LIBRA-phase1 results and perspectives of the phase2,” *EPJ Web Conf.*, vol. 95, p. 03001, 2015.
- [57] G. Aad, T. Abajyan, B. Abbott, J. Abdallah, S. Abdel Khalek, A. Abdelalim, O. Abdinov, R. Aben, B. Abi, M. Abolins, and *et al.*, “Observation of a new particle in the search for the standard model higgs boson with the atlas detector at the lhc,” *Physics Letters B*, vol. 716, p. 1–29, Sep 2012.
- [58] A. Boveia and C. Doglioni, “Dark matter searches at colliders,” *Annual Review of Nuclear and Particle Science*, vol. 68, p. 429–459, Oct 2018.
- [59] M. Aaboud *et al.*, “Search for dark matter and other new phenomena in events with an energetic jet and large missing transverse momentum using the ATLAS detector,” *JHEP*, vol. 01, p. 126, 2018.
- [60] L. E. Strigari, “Galactic Searches for Dark Matter,” *Phys. Rept.*, vol. 531, pp. 1–88, 2013.
- [61] G. Steigman, B. Dasgupta, and J. F. Beacom, “Precise Relic WIMP Abundance and its Impact on Searches for Dark Matter Annihilation,” *Phys. Rev. D*, vol. 86, p. 023506, 2012.
- [62] J. M. Gaskins, “A review of indirect searches for particle dark matter,” *Contemp. Phys.*, vol. 57, no. 4, pp. 496–525, 2016.
- [63] M. Ackermann *et al.*, “Fermi LAT Search for Dark Matter in Gamma-ray Lines and the Inclusive Photon Spectrum,” *Phys. Rev. D*, vol. 86, p. 022002, 2012.

- [64] A. Abramowski *et al.*, “Search for a Dark Matter annihilation signal from the Galactic Center halo with H.E.S.S.,” *Phys. Rev. Lett.*, vol. 106, p. 161301, 2011.
- [65] S. Desai *et al.*, “Search for dark matter WIMPs using upward through-going muons in Super-Kamiokande,” *Phys. Rev. D*, vol. 70, p. 083523, 2004. [Erratum: *Phys. Rev. D* 70, 109901 (2004)].
- [66] M. Aartsen *et al.*, “Improved limits on dark matter annihilation in the Sun with the 79-string IceCube detector and implications for supersymmetry,” *JCAP*, vol. 04, p. 022, 2016.
- [67] S. Adrian-Martinez *et al.*, “Limits on Dark Matter Annihilation in the Sun using the ANTARES Neutrino Telescope,” *Phys. Lett. B*, vol. 759, pp. 69–74, 2016.
- [68] G. Wikstrom and J. Edsjo, “Limits on the WIMP-nucleon scattering cross-section from neutrino telescopes,” 2009.
- [69] M. Boudaud, S. Aupetit, S. Caroff, A. Putze, G. Belanger, Y. Genolini, C. Goy, V. Poireau, V. Poulin, S. Rosier, and et al., “A new look at the cosmic ray positron fraction,” *Astronomy Astrophysics*, vol. 575, p. A67, Feb 2015.
- [70] P. D. Serpico, “Astrophysical models for the origin of the positron ‘excess’,” *Astropart. Phys.*, vol. 39-40, pp. 2–11, 2012.
- [71] D. Hooper, I. Cholis, T. Linden, and K. Fang, “HAWC Observations Strongly Favor Pulsar Interpretations of the Cosmic-Ray Positron Excess,” *Phys. Rev. D*, vol. 96, no. 10, p. 103013, 2017.
- [72] A. Fitzpatrick, W. Haxton, E. Katz, N. Lubbers, and Y. Xu, “The Effective Field Theory of Dark Matter Direct Detection,” *JCAP*, vol. 02, p. 004, 2013.
- [73] N. A. Larson, *An Effective Field Theory Analysis of the First LUX Dark Matter Search*. PhD thesis, 2016.
- [74] D. S. Akerib *et al.*, “An Effective Field Theory Analysis of the First LUX Dark Matter Search,” *arXiv: Cosmology and Nongalactic Astrophysics*, 2016.
- [75] E. Aprile and T. Doke, “Liquid xenon detectors for particle physics and astrophysics,” *Rev. Mod. Phys.*, vol. 82, pp. 2053–2097, Jul 2010.
- [76] A. Hitachi, T. Takahashi, N. Funayama, K. Masuda, J. Kikuchi, and T. Doke, “Effect of ionization density on the time dependence of luminescence from liquid argon and xenon,” *Phys. Rev. B*, vol. 27, pp. 5279–5285, 1983.
- [77] C. Dahl, *The physics of background discrimination in liquid xenon, and first results from Xenon10 in the hunt for WIMP dark matter*. PhD thesis, 2009.
- [78] P. Sorensen and C. E. Dahl, “Nuclear recoil energy scale in liquid xenon with application to the direct detection of dark matter,” *Phys. Rev. D*, vol. 83, p. 063501, 2011.
- [79] E. Aprile *et al.*, “Design and Performance of the XENON10 Dark Matter Experiment,” *Astropart. Phys.*, vol. 34, pp. 679–698, 2011.

- [80] V. Lebedenko *et al.*, “Result from the First Science Run of the ZEPLIN-III Dark Matter Search Experiment,” *Phys. Rev. D*, vol. 80, p. 052010, 2009.
- [81] D. Akerib *et al.*, “The Large Underground Xenon (LUX) Experiment,” *Nucl. Instrum. Meth. A*, vol. 704, pp. 111–126, 2013.
- [82] D. Akerib *et al.*, “Tritium calibration of the LUX dark matter experiment,” *Phys. Rev. D*, vol. 93, no. 7, p. 072009, 2016.
- [83] D. Akerib *et al.*, “Technical Results from the Surface Run of the LUX Dark Matter Experiment,” *Astropart. Phys.*, vol. 45, pp. 34–43, 2013.
- [84] A. Manalaysay, “Dark-matter results from 332 new live days of LUX data,” *11th International Identification of Dark Matter Conference*, 2016.
- [85] D. Akerib *et al.*, “3D Modeling of Electric Fields in the LUX Detector,” *JINST*, vol. 12, no. 11, p. P11022, 2017.
- [86] D. Woodward, *Simulations of cosmic muons and background radiations for muon tomography and underground experiments*. PhD thesis, 2017.
- [87] A. Bolozdynya, P. Brusov, T. Shutt, C. Dahl, and J. Kwong, “A chromatographic system for removal of radioactive Kr-85 from xenon,” *Nucl. Instrum. Meth. A*, vol. 579, pp. 50–53, 2007.
- [88] C. Zhang, D. M. Mei, V. Kudryavtsev, and S. Fiorucci, “Cosmogenic Activation of Materials Used in Rare Event Search Experiments,” *Astropart. Phys.*, vol. 84, pp. 62–69, 2016.
- [89] “Luxsim: A component-centric approach to low-background simulations,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 675, pp. 63 – 77, 2012.
- [90] S. Agostinelli *et al.*, “GEANT4—a simulation toolkit,” *Nucl. Instrum. Meth. A*, vol. 506, pp. 250–303, 2003.
- [91] M. Szydagis, N. Barry, K. Kazkaz, J. Mock, D. Stolp, M. Sweany, M. Tripathi, S. Uvarov, N. Walsh, and M. Woods, “NEST: A Comprehensive Model for Scintillation Yield in Liquid Xenon,” *JINST*, vol. 6, p. P10002, 2011.
- [92] D. Akerib *et al.*, “Improved modeling of β electronic recoils in liquid xenon using LUX calibration data,” *JINST*, vol. 15, no. 02, p. T02007, 2020.
- [93] N. N. D. Center, “NuDat 2.” <https://www.nndc.bnl.gov/nudat2/chartNuc.jsp>. Accessed: 2018-09-30.
- [94] D. Akerib *et al.*, “Radiogenic and Muon-Induced Backgrounds in the LUX Dark Matter Detector,” *Astropart. Phys.*, vol. 62, pp. 33–46, 2015.
- [95] D. Akerib *et al.*, “Constraints on Effective Field Theory Couplings Using the Full LUX Exposure,” *Preprint*, 2020.

- [96] F. R. Unterweger, M.P., “Update Of NIST Half-Life Results Corrected For Ionization Chamber Source-Holder Instability,” *Applied Radiation and Isotopes*, vol. 87, pp. 92–94, 2014.
- [97] A. Ringbom, “Calculation of electron-photon coincidence decay of ^{131m}Xe and ^{133m}Xe including atomic relaxation,” *Applied Radiation and Isotopes*, vol. 70, pp. 1499–1508, 2012.
- [98] T. Kibédi, T. Burrows, M. Trzhaskovskaya, P. Davidson, and C. Nestor, “Evaluation of theoretical conversion coefficients using bricc,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 589, no. 2, pp. 202 – 229, 2008.
- [99] M. F. L’ANNUNZIATA, “1 - nuclear radiation, its interaction with matter and radioisotope decay,” in *Handbook of Radioactivity Analysis (Second Edition)* (M. F. L’Annunziata, ed.), pp. 1 – 121, San Diego: Academic Press, second edition ed., 2003.
- [100] O. Duparc, “Pierre Auger – Lise Meitner: Comparative contributions to the Auger effect,” *International Journal of Materials Research*, vol. 100, no. 9, pp. 1162–1166, 2009.
- [101] A. Kaboth. Private Communication, 2018.
- [102] M. Berger, J. Hubbell, S. Seltzer, J. Chang, J. Coursey, R. Sukumar, D. Zucker, and K. Olsen, *XCOM: Photon Cross Section Database (version 1.5)*. 100 Bureau Drive, Stop 1070, Gaithersburg, MD 20899-1070.: National Institute of Standards and Technology, 2010.
- [103] G. Feldman and R. Cousins, “Unified approach to the classical statistical analysis of small signals,” *Physical Review*, vol. 57(7), p. 3873, 1998.
- [104] R. Watson. Private Communication, 2020.
- [105] “COMSOL.” <https://uk.comsol.com/>. Accessed: 2020-02-12.
- [106] J. B. Albert *et al.*, “Measurement of the drift velocity and transverse diffusion of electrons in liquid xenon with the EXO-200 detector,” *Phys. Rev. C*, vol. 95, p. 025502, Feb 2017.
- [107] L. S. Miller, S. Howe, and W. E. Spear, “Charge Transport in Solid and Liquid Ar, Kr, and Xe,” *Phys. Rev.*, vol. 166, pp. 871–878, Feb 1968.
- [108] E. M. Gushchin, A. A. Kruglov, and I. M. Obodovskil, “Electron dynamics in condensed argon and xenon,” *Zh. Eksp. Teor. Fi*, vol. 82, pp. 1114–1125, Apr 1981.
- [109] P. Sorensen, “Anisotropic diffusion of electrons in liquid xenon with application to improving the sensitivity of direct dark matter searches,” *Nucl. Instrum. Meth. A*, vol. 635, pp. 41–43, 2011.
- [110] E. Aprile *et al.*, “Analysis of the XENON100 Dark Matter Search Data,” *Astropart. Phys.*, vol. 54, pp. 11–24, 2014.

- [111] D. S. Akerib *et al.*, “First Results from the LUX Dark Matter Experiment at the Sanford Underground Research Facility,” *Phys. Rev. Lett.*, vol. 112, p. 091303, Mar 2014.
- [112] V. Solovov. Private Communication, 2019.
- [113] D. Tiedt, *The Active Veto of the Large Underground Xenon Detector and the Study of Muons Deep Underground*. PhD thesis, South Dakota School of Mines and Technology, 2018.
- [114] B. Lenardo, *Measurements and Modeling of Low Energy Nuclear Recoils in Liquid Xenon for Dark Matter and Neutrino Detection*. PhD thesis, 2018.
- [115] E. Aprile *et al.*, “Observation of two-neutrino double electron capture in ^{124}Xe with XENON1T,” *Nature*, vol. 568, no. 7753, pp. 532–535, 2019.
- [116] A. Lindote. Private Communication, 2019.
- [117] S. Alsum. Private Communication, 2019.
- [118] M. Szydagis *et al.*, “Enhancement of NEST Capabilities for Simulating Low-Energy Recoils in Liquid Xenon.,” *Journal of Instrumentation*, vol. 8(10), p. C10003, 2013.
- [119] “Xgboost documentation.” <https://xgboost.readthedocs.io/en/latest/>. Accessed: 2019-09-30.
- [120] G. Rischbieter. Private Communication, 2019.
- [121] T. Arnold and T. Emerson, “Nonparametric Goodness-of-Fit Test for Discrete Null Distributions,” *The R. Journal*, vol. 3-2, pp. 34–39, 2011.
- [122] D. Akerib *et al.*, “Low-energy (0.7-74 keV) nuclear recoil calibration of the LUX dark matter experiment using D-D neutron scattering kinematics,” *arxiv*, vol. 1608.05381, 8 2016.