

Improving the Low Energy Searches of the LUX-ZEPLIN Experiment with Advanced Waveform Simulation

Nicholas Fieldhouse

St. Peter's College



University of Oxford

A thesis presented for the degree of

Doctor of Philosophy

Trinity 2026

For my parents.

Everything grand is made from a series of ugly little moments. Everything worthwhile by hours of self-doubt and days of drudgery. All the works by people you and I admire sit atop a foundation of failures.

— Pierce Brown

Acknowledgements

My time at Oxford has taught me many things about physics, the world and also myself. Throughout the last four years, I've had experiences that I could have only ever dreamed of without this opportunity. So first and foremost, I would like to thank my supervisor, Kimberly Palladino, for accepting me into your group and providing valuable oversight for the last four years.

To the Oxford group, Amy, Niamh, Aiham, Dan, Josh, Sparshita, Kevin, Wenbo, Antoine and Noah, thank you for all of your guidance and friendship. Dan, we have had endless conversations that have inspired new ideas and really pushed the boundaries of the quality of my work. Kevin and Noah, thank you for making the final year so enjoyable. You have both helped immensely both in and out of work. Josh, thank you for being a constant force of whimsy and energy that fuels everyone around you. And finally, Sparshita, thank you for being the perfect example of both an academic weapon and a great friend. Some of my fondest memories in Oxford are those cheeky post-gym kebab nights. I look forward to many more adventures across the pond.

I would also like to thank my LTA supervisor, Theresa Fruth, for allowing me to work with you and your group. I am extremely grateful for your guidance throughout the start of my PhD through to my time in Sydney. To the group in Sydney: Sharry, Chiara, Christian, Josh, Anthony, Ellen, Ciaran and Elden, thank you all for such a warm welcome and for making me feel like I had a second family so far away from home. I wish you all the best and hope to see you all again someday. To everyone who I shared time with in South Dakota: George, Ellie, Linda, Olivia, Gavin, Harvey and Nathan, thank you for making my time there so special. Whether we were in a mine or up a rock face, you managed to turn this smorgasbord of people into a tight-knit community.

To my oldest friends, Llama and James, thank you for always being there when I needed it, either for a distraction or for giving me advice in my car till 4 in the morning. I am deeply appreciative of everything you have done for me. To Josie, thank you for your endless patience and for inspiring me in new ways every day. I sincerely hope you know how grateful I am.

And finally, to my family, thank you for listening to me yap about physics all of those times and for never once doubting me even when I doubted myself. I could not have done this without you building me up to get here for all those years. *Thank you.*

Declaration

I, Nicholas Fieldhouse, declare that the work presented in this thesis is my own, and that any sources of information have been acknowledged by means of references.



Friday 2nd October, 2026

Signature *Date*

Abstract

The LUX-ZEPLIN (LZ) experiment is a dual-phase xenon time projection chamber primarily designed to search for weakly interacting massive particles (WIMPs) directly through their scattering interactions with xenon nuclei. LZ sets the current world-leading limits for WIMP masses above $5\,\mathrm{GeV/c^2}$ with the strongest spin-independent exclusion of $2.2 \times 10^{-48}\,\mathrm{cm^2}$ at the 90% confidence level, and a best median sensitivity of $5.1 \times 10^{-48}\,\mathrm{cm^2}$ for a $40\,\mathrm{GeV/c^2}$ mass WIMP. The LZ exclusion limit extends down to WIMP masses of $3\,\mathrm{GeV/c^2}$, limited by the analysis threshold set by the charge signal produced during WIMP-nucleon scattering interactions, a limit of $3.5\,\mathrm{electrons}$. LZ has sensitivity down to single electrons, limited by our understanding of low-energy detector effects and background-rejection techniques.

This thesis develops waveform-level simulation and analysis techniques for low-energy searches in the LZ experiment. Simulations in LZ are split into the fast chain, which is the most widely used and produces only the analysis variables without the waveforms, and the full chain, which provides all the analysis variables in addition to the waveforms. Prior to the work presented in this thesis, the waveform simulations had not been used routinely for analysis. In this thesis, the full chain simulation is improved through revised models and implementation. These developments substantially improve the agreement between simulated secondary proportional scintillation pulse (S2) pulses and calibration data for the observables used in the analysis. This tuned simulation is then used to characterise the effects of pulse fragmentation on the S2, where diffusion causes small S2 pulses to be reconstructed as multiple pulses, thereby biasing reconstructed pulse areas and widths. Methods are proposed to model signal loss, reject affected events, and reconstruct the original S2 from its fragments, evaluating the impact through changes in the ^{8}B CEνNS discovery significance.

Finally, waveform-based machine learning methods are investigated for S2-only searches. Neural networks are utilised to study background classification and drift-position prediction. UMAP clustering is used for anomaly detection and background rejection applications. Together, these studies provide improved tools for low-threshold dark matter searches in LZ and future xenon experiments.

Table of Contents

List of Abbreviations **vi**

List of Figures **xxiii**

List of Tables **xxiv**

Introduction **1**

1 Dark Matter **3**

1.1 Experimental Evidence for Dark Matter . . . 3

1.1.1 Galactic Motion . . . 3

1.1.2 Galaxy Rotation Curves . . . 4

1.1.3 Gravitational Lensing . . . 5

1.1.4 Cosmic Microwave Background . . . 6

1.1.5 Properties of Dark Matter . . . 7

1.2 Dark Matter Candidates . . . 7

1.2.1 Modified Newtonian Dynamics . . . 7

1.2.2 Massive Astrophysical Compact Halo Objects . . . 7

1.2.3 Axions . . . 8

1.2.4 Sterile Neutrinos . . . 9

1.2.5 Weakly Interacting Massive Particles . . . 9

1.3 Dark Matter Detection Methods . . . 10

1.3.1 Indirect Detection . . . 11

1.3.2 Collider Production . . . 12

1.3.3 Direct Detection . . . 12

1.3.3.1 Recoil Kinematics . . . 13

1.3.3.2 WIMP Parameter Space . . . 14

2 The LUX-ZEPLIN Experiment **16**

2.1 The LZ Detector . . . 17

2.2 Signal Model . . . 18

2.3 Calibrations . . . 21
2.4 Backgrounds . . . 24
2.5 Bias Mitigation . . . 25
2.6 WS2022 + WS2024 Results . . . 26
2.7 ^{8}B CEνNS/Low Mass WIMP Results . . . 26
2.8 Towards XLZD . . . 28

3 LZ Simulation Packages 31
3.1 Full Chain Simulation . . . 32
3.1.1 BACCARAT . . . 33
3.1.1.1 GEANT4 . . . 33
3.1.1.2 Physics List . . . 33
3.1.1.3 Event Process . . . 34
3.1.1.4 Clustering . . . 35
3.1.1.5 Particle Propagation . . . 36
3.1.1.6 BaccMCTruth . . . 37
3.1.2 Detector Electronics Response (DER) . . . 37
3.1.2.1 SPE Modelling . . . 38
3.1.2.2 Secondary PMT Effects . . . 40
3.1.2.3 Data Acquisition . . . 44
3.1.3 LZAP . . . 46
3.2 Noble Element Simulation Technique (NEST) . . . 48
3.3 LZ Electric Field (LZEF) . . . 48
3.3.1 Field Maps . . . 49
3.3.2 Drift Maps . . . 49

4 Simulating Single Electrons 51
4.1 Single Electron Data Selection . . . 52
4.2 Electric Field Modelling . . . 54
4.2.1 Grid Deflection . . . 55
4.2.2 Electric Field . . . 57
4.3 Electron Drift Speed . . . 59

4.4 Photon Production Distribution 60
4.4.1 GARFIELD++ Simulation 61
4.5 Scintillation Timing 63
4.6 SPE Modelling 65
4.6.1 Secondary PMT Effects 66
4.6.2 PMT Gain 67
4.6.3 SPE Pulse Shapes 68
4.7 Electroluminescence Yield 72
4.7.1 Araújo and Chepel 72
4.8 S2 Light Maps 72
4.8.1 Dey Model 75
4.9 Tuning Results 77

5 Simulating S2 Pulses **80**
5.1 Clustering and Yields 80
5.1.1 Light and Charge Yields 81
5.2 Charge Transport 83
5.2.1 Diffusion 84
5.2.2 Electron Lifetime 91
5.2.3 Extraction Efficiency 94
5.3 Afterpulsing 97
5.4 ParticleBomb Generator 102
5.5 Calibration Validations 104
5.5.1 Tritium 104
5.5.2 Neural Network 105
5.5.3 High Energy 109

6 Fragmentation **111**
6.1 Characterising Fragmentation 112
6.1.1 Simulation Measurements 113
6.1.2 AmBe Validations 115
6.2 Signal Modelling 118

6.3 Fragmentation Cut . . . 120
6.3.1 Cut Logic . . . 120
6.3.2 Signal Acceptance . . . 121
6.3.3 Background Rejection . . . 123
6.3.4 S2-Only Impact . . . 125
6.3.5 ^{8}B CEνNS Impact . . . 128
6.3.5.1 Cuts Correlation . . . 131
6.4 Fragmentation Tag . . . 132
6.4.1 Tag Logic . . . 133
6.4.2 Post-Processing Reconstruction . . . 137
6.4.3 S2-Only Impact . . . 142
6.4.4 ^{8}B CEνNS Impact . . . 144

7 Low Threshold S2-Only 147
7.1 S2-Only Searches . . . 147
7.1.1 Bias Mitigation . . . 147
7.1.2 Backgrounds . . . 148
7.2 Multiple Single Electron S2s . . . 149
7.2.1 Cut Development . . . 150
7.2.2 Neural Network Architecture . . . 153
7.3 Drift Time NN . . . 155
7.3.1 Neural Network Architectures . . . 156
7.3.2 Results . . . 158
7.4 Anomaly Detection . . . 161
7.4.1 UMAP Theory . . . 161
7.4.2 Clustering Algorithms . . . 163
7.4.3 Identifying Anomalous Clusters . . . 163
7.4.4 Background Rejection . . . 169

8 Conclusion 175
8.1 Waveform Simulations . . . 175
8.1.1 Alternative Developments . . . 176

8.1.2 Future Outlook 177
8.2 Fragmentation 177
8.3 S2-Only Analysis 178

References **180**

List of Abbreviations

General Terminology

$0\nu\beta\beta$ ..... Neutrino-less double beta decay

ADC ..... Analogue-to-Digital Converter

AFT ..... Area Fraction Time

APP ..... Afterpulse probability

APR ..... Afterpulse ratio

APS ..... Afterpulse size

CDF ..... Cumulative distribution function

CEνNS ..... Coherent elastic neutrino-nucleus scattering

CMB ..... Cosmic Microwave Background

CSD ..... Calibration source deployment tubes

DEC ..... Double electron capture

DD ..... Deuterium-deuterium

EL ..... Electroluminescence

ER ..... Electronic recoil

GXe ..... Gaseous xenon

ICV ..... Inner cryostat vessel

LCE ..... Light collection efficiency

LMW ..... Low mass WIMP

LXe ..... Liquid xenon

ML ..... Machine learning

MSE ..... Multiple Single Electron

NN . . . . . Neural network

NR . . . . . Nuclear recoil

OCV . . . . . Outer cryostat vessel

PDF . . . . . Probability density function

PMT . . . . . Photomultiplier tube

POD . . . . . Pulse-only digitisation

PTFE . . . . . Polytetrafluoroethylene

PLR . . . . . Profile likelihood ratio

QE . . . . . Quantum efficiency

ROI . . . . . Region of interest

RMS . . . . . Root mean square

RQs . . . . . Reduced quantities

SI . . . . . Spin-independent

TBA . . . . . Top-bottom asymmetry

TPC . . . . . Time projection chamber

UMAP . . . . . Uniform Manifold Approximation and Projection

VUV . . . . . Vacuum ultraviolet

WIMP . . . . . Weakly Interacting Massive Particle

XLZD . . . . . XENON, LUX-ZEPLIN, DARWIN

LZ Software Packages

BACCARAT . . . . Basically, A Component Centric Analogue Response to AnyThing. LZ's GEANT4 wrapper.

DER . . . . . Detector Electronics Response. Simulation package which produces LZ's PMT, electronics and data acquisition systems.

LZAP ..... LZ analysis package. A Gaudi framework which reads in the raw channel level waveforms and generates reconstructed quantities for analysis.

LZLAMA ..... LZ Light Analysis Monte-carlo Application. A fast simulation software which reads in energy deposits in the detector and uses parametrised models to produce reduced quantities.

NEST ..... Noble Element Simulation Technique

Pulse Types

S1 ..... Primary scintillation pulse.

$S1_c$ ..... Spatially corrected S1 pulse

S2 ..... Secondary scintillation pulse.

$S2_c$ ..... Spatially corrected S2 pulse

SE ..... Single Electron

SPE ..... Single PhotoElectron

MPE ..... Multiple PhotoElectron

Event Types

SS ..... Single Scatter

S2O S2-Only

MS ..... Multiple Scatter

List of Figures

1.1 Rotational velocity measured in the M31 spiral galaxy as a function of distance from the centre of the galaxy. The rotational velocity does not fall with radius as is expected by Equation 1.2, which implies that the mass continues to increase with radius. Figure from [4]. 4

1.2 (right) Observations of the Bullet Cluster from the Magellan images. The white bar indicates 200 kpc at the distance of the cluster. (left) Chandra X-ray observations of the same cluster. In both images, the green contours show the mass density reconstructed from weak lensing measurements. The heat map in the right image shows the density profile of the luminous matter in both clusters. Figure from [5] . 5

1.3 Temperature-temperature auto-correlation power spectrum from Planck 2018 across the multipole moments, l. Data from $l \geq 30$ are taken from the Plik cross-half-mission likelihood. Data for $2 \leq l \leq 29$ are from estimates produced by the Commander component-separation algorithm. The red dots depict the data and the blue line is the fit prediction from the ΛCDM cosmological model. Figure from [6]. 6

1.4 Brief summary of the mass regimes of a subset of dark matter candidates split into the wave-like and particle-like candidates at the $1\,\mathrm{eV}$ mass scale. Figure from [8]. . 8

1.5 Interaction processes utilised for different dark matter experiments. The arrows show the direction in which the interactions must take place for a given mode. The mediator and underlying interaction are unknown. Adapted from [16]. 11

1.6 DM-nucleon spin-independent cross-section limits (solid lines) versus dark matter mass set by the experiments labelled. The parameter space excluded by these experiments is shaded above the lines. The neutrino fog is shown for a xenon target in the shaded grey at the bottom. Projected limits are shown for various experiments with the dashed lines. White space indicates parameter space which is yet to be explored. Adapted from [35] to include the latest LZ results down to $3\,\mathrm{GeV}$ [1]. 15

2.1 Technical rendering of a cross-section of the LZ detector in the Davis Cavern at SURF. The TPC is displayed in the centre of the figure, with the skin veto immediately surrounding it. The green regions show the Gd-loaded liquid scintillator acrylic tanks. The white region is the water tank instrumented with 120 PMT domes. 18

2.2 An example event containing a single S1 and S2 waveform annotated as such. . . . 20

2.3 Doke plot of $S1_c$ per unit energy against $S2_c$ per unit energy to extract the values of detector parameters g_1 and g_2^{bot} for the WS2022 detector configuration [46]. g_2^{bot} refers to the electron gain as measured from the bottom PMT array. This is performed to circumvent PMT saturation occurring in the top array. Generally, $g_2 = 3g_2^{bot}$. Figure from [47]. 23

2.4 ER and NR bands in $S1$-$log(S2)$ space, as is custom for LZ, produced by tritium and DD calibrations, respectively. There is 99.9% discrimination of ER and NR events below the $40\,\mathrm{GeV/c^2}$ WIMP median. 23

2.5 Final dataset (black points) from the WS2024 result in $S1$-$log(S2)$ space. Also shown are the contours of the $40\,\mathrm{GeV}$ WIMP signature as well as the accidentals and ^{124}Xe double electron capture backgrounds. The ER band is shaded in grey. Recoil energy lines are given in the solid grey lines. 27

2.6 90% confidence level SI WIMP-nucleon cross-section limit as a function of WIMP mass from the WS2024 dataset in Figure 2.5. The solid black line shows the -1σ power-constrained limit, while the dashed grey line shows the unconstrained limit. The green and yellow shaded regions show the range of expected upper limits from 68% and 95% of background-only experiments, while the dashed black line shows the median expectation, obtained with post-fit background estimates. Figure from [36]. 28

2.7 Final dataset used for both the ^{8}B and low-mass WIMP searches. 1σ and 2σ contours are shown for the accidentals background (yellow), ^{8}B neutrino CEνNS signal (green) and a $3\,\mathrm{GeV/c^2}$ WIMP (pink). Each data point is represented by a pie chart indicating the fit contributions of each event to each model. The relative size of each pie chart is proportional to its contribution to the $3\,\mathrm{GeV/c^2}$ WIMP profile. The data and best fit models are shown as projections in both $S1_c$ and $S2_c$. Figure from [1]. 29

2.8 Low-mass WIMP upper limit (90% CL) observed from the dataset shown in Figure 2.7 with an exposure of 5.7 tonne-years (black line). The green and yellow bands show the upper limits of background-only for 68% and 95% of the experiments, respectively. The dashed black line shows the median expected upper limit, while the dotted black line shows the median 3σ discovery potential, using the post-fit model. Other experimental limits are also shown. The shaded grey region depicts the neutrino fog. Figure from [1]. 30

3.1 Simplified schematic showing the flow of both the full and fast chain methodologies in LZ simulations. BACCARAT is the cornerstone of both paths. The outputs of the simulation paths differ in the relative capability and detail within the full chain vs the fast chain. 32

3.2 Simplified flow chart describing the stages of processing through BACCARAT. . . 34

3.3 Simplified schematic of the LZ data acquisition pipeline from the PMT through the amplifier and through to FADR. The FADR process includes digitisation of the voltage signal, data sparsification, triggering, and event building [58]. The DER only needs to map the top chain from this figure, which is mirrored in the simple flow diagram in Figure 3.4. 38

3.4 Schematic showing the general steps taken through the DER process from initial PMT hits from BACCARAT through to the event output file. This output file is then ready to be passed into LZAP. 39

3.5 Heat map displaying the quantum efficiency of all PMTs in the top and bottom array of the LZ TPC. There are visible gaps in these arrays; they indicate where PMTs have been disabled due to high afterpulsing rates, which would otherwise degrade data quality. Both arrays follow the colour bar on the right side of the plot. 39

3.6 Circuit diagram for the output voltage from the R11410-22 Hamamatsu PMT base. The expression for $V(t)$ is given in Equation 3.1. 40

3.7 Voltage output from the PMT base for a single photon incident on the PMT face. The nominal values of all parameters used in the simulation are provided for this pulse, showing the relevant timing scales for the specified rise and fall times. The voltage scale shown motivates the necessity for amplification in the signal chain. . 41

3.8 Calibration fits of a high gain R11410-22 PMT from the Imperial College London test stand [59]. This PMT was set to a gain of 3.5×10^6, the expected gain of the LZ PMTs before commissioning. The plot shows the separation of single-photoelectron (SPE), double-photoelectron (DPE), and inelastic-scatter (IES) pulses. The additional peak in pulse area at $0\,\mathrm{nV\,s}$ is a result of first dynode hits. 42

3.9 Alterations made by passing the base PMT pulse through each stage of the data acquisition chain. The full characteristics of each change are approximated by applying a series of transfer functions described in Table 3.1. Note that in this plot, the low-gain response has been shifted by $80\,\mathrm{ns}$ for readability. 45

3.10 An example S2 waveform showing the integrated area (blue) used to define the S2 pulse area. Two common width metrics using AFT parameters are shown. An identical process is performed for all other pulse types. 47

4.1 Separating the SE pulse selection from the Other pulses by applying a TBA cut. The distribution observed here is from the WS2022 science run and hence differs slightly from the WS2024 results shown in this chapter due to the higher electric field. 53

4.2 Original (BACCARAT-V6.3.0, 2022) state of the simulation compared to a real selection of SE pulses in data. This comparison motivates the need for this tuning; discrepancies in SEs are amplified in S2s. 54

4.3 Simplified schematic of the LZ extraction region showing the scale of the liquid-gas xenon interface as well as the gate and anode grids and their relative deflection. The figure also shows the electric field as modelled in the gas region for the WS2024 detector configuration, showing the radial dependence induced by the small deflection of the grid wires. 55

4.4 Difference in electric field uniformity when correcting the implementation of the deflection model in the extraction region. 58

4.5 NEST drift speed model for electrons in gaseous xenon [41]. This is given in terms of the electric field by assuming a constant gas density and an ideal gas at LZ operating conditions. 60

4.6 Individual cell generated in GARFIELD++ geometry. A simulated electron track from the liquid surface to the gas is also shown. Note that this cell is generated in a 2-dimensional cross-section and does not include the additional field strength from the wires in the x-direction. 61
4.7 Fit to the observed photon production distribution along the path of a single electron through the gaseous xenon. The exponential rise is observed near the anode due to the increased field strength near the high-voltage wires. 62
4.8 Comparison of SPE data from LED calibrations to the default DER settings for SPEs. There is a clear systematic difference between the modelling in simulation and the calibration data. 65
4.9 Comparison of Imperial test stand characterisation of the PMT response with the observed response from a single channel in the top array of LZ. Note that the additional peak at $0\,\mathrm{nV\,s}$ in the Imperial test stand is not present in the LZ data. A triple Gaussian fit to the LZ data to encompass the IES contribution was tested and worsened the goodness of fit in all tested PMTs. Only the SPE and DPE peaks can be observed in the LZ data, with the DPE effect having a larger variance than the Imperial test stand. 67
4.10 Comparison of the average SPE waveform observed in PMT 10 with the SPE pulse using the nominal timing characteristics of every PMT pulse in the TPC from the DER. The variation in SPE size is entirely driven by the gain of each PMT. 68
4.11 SPE pulse area comparisons for individually tuned timing constants for each of the PMTs. Average response has improved but has much larger variance than desired for this result. 69
4.12 Final SPE pulse area comparisons with fixed $\tau_r = 0.48\,\mathrm{ns}$ and $\tau_f = 5.36\,\mathrm{ns}$. Some large outliers remain for the lower gain PMTs, but the majority of TPC PMTs align well with the LED calibrations. 70
4.13 Comparison of tuned time constants for the same PMT 10 example as in Figure 4.10. The plot shows the full distribution of all SPEs contributing to the mean SPE pulse from data, showing some consistency with the simulation, despite the simulation having a larger pulse area. 71

4.14 PMT hit probability map for PMT 30 on the top array of the TPC. Each data point shows the starting position of a photon which hits the PMT highlighted in red. The colour scale here is to allow the reader to visualise the depth. While there is a higher density of points below the PMT, the plot demonstrates that there is a broad distribution of potential starting positions. 74
4.15 Grid search loss matrix showing the difference from the expected convolution of LCE·QE from the measured value from the simulation output. The variation in these values indicates the level of fine-tuning occurring here, and the optimal configuration is identified as the minimum value. 76
4.16 Comparison of the SE pulse areas observed split into the total SE (top + bottom) and each of the top and bottom PMT arrays. The mean SE pulse area is written in the top-right corner of each plot. 78
4.17 Comparison of different SE width metrics including RMS width, AFT75-AFT25 and AFT99-AFT1. Notice that the RMS and 99-1 widths are still slightly overestimated. 78
4.18 Comparison of the average SE pulse shape from simulation and data. The data SE waveforms here are a small sample ($\mathcal{O}(100)$ events) taken from a CH_3T calibration run, where all criteria derived in Section 4.1 are applied along with a limit of 10 SE pulses per event to eliminate electron train backgrounds. The average waveforms are aligned with 0 at AFT1. The tail effects in the data are more prominent than in the simulation, likely due to the inclusion of effects such as stingers, electron trains, and photoionisation in the data but not in the simulation. 78

5.1 Direct comparison of the pulse area and AFT90 - AFT10 widths for simulated and calibration tritium, CH_3T, S2 pulses with a tuned SE response following Chapter 4. 81
5.2 Light and charge yield models from NEST shown for the four different response models introduced in this thesis. The $300\,\mathrm{keV}$ line represents the upper bound of the NEST- tuned parameters for the NR response fixed by the endpoint of the AmBe neutron spectrum. Above this line, the NR model is an extrapolation. The beta model is tuned to a combination of the CH_3T and ^{214}Bi beta spectra extending up to $\sim 3\,\mathrm{MeV}$. 83

5.3 Transverse radial spread of an electron cloud as a function of the drift length through the TPC. Three example transverse diffusion coefficients are simulated ranging from 20–40 $\mathrm{cm^2\,s^{-1}}$ to encompass the uncertainty on D_T. 86
5.4 Measured fit from CH_3T calibration data. For this fit, only D_L is allowed to float and $v_d = 1.4\,\mathrm{mm\,\mu s^{-1}}$ and $\sigma_{SE} = 562\,\mathrm{ns}$ remain fixed. A drift time cut is applied, removing $0\,\mu\mathrm{s}$ to $60\,\mu\mathrm{s}$ to remove events occurring in the gaseous xenon since they do not obey the same diffusion relationship. 88
5.5 Various models for the longitudinal diffusion coefficient as a function of field strength in the liquid xenon. Initial Model refers to the model present in the simulation prior to this tuning; the alpha fit is a model derived from fitting to alpha peaks in the detector, and similarly, the beta fit refers to modelling performed on a beta spectrum. Both the alpha and beta measurements shown are derived by collaborators external to this thesis work. The NEST model is also provided here as a reference for a global fit to the latest measurements in the field. The LZ WS2024 result from Figure 5.4 is shown as a single data point derived by the previous fitting method from a CH_3T calibration. 89
5.6 Measurements of the longitudinal diffusion coefficient from different energy calibration sources. The measurements from simulation are also shown with agreement with the general energy trend despite having a fixed input diffusion coefficient of $22\,\mathrm{cm^2\,s^{-1}}$. 90
5.7 Comparison of diffusion properties from a CH_3T calibration source (black contours) and two runs of the simulation. The blue contours show the outcome of simulating with the measured coefficient from the data, indicating an overestimation of the widths. The red contours show the result obtained with the tuned coefficient. The diffusion coefficients measured from this plot are given at the top for each corresponding curve. 91
5.8 Survival probability for various values of the attenuation length used in Equation 5.8. The area of the shaded region under each curve is representative of the survival proportion of electrons. The initial default value in the simulation was $1\,\mathrm{m}$. 93
5.9 1σ contour comparison of real and simulated CH_3T S2 size with drift time, demonstrating the agreement due to the electron lifetime modelling. 94

5.10 Simplified diagram showing the effect of increased electric field on the potential barrier from liquid to gaseous xenon. The stronger the field strength, the lower the barrier height and the higher the electron extraction efficiency. 95
5.11 Measured extraction efficiency curve from CH_3T calibration data compared with the accepted NEST model. 96
5.12 Heat map of the reconstructed positions of simulated S2 pulses. Despite the true position distribution being flat, notable areas of higher density arise localised to PMTs with high afterpulsing rates. 98
5.13 LED calibration data time-aligned to the beginning of the LED-induced S1 pulse, displaying the time windows for each afterpulsing species. All small pulses shown to follow the initial LED-induced S1 pulse are taken to be afterpulses. Figure from [93]. 100
5.14 Mean afterpulsing size for each ion species simulated in the DER. The error bars show the error on the mean from the fitting procedure. Note that the afterpulsing size is measured in photoelectrons, phe. 101
5.15 Investigation into the impact of applying afterpulsing effects on the S1 size for simulated CH_3T data. The red dashed line shows the mean double-photoelectron probability for PMTs in LZ, revealing a logical fault between the afterpulsing and double-photoelectron modelling. 102
5.16 Direct comparison between simulation and calibration CH_3T data for the pulse area, AFT width and total reconstructed energy. These represent the core RQs used in the analysis and demonstrate the agreement between the simulation and the data. 105
5.17 Top-bottom asymmetry comparison between simulation and calibration CH_3T data. The left-hand plot shows the total disagreement between the two populations, while the right-hand plot shows the pulse area observed in each PMT array. The main discrepancy arises from the bottom PMT array being overestimated by the simulation. 106

5.18 Neural network output for classifying a sample of testing data and simulation S2s. Test Accuracy here gives the percentage of correctly identified labels. For a perfect simulation, this should result in 50% accuracy. This plot demonstrates the model has difficulty distinguishing between real and simulated S2 pulses. 107
5.19 Permutation feature importance for all RQs used to train the classification NN. The most important RQs used to distinguish are AFT99 and TBA. 108
5.20 Impact of removing AFT99 and TBA parameters from the training data in Figure 5.18. These two RQs are known to be inaccurate or missing models in the simulation run, and further tuning work is left for future analysis. 108
5.21 Mean S2 pulse area measured by the bottom PMT array for the labelled gamma calibration sources. Pulses in the bottom array are chosen to remove the PMT saturation effects observed in the top array PMTs for large S2s. 109

6.1 Simulated S2 waveform showing the fragmentation effect. Both pulses originate from the same position and time in the detector but are reconstructed as a single prominent S2 and a fragment. Fragments can be identified by LZAP as SE or S2 pulses. 111
6.2 Percentage of events generated using the ParticleBomb S2 generator which result in either 1 prominent S2, more than one prominent S2 or no prominent S2s. A single prominent S2 is used as a proxy for a single scatter event, while more than one is used as a multiple scatter event. The percentage loss across multiple scatters is negligible. 113
6.3 Probability of fragmentation of an S2 pulse consisting of 3–15 electrons in drift time bins spanning the total liquid xenon volume. Fragmentation remains a significant problem near the bottom of the detector, as expected from the diffusion relation. 115
6.4 Distribution of pulse area lost as a percentage of the total area of the S2 pulses for each S2 size in the low-energy ROI. The marker shows the mean percentage loss on each distribution. 116

6.5 Time and radial separations between the prominent S2 in each simulated event and any adjacent SE or S2 fragments shown for both pre- and post-S2. Time separation is defined as AFT5 of one pulse to AFT95 of the other. Radial separation is determined from the reconstructed positions of each pulse. 117
6.6 Probability of fragmentation to multiple S2 pulses. AmBe calibration S2s are displayed as triangles, compared with the simulation measurements shown as circles. The error bars are large for the AmBe calibration data due to limited statistics. Nevertheless, the simulation is consistent with the calibration data. 118
6.7 Impact of applying the fragmentation probabilities and area loss distributions to the ^8B PDF generated using LZLAMA. This area loss causes some signal loss by pushing signal events out of the ^8B ROI (below $3.5\,\mathrm{electrons}$ in $S2_c$). Figure from Alberto Usón. 119
6.8 (top) Example S2 pulse with time boundaries explicitly shaded out. Any other S2 or SE pulses that exist in this time-space will cause this event to fail this cut so long as they also satisfy the spatial cut in (bottom). The spatial cut shows the $10\,\mathrm{cm}$ radius drawn around the prominent S2 position in the TPC. Any other S2 or SE pulses that are in this space, in addition to the time space, will be cut. The radius of the entire detector is given here for reference. 121
6.9 Fragmentation parameter space with the spatial and time separation between pulses illustrated. The red simulated data points indicate where the fragmented events lie, and the white dashed lines indicate the region of this parameter space removed by this cut. 122
6.10 Signal acceptance of applying the fragmentation cut to a sample of CH_3T + AmLi ChopStitched data. This is constructed by sampling the S1 from the AmLi neutron calibrations and the S2 from the CH_3T calibrations to form NR events. 123
6.11 Percentage of accidental ChopStitch events that fail the fragmentation cut as a function of the spatio-temporally corrected S2 size in terms of electrons. 124

6.12 Cut efficiency for successfully removing simulated fragmented S2s binned by prominent S2 size in terms of electrons extracted. This results in a flat efficiency of 99.3% for targeting fragmented S2s. This also demonstrates that S2s up to $40\,\mathrm{electrons}$ can still exhibit fragmentation though its impact is less important since these lose a smaller fraction ($\sim$3%) of the total pulse area. 125
6.13 Showing the proportion of SS events removed by the fragmentation cut. A secondary peak is shown here, due to events from the resistor chain, as demonstrated by applying the resistor chain cut (FCRES), which removes any events with S2s near the (x, y) positions of the resistors. The final distribution shows a gradual increase in the number of events cut as drift time increases. 126
6.14 RMS width against S2 size phase space for all S2-only events in the 10% unblinded dataset. Areas outside of the region of interest are shaded in grey. The black data points here show the events in the search that pass all cuts except the ROI cuts in RMS width and S2 pulse area. The red data points show the events that uniquely fail the fragmentation cut and pass all other cuts applied here. In the highlighted region of interest, 67 events are uniquely removed by the fragmentation cut, 12% of the total number of events. 127
6.15 Final ^{8}B/LMW dataset with salt removed. The fragmentation cut would have removed two events. The contours show the ^{8}B signal model contours from simulation. 129
6.16 2-dimensional projection of the reconstructed position of all events in the final dataset. Highlighted are the two events which fail the fragmentation cut. 130
6.17 Final impact of the fragmentation cut on the ^{8}B CEνNS observed significance. . . . 131
6.18 Correlation matrix between the cuts applied to the ^{8}B analysis. The fragmentation cut is highlighted with red borders, showing minimal correlation with most cuts. All correlated cuts are S2-shape cuts. 132
6.19 Cuts applied sequentially to a selection of S2-only data and simulation, showing the impact of each cut on the time between consecutive pulses pre- and post-S2. A full description of the cuts is omitted here, and only specific cuts will be discussed in the text where relevant. 134

6.20 (top) Example MSE event window. The red lines indicate SE pulses within the event window, and the blue line shows the S2. (bottom) The same event shown in the reconstructed position space where the colours remain the same. Since all SE pulses are so close to the S2 pulse position, this is most likely an S2 caused by coincident overlap of multiple SE pulses. A 10 cm circle is drawn around the prominent S2 position to illustrate the radius within which electrons are tagged as MSE, motivated by the simulation. 136
6.21 Simple diagram of a boxcar distribution as defined in Equation 6.2. 138
6.22 Extreme example of the fragmentation effect on the width of a simulated S2 pulse. The 'Original AFT95 - AFT5' width indicated here shows the width of this pulse taken directly from LZAP classification of the prominent S2. The new quantiles derived from the post-reconstruction processing algorithm are indicated along the bottom of the pulse. The 'New AFT95 - AFT5' width according to this reconstruction is given by the blue arrows. The shaded green area of the pulse is a result of the integrated width of the waveform. 140
6.23 Absolute and fractional residuals from the reconstructed AFT95 - AFT5 widths and the actual integrated widths from the simulated waveforms. 141
6.24 Impact of the width reconstruction on a distribution of simulated S2s of 4–$5\,\mathrm{electrons}$. The heat map shows the total distribution, and the red crosses indicate the previously fragmented widths that then become the green dots as the corrected widths, now conforming to the expected drift time-width relationship. 141
6.25 Effect of reconstruction on the final S2-only dataset demonstrating the gain and loss of events in this parameter space. 143
6.26 Impact of the reconstruction algorithm on the evaluation of all data quality cuts for the S2-only analysis. 143
6.27 Two events transformed into new corrected S2 space; all other events remain in their previous positions. No new events were added to the dataset, and no events were taken away. 144
6.28 Final impact of the reconstruction algorithm on the ^{8}B observed significance. . . . 145

7.1 Background spectrum for the S2O regime from WS2022 data with 60 live-days of exposure. The largest contributions are from the gate and cathode. 148
7.2 ^{8}B CEνNS interaction discovery significance as a function of time for 3 different levels of uncertainty on the background model from the cathode. The significance is evaluated with both a signal-to-background ratio (dashed) and the Asimov test statistic (solid) [99]. The current best estimate of the uncertainty for the S2O regime is 7%. 149
7.3 Number of SE pulses identified within $10\,\mathrm{cm}$ of the prominent S2 in the event window, further than $5000\,\mathrm{ns}$ away. Most of the MSE pulses are below the analysis threshold of $3.5\,\mathrm{electrons}$. Above this, the most likely outcome is one SE pulse within the event window. 151
7.4 Percentage of MSE events removed by varying the threshold for the number of SE pulses in the event window for all MSE events up to $15.5\,\mathrm{electrons}$ (light blue) and just the fraction above the $3.5\,\mathrm{electron}$ threshold (dark blue). 152
7.5 Final results from training the NN with a selection of MSE events and simulated S2s. The NN achieves a classification accuracy of 97.6% and separates the MSE population from the simulated S2s. 154
7.6 Diffusion relation contours shown for increasing S2 sizes. The smallest S2 sizes (3–$25\,\mathrm{electrons}$) show a much larger variance. Above $75\,\mathrm{electrons}$, the correlation becomes slightly stronger, approaching the Mean Diffusion Relation line. A drift time cut is applied to the top $60\,\mu\mathrm{s}$ to remove a large number of gas and gate events which dominate the contours. The general trends shown extend down to $0\,\mu\mathrm{s}$ in drift time. 155
7.7 Comparison of the root mean squared error of each training epoch for each architecture approach. A significant improvement is achieved by training on the waveforms using a CNN. The FiLM ResNet achieves the best performance, with the smallest error and the shortest training time to reach stability. 157

7.8 Predicted z height of an S2 against its true z height for the FiLM residual CNN trained on simulated S2 waveforms. Four regression loss metrics are evaluated in the top left corner, where RMSE denotes the Root Mean-Square Error; MAE denotes the Mean Absolute Error; R^2 denotes the coefficient of determination, and the Pearson r is the Pearson correlation coefficient. The residuals of Predicted z - True z are shown in the bottom panel. 158

7.9 Comparison of the original and post-calibration outputs from the FiLM ResNet. The largest improvement came from the GBDT approach, reducing the RMSE by $\sim 24\,\mathrm{mm}$. The calibrated output does not show a significant improvement over the standard output, except for a better-calibrated upper edge. 160

7.10 RQ-WF contrast embedding for all S2O data after the preliminary cuts are applied. 25 clusters of events are identified. 164

7.11 UMAP RQ-WF embeddings with colour bars showing Negative Area, TBA, RMS and hotspot index distributions in the UMAP dimensional space. Each RQ is used to identify the likely source of each cluster. 165

7.12 RQ - WF mapping from Figure 7.10 highlighting only the events that pass all the current analysis cuts and the clusters to which they belong. 167

7.13 WF - RQ contrast mapping for all events passing all analysis cuts. Three anomalous regions are separated from the 3 main clustered regions. 168

7.14 UMAP trained on S2 waveforms with all analysis cuts applied except RMS width. The grey data points show the output of all training data, while the highlighted points show the events that pass the RMS cut, coloured by cluster. 170

7.15 Direct comparison of the mean waveform of S1-tagged cathode S2 selection against the mean waveform of each of the clusters in Figure 7.14. Cluster 0 aligns very well with the cathode S2s, whereas the other two clusters disagree. 171

7.16 Direct comparison of the mean waveform of Cluster 2 and a glue-ring-contaminated selection of gate S2s. The 1σ bands of the mean waveforms are shaded in the corresponding colour. 171

7.17 Comparison of manual selections of Cluster 2 waveforms with the mean gate and gas S2 waveforms. The Cluster 2 waveforms show substructure within the cluster, suggesting a potential mixture of gate and gas events near the top of the detector. . 172

7.18 Position distribution of each of the clusters from Figure 7.14 with the position of the remaining hotspot highlighted in the red circle. 173

List of Tables

2.1 Common calibration sources deployed in LZ separated into ER and NR responses. IC refers to internal capture, and AE refers to Auger electrons. Note that the energies listed for each source are the energies relevant for LZ and are not exhaustive. Adapted from [42]. 22

2.2 Background table for the relevant backgrounds of the WS2024 result, including the pre-fit expectation and post-fit result. The detector neutrons pre-fit expectation is derived from a sample of veto-tagged events and is not explicitly used in the final combined fit. Table from [36]. 25

3.1 Transfer functions used by the DER in modelling steps through stages of the data acquisition flow. In these functions, ω refers to the cut-off frequency of the filter and τ refers to the RC time constant of the circuit schematic. These values will differ between each application stage in the DAQ pipeline. 44

5.1 Afterpulsing species in TPC PMTs and their initial expected timings and number of photoelectrons produced. The timings presented here should be understood as a time window around the mean as shown in Figure 5.13 [93]. The number of photoelectrons is presented as the mean and standard deviation, following a Gaussian distribution. 99

7.1 Post-training calibration methods applied to Figure 7.8 where z_p is the predicted z position and z_c is the calibrated z position. 159

Introduction

Over the course of the last century, physicists have been perplexed by the existence of dark matter: the missing mass hidden across the Universe. While able to detect its interactions through gravity with the dynamics of galaxies and gravitational lensing, no direct observation of any dark matter particle has ever been made. The LUX-ZEPLIN (LZ) experiment is one of the leading dark matter direct detection experiments attempting to make direct observations of the interactions between dark matter and Standard Model matter. Primarily searching for weakly interacting massive particles (WIMPs), the LZ detector is a dual-phase xenon time projection chamber (TPC) operating at a depth of $1.5\,\mathrm{km}$ underground at the Sanford Underground Research Facility in South Dakota, USA. LZ and other dual-phase xenon TPCs are well suited to detecting the scintillation and ionisation signals produced by particle interactions in liquid xenon.

The sensitivity of the LZ detector places it at the forefront of the dark matter field, holding world-leading limits on the spin-independent WIMP-nucleon scattering cross section for WIMP masses $>5\,\mathrm{GeV/c^2}$ [1]. The low WIMP mass limit is imposed by the analysis threshold, determined by the size of the charge signal, the S2, within LZ. With the LZ detector sensitive to individual electrons, we aim to utilise the full sensitivity of LZ to dark matter masses to improve WIMP discovery potential.

The primary question explored during this work is how the LZ waveform simulation can be used to accurately model and study detector behaviour as the S2 signals approach the few-electron threshold. Consequently, can waveform-level analysis be used to improve background discrimination beyond the conventional RQ analysis? The studies presented in this thesis examine the physical mechanisms driving the S2 signal and aim to utilise the understanding gained to push the LZ threshold to lower energies and enhance the WIMP search region.

Chapter 1 provides the underlying scientific evidence and challenges posed to the dark matter problem. We establish the evidence gained over the last century, including galaxy dynamics, cosmic microwave background anisotropies and gravitational lensing. Several dark matter candidates are introduced with their motivations and current exclusions where relevant. Finally, the dark matter detection methods are presented, noting the interplay among them, their leading experiments, and current results.

Chapter 2 details the LZ experiment, the fundamental physics it relies upon to function and outlines the most recent world-leading results in the field. A detailed account of the detector setup is given, including the calibration techniques, backgrounds, and bias mitigation, all of which are necessary for the high-fidelity analysis LZ provides.

Chapter 3 introduces the LZ simulation framework, with a particular focus on waveform simulation and the full chain. This chapter details the two main software components, BACCARAT and DER, which simulate all interactions in the xenon and all electronics responses from the data acquisition chain, respectively. Additionally, we introduce the LZ analysis package, LZAP, through which all LZ data is analysed, as well as the LZ Electric Fields (LZEF) and Noble Element Simulation Technique (NEST) packages, which are integrated into BACCARAT.

Chapters 4 and 5 detail significant improvements implemented in the BACCARAT and DER simulation packages to refine the accuracy of the single electron and S2 responses, respectively. The improvements made in these chapters are crucial to produce the results presented in Chapters 6 and 7.

Chapter 6 details the investigation of the fragmentation effects of S2 pulses at low-energy. The effects of fragmentation are characterised utilising the full extent of the tuned waveform simulation. Three approaches to resolving the loss of signal due to fragmentation are proposed, and the statistical impact is evaluated by changes in the ^{8}B CEνNS discovery significance.

Chapter 7 outlines three machine learning techniques applied to low-energy S2s to extract event-level information from the encoded physics within S2 waveforms and improve background rejection capabilities. This inherently enhances the sensitivity of LZ to low-energy signals that are near the threshold and are subject to additional backgrounds.

Chapter 8 finally concludes with a reiteration of the contributions to LZ and the dark matter field encompassed in this thesis. It also indicates future advancements that could be applied to build on this work, both for simulation upgrades and for enhanced analysis techniques.

1 | Dark Matter

One of the most prominent lacunae in our knowledge of the Universe is the dark matter phenomenon. Astrophysical observations over the past century have revealed a vast amount of a new type of matter in our Universe that constitutes $\sim$85% of the total matter density. To date, the nature of this phenomenon remains unclear, with many theories posed but none yet confirmed. This chapter will provide a brief introduction to the evidence accrued over the last century and the proposed detection methods and experiments currently dedicated to this search.

1.1 Experimental Evidence for Dark Matter

Whilst no direct detection of any dark matter candidate has been observed, evidence of the existence of dark matter is abundant. Each advancement in our knowledge imposes new constraints on the search for these elusive particles. Unless otherwise stated, much of this evidence is given with reference to the dark matter overview text by Bertone and Hooper, which contains a more detailed and complete account of the history of the dark matter conundrum [2].

1.1.1 Galactic Motion

One of the earliest indications for the existence of dark matter was provided by Fritz Zwicky in 1933 whilst observing the motion of galaxies in the Coma cluster [3]. By measuring the speeds of galaxies, he noticed they were moving too quickly for the amount of matter they appeared to contain. He became the first physicist to apply the Virial Theorem,

$$2\langle T\rangle + \langle U\rangle = 0, \tag{1.1}$$

where $\langle T\rangle$ is the mean kinetic energy and $\langle U\rangle$ is the mean gravitational potential energy, to calculate the mass of the galaxy cluster. This method is agnostic about whether that mass is observable. The result of his calculation implied that the total mass of the cluster is much larger than the mass of the visible matter present. From this, he concluded that these clusters must contain a large amount of mass that cannot be seen, which he referred to as 'Dunkle Materie': *Dark Matter*.

1.1.2 Galaxy Rotation Curves

In the 1970s, independent evidence on galactic scales emerged from measurements of galaxy rotation curves by Vera Rubin and Kent Ford [4]. They measured the orbital speeds of stars and gas at different radii in the M31 spiral galaxy. The rotation speed of mass in gravitational orbit is given by

$$v(r) = \sqrt{\frac{GM(r)}{r}}, \tag{1.2}$$

where all symbols have their usual meanings. From this, if all of the matter were mostly compacted at the centre of the galaxy, as is the case for the visible matter, the speed should fall with radius as $v \propto r^{-1/2}$. Rubin and Ford's measurements are displayed in Figure 1.1. From this observation, it was concluded that the mass distribution, $M(r)$, is not concentrated at the centre with the visible matter but rather extends out in a manner where for large radii $M(r) \propto r$. This provides critical evidence for the existence of dark matter haloes around galaxies, whose gravitational interaction enables indirect observations.

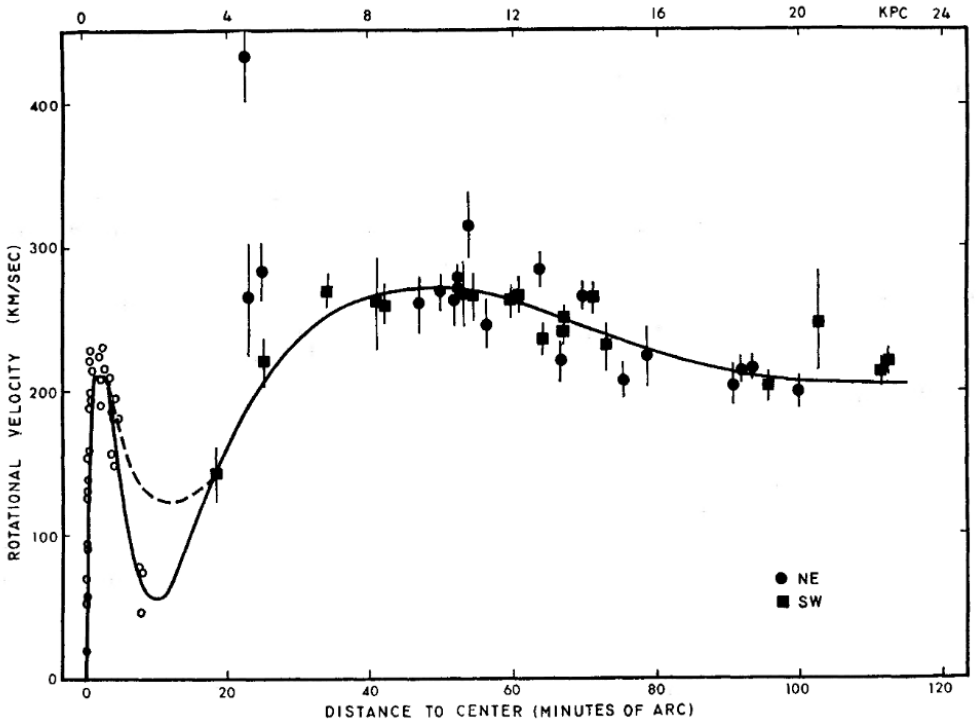


Figure 1.1: Rotational velocity measured in the M31 spiral galaxy as a function of distance from the centre of the galaxy. The rotational velocity does not fall with radius as is expected by Equation 1.2, which implies that the mass continues to increase with radius. Figure from [4].

1.1.3 Gravitational Lensing

In 1915, with his Theory of General Relativity, Einstein proposed that light can be deflected from its straight-line paths by massive objects that curve spacetime. Since then, gravitational lensing has become a frequently utilised technique in astrophysics. In dark matter searches, this is particularly useful for determining the mass distribution of a region of space by observing the distortion of light caused by the gravitational field. An important case of this is from the Bullet Cluster. The Bullet Cluster observations from 2006, shown in Figure 1.2, depict the aftermath of a collision between two large galaxy clusters.

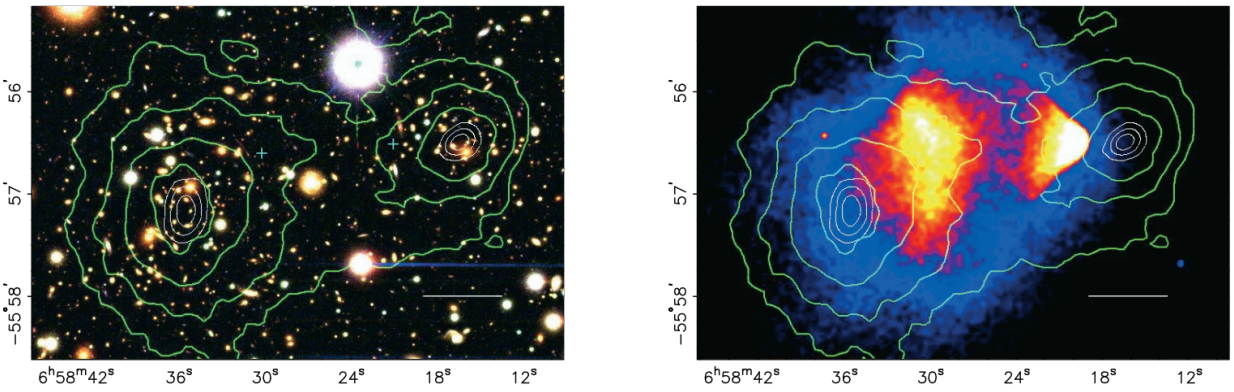


Figure 1.2: (right) Observations of the Bullet Cluster from the Magellan images. The white bar indicates 200 kpc at the distance of the cluster. (left) Chandra X-ray observations of the same cluster. In both images, the green contours show the mass density reconstructed from weak lensing measurements. The heat map in the right image shows the density profile of the luminous matter in both clusters. Figure from [5]

The Chandra X-ray images of the Bullet Cluster in Figure 1.2 visually demonstrate where the majority of the luminous and non-luminous mass lies in the aftermath of this galaxy merger. The conical shape of the luminous mass on the right-hand side of the plot suggests a level of drag produced by the collisions when the two galaxies passed through one another. Thus, most of the baryonic mass resides in the hot X-ray-emitting gas. With this additional context, we see that the weak-lensing mass peaks on either side of the plot are not conical and are further from the centre of the merger. This separation constrains the dominant component of these peaks to be weakly interacting. Thus, the dark matter in these two galaxies does not experience the same drag forces as the luminous matter, indicating that it must be very weakly interacting and therefore non-baryonic.

1.1.4 Cosmic Microwave Background

Precise measurements of the dark matter abundance can be made using the Cosmic Microwave Background (CMB) anisotropies. Baryonic acoustic oscillations (BAO) in the early Universe are observable through the power spectrum of the CMB, shown in Figure 1.3.

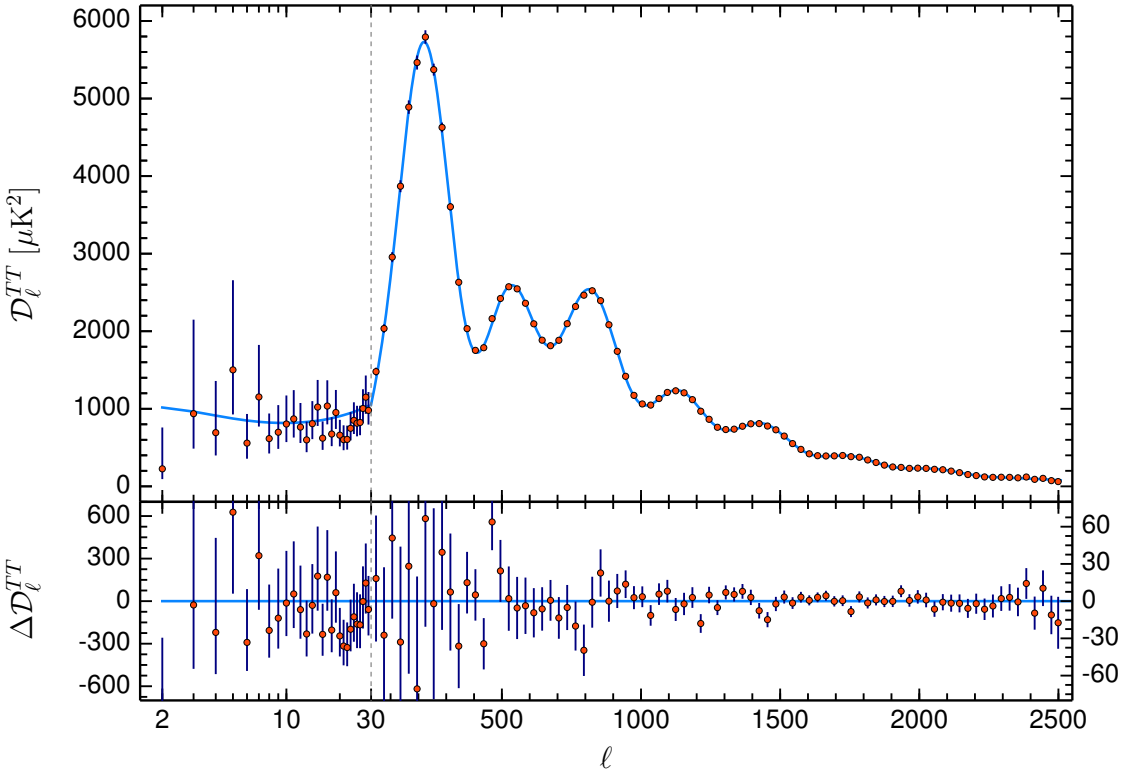


Figure 1.3: Temperature-temperature auto-correlation power spectrum from Planck 2018 across the multipole moments, l. Data from $l \geq 30$ are taken from the Plik cross-half-mission likelihood. Data for $2 \leq l \leq 29$ are from estimates produced by the Commander component-separation algorithm. The red dots depict the data and the blue line is the fit prediction from the ΛCDM cosmological model. Figure from [6].

The peaks observed in Figure 1.3 originate from acoustic oscillations in the photon-baryon plasma in the early Universe, where the relative heights and positions of the peaks depend on the amounts of baryons, radiation, dark matter, and dark energy. This figure also shows a strong fit to the Λ Cold Dark Matter (ΛCDM) cosmological model in which the dark matter is non-relativistic, hence 'cold'. This is initially constrained by the existence of structure in the Universe which could not have formed in the hot dark matter paradigm [7]. Removing the dark matter from the ΛCDM model fails to reproduce the observed peaks. Two important parameters of the ΛCDM model are the baryonic matter density, Ω_b, and the dark matter density, Ω_{DM}. The fits of the ΛCDM model

return $\Omega_b = 0.049$ and $\Omega_{DM} = 0.26$, leading to the ratio of dark matter to baryonic matter of 5:1. Hence, it is concluded that dark matter constitutes 85% of the total matter density in the Universe and Standard Model matter makes up the remaining 15%.

1.1.5 Properties of Dark Matter

From the presented pieces of evidence, several conclusions can be drawn which constrain the nature of dark matter. These properties are stated below.

- Interacts gravitationally
- Non-relativistic
- Non-baryonic
- Mostly non-luminous
- Long-lived
- Rarely interacting

1.2 Dark Matter Candidates

With these properties of dark matter outlined by the prior evidence, many theories are proposed to explain the nature of dark matter, spanning a wide variety of masses shown in Figure 1.4.

1.2.1 Modified Newtonian Dynamics

Before discussing the particle-like dark matter, we must first introduce the alternative theory to dark matter: modified Newtonian dynamics (MOND). This theory suggests that instead of there existing some vast amount of unknown mass throughout the Universe, perhaps these observations indicate that our understanding of gravity is flawed [9]. Currently, MOND has produced results which could explain the missing mass on a galactic scale, being able to reproduce the flat rotation curves observed in Figure 1.1 [10]. However, it is unable to explain the results of the CMB anisotropies on a cosmological scale. Because of this, it is deemed more likely by the dark matter community that dark matter has a particle nature, and our understanding of gravity holds.

1.2.2 Massive Astrophysical Compact Halo Objects

Massive astrophysical compact halo objects (MACHOs) were among the first dark matter candidates proposed to explain the missing mass in galaxies. MACHOs describe a class of astrophysical

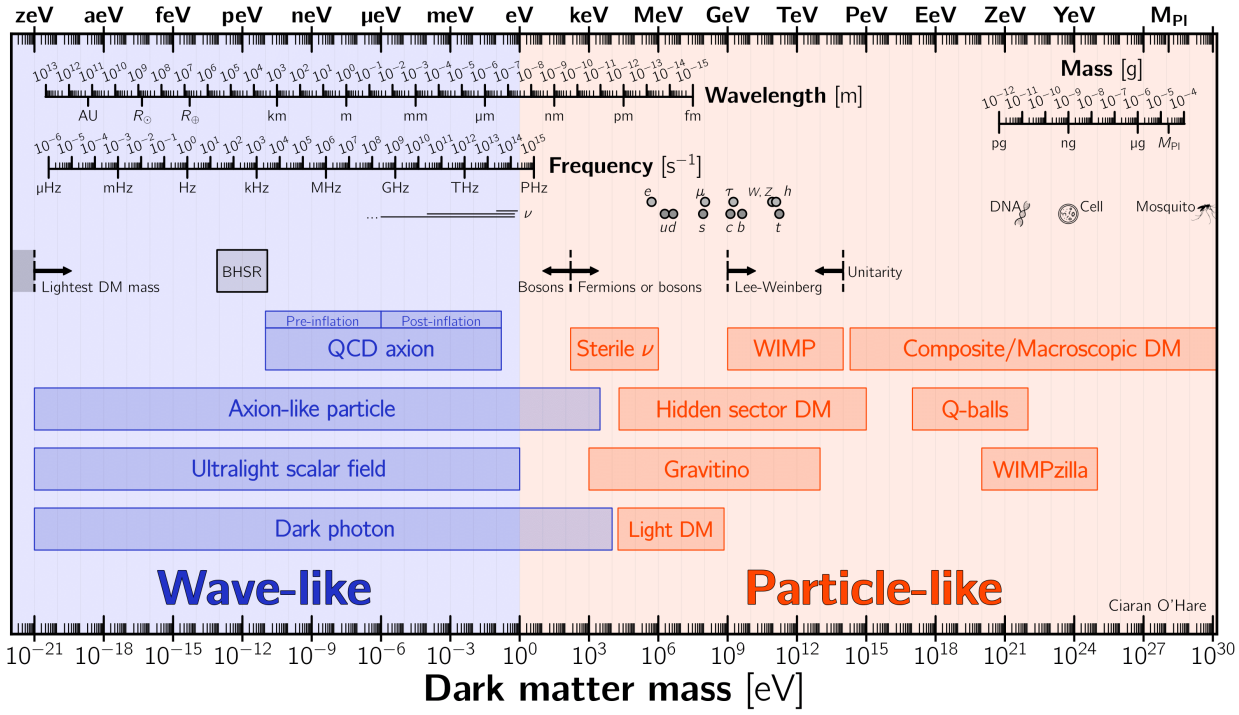


Figure 1.4: Brief summary of the mass regimes of a subset of dark matter candidates split into the wave-like and particle-like candidates at the $1\,\mathrm{eV}$ mass scale. Figure from [8].

objects such as planets, brown dwarfs, red dwarfs, white dwarfs, neutron stars, and black holes which are less luminous than ordinary stars and thus were not observable as luminous mass in the galaxies. This would also explain the gravitational interactions.

However, the modern consensus on MACHOs is that, if they exist, they will only contribute a small proportion of the dark matter mass. This is mostly constrained by microlensing surveys of the Magellanic clouds in the Milky Way. It would be expected that one should see these objects passing in front of the clouds and thus being able to observe peaks in the brightness of the background stars from gravitational microlensing. The observations from EROS, MACHO and OGLE conclude that there are insufficient observations of this for it to be the dominant dark matter mass [11–13].

1.2.3 Axions

Axions were originally proposed as a solution to the strong CP problem in QCD. In the years following its conception it became heavily constrained by the rates of exotic meson decays, cooling of red giant stars and supernovae observations [14]. The remaining hypothetical particle can have a mass between $\sim 10^{-11}$ and $1\,\mathrm{eV}$ and interact very weakly with Standard Model particles. If this

particle exists in abundance, it becomes a viable candidate for dark matter.

The coupling between the axion and the photon can be written as

$$g_{a\gamma\gamma} = \frac{\alpha}{2\pi f_a} C_{a\gamma}, \tag{1.3}$$

where α is the fine-structure constant, f_a is the scale at which the new $U(1)$ global symmetry introduced to solve the strong-CP problem breaks, and $C_{a\gamma} \sim \mathcal{O}(1)$ encodes the model dependence. Experiments can then be devised to utilise the coupling in Equation 1.3 in order to attempt to detect axion particles via their interactions with photons in the presence of a strong magnetic field.

The extension of the QCD axion described above is the field of axion-like particles (ALPs), which are capable of much larger mass ranges, as can be seen in Figure 1.4. The main difference between the QCD axion and the ALPs is that the ALPs are not constrained by the QCD coupling strength and hence can occupy a broader definition of particles which could be light dark matter candidates.

1.2.4 Sterile Neutrinos

The Standard Model neutrinos exhibit many of the properties necessary for the missing mass; they are electrically neutral, stable on cosmological timescales, and are weakly interacting. Despite this, their sub-eV masses imply that they were relativistic during the epoch of structure formation and hence these neutrinos are unable to account for the Universe's large-scale structure and thus cannot be a dark matter candidate. However, extensions of the neutrino sector can introduce heavier, sterile neutrinos which only interact by mixing with the active neutrino flavours. At the constrained masses of $\mathcal{O}(\mathrm{keV})$ these neutrinos would behave as warm dark matter particles and thus many simple production scenarios are heavily constrained. Nevertheless, they remain a well-motivated warm dark matter candidate. Searches for sterile neutrinos are predominantly studying the theoretical radiative decay of sterile neutrinos to X-ray photons.

1.2.5 Weakly Interacting Massive Particles

One of the most well-motivated dark matter candidates is the Weakly Interacting Massive Particle (WIMP). This is a class of particles ranging in mass from GeV to TeV scales with couplings on the weak-scale. WIMPs are motivated by the notion that a long-lived particle with weak-scale

interactions would naturally be produced in the early Universe with the correct relic abundance required of dark matter. This is known as the WIMP miracle. In the early Universe, WIMP particles of this mass would be held in thermal equilibrium with the plasma by the annihilation and pair-production processes into and from Standard Model fermions.

$$\chi + \overline{\chi} \rightleftharpoons f + \overline{f} \tag{1.4}$$

where χ represents a WIMP and f denotes a Standard Model fermion. As the Universe expanded and cooled, the annihilation rate falls below the Hubble expansion rate and freezes out, leaving the relic abundance of WIMPs remarkably close to the dark matter abundance observed in the Universe [15]. This can be shown by solving the Boltzmann equation for the asymptotic yield, resulting in the energy density of dark matter, which can be written as

$$\Omega_\chi h^2 \simeq \frac{3 \times 10^{-27}\ cm^3\ s^{-1}}{\langle \sigma v \rangle}, \tag{1.5}$$

where $\langle \sigma v \rangle$ denotes the thermally averaged annihilation cross-section. Current measurements of $\Omega_{DM} h^2 \simeq 0.12$ imply that the dark matter annihilation cross-section must be $\sim 3 \times 10^{-26}\,\mathrm{cm}^3\,\mathrm{s}^{-1}$, which is within an order of magnitude of the expected weak-annihilation cross-section, $10^{-25}\,\mathrm{cm}^3\,\mathrm{s}^{-1}$. This means that the dark matter abundance naturally emerges at the electroweak scale for a stable particle with weak-scale annihilation cross-sections, making it a well-motivated candidate.

Searches for WIMPs take many forms, all utilising the production and annihilation interactions proposed in Equation 1.4.

1.3 Dark Matter Detection Methods

The weakly-interacting nature of dark matter makes it extremely challenging to detect. While the effects of dark matter on gravity have been measured in abundance, there has yet to be a recorded interaction of the particle itself. Many of the methods devised rely on the interactions demonstrated in Equation 1.4. Three approaches are taken to attempt to observe these interactions. These are depicted in Figure 1.5. These three detection methods are utilised in cooperation with one another since they directly probe the same underlying physics and are able to collectively constrain the

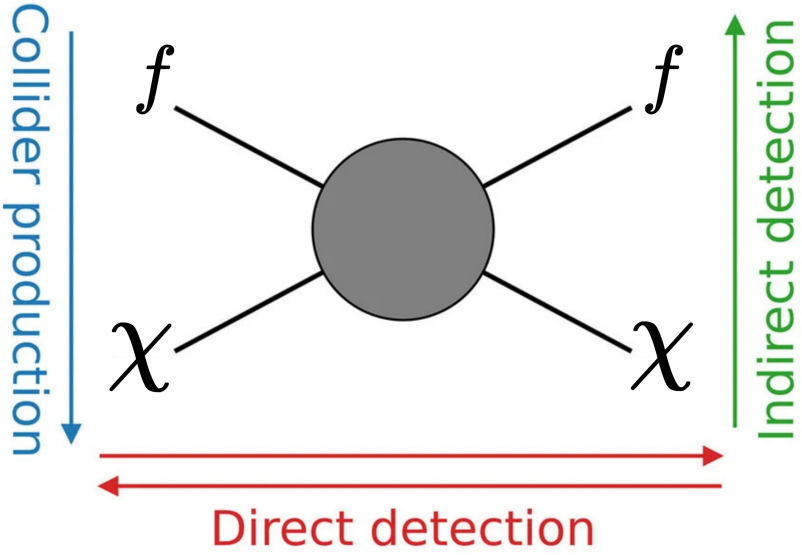


Figure 1.5: Interaction processes utilised for different dark matter experiments. The arrows show the direction in which the interactions must take place for a given mode. The mediator and underlying interaction are unknown. Adapted from [16].

search for dark matter.

1.3.1 Indirect Detection

Indirect detection methods rely on the annihilation of dark matter particles producing Standard Model fermions, which can be observed by detectors, most commonly gamma rays, positrons, antiprotons, or neutrinos. The Bullet cluster indicates that, if the WIMP particle does self-interact, it does so very weakly and hence to observe such an effect one must look at areas of the Universe where the density of dark matter is higher; galaxies, galaxy clusters or the Galactic Centre itself.

Indirect detection searches are carried out by experiments including Fermi-LAT, H.E.S.S., PAMELA, IceCube and Super-Kamiokande [17–21]. These searches already rule out thermal relic annihilation in some low-mass WIMP scenarios. The strongest limits in this regime come from Fermi-LAT observations of dwarf spheroidal galaxies, which are particularly useful because of their low astrophysical backgrounds. At higher WIMP masses, searches towards the Galactic Centre with H.E.S.S. become more competitive. Nevertheless, over the broader GeV–TeV mass range, a substantial part of the WIMP parameter space remains uncharted, especially for annihilation channels involving neutrinos.

1.3.2 Collider Production

Particle colliders are able to perform WIMP searches through the pair production channels of Standard Model fermions to WIMPs. With the high collision rate produced in high-energy particle colliders, they are able to probe rare production channels despite their diminutive cross-sections and couplings. If WIMPs are produced in the interactions, they will not be visible to the surrounding detectors, which rely on electromagnetic interactions. For this reason, the detection of dark matter particles in colliders is often inferred from missing transverse momentum originating from some initial-state object such as a jet, photon or vector boson [22, 23].

Since the underlying interaction between the Standard Model particles and the dark matter is unknown, many theories have been devised as to the nature of these interactions. The simplest is the existence of a single mediator between the quarks or gluons and dark matter particles with some mass and coupling strength. Alternatively, there may exist visible mediators which produce dijet or dilepton resonances within the events. Many Higgs-portal searches are also sought after [24]. It is possible to conceive of a Higgs decay to dark matter which probes the low masses and small couplings.

Colliders have set strong limits on the mediator masses, coupling strengths, and Higgs decay branching ratios, yet a dark matter event remains to be observed. This detection method complements the direct detection method by probing the low mass, strong coupling strengths, whereas the direct detection methods thrive at high mass and weak coupling strengths.

1.3.3 Direct Detection

The final method of detecting dark matter is through direct scattering between Standard Model particles and dark matter. The conventional WIMP search considered here targets nuclear recoils. To enhance the spin-independent (SI) cross-section of the scattering interaction, we choose high-mass targets since the cross-section roughly scales with the square of the mass number, A^2 [25]. In order for a detector to be sensitive to spin-dependent (SD) scattering interactions, the target must have a total nuclear spin, $J \neq 0$. The most common targets for WIMP detectors are noble elements such as argon and xenon because of their large nuclei and chemical nobility.

Collisions between particles and the targets in these detectors produce three forms of detectable

signal: light, charge and heat. Experiments are thus designed to detect these signals. One detection method is to probe the annual modulation induced by the Earth's orbit around the Sun through the so-called 'dark matter wind'. This should cause a peak in the dark matter flux through Earth in June and a minimum flux in December, following a sinusoidal dependency with time. Searches like this most commonly use NaI crystal scintillation detectors in experiments like DAMA/LIBRA, ANAIS, COSINE and SABRE [26–29]. Expanding on this idea, some experiments utilise directionality to determine the source from which the dark matter particles would come. One example of this is the CYGNUS collaboration, aptly named since the dark matter wind is precisely in the direction of the Cygnus constellation [30]. The added directionality dimension allows for further discrimination of dark matter against backgrounds such as neutrons or neutrinos by ensuring that they originate from the direction of Cygnus.

Currently, the most sensitive detectors take the form of tonne-scale liquid noble dual-phase time projection chambers (TPCs) which are sensitive to light and charge but not heat. Examples of these include LUX-ZEPLIN (LZ), XENON, PandaX and DarkSide [31–34]. They aim to detect direct scattering interactions with dark matter by searching for excesses above extremely low background environments. To achieve these low backgrounds, these TPCs are placed in underground laboratories around the world. Additionally, this technology is able to distinguish between nuclear recoil (NR) and electronic recoil (ER) scatters by measuring the ratio of light yield to charge yield, allowing for large background rejection from beta and gamma interactions.

1.3.3.1 Recoil Kinematics

The sensitivity of the direct detection experiments is governed by the kinematics of elastic scattering interactions between the dark matter particles and the target nuclei. For a WIMP travelling at speed v, the maximum nuclear recoil energy that can be produced is given by

$$E_R^{max} = \frac{2\mu_{\chi N}^2 v^2}{m_N}, \tag{1.6}$$

where $\mu_{\chi N}$ is the WIMP-nucleus reduced mass given by

$$\mu_{\chi N} = \frac{m_\chi m_N}{m_\chi + m_N}, \tag{1.7}$$

m_N is the mass of the nucleus and m_χ is the WIMP mass. Therefore, in general, lighter dark matter masses produce less nuclear recoil energy, inherently making detection more difficult.

The differential event rate can be expressed as

$$\frac{dR}{dE_R} = \frac{\rho_\chi \sigma_{\chi N}}{2 m_\chi \mu_{\chi N}^2} F^2(E_R) \int_{v_{min}}^{\infty} \frac{f(v)}{v} dv, \tag{1.8}$$

where $\rho_\chi \simeq 0.3\,\mathrm{GeV\,cm^{-3}}$ is the local dark matter density, $\sigma_{\chi N}$ is the WIMP-nucleus scattering cross-section, $F(E_R)$ is the nuclear form factor, v_{min} is the minimum WIMP speed necessary to produce a nuclear recoil with energy E_R, and $f(v)$ is the dark matter velocity distribution. Thus, it defines a direct relationship between the event rate, nuclear recoil energy, WIMP-nucleus cross-section and WIMP mass that can be utilised to produce limits from observed events.

1.3.3.2 WIMP Parameter Space

The WIMP parameter space, shown in Figure 1.6, has been ruled out by many experiments spanning multiple decades of progress, with recent limits set by LZ, PandaX and XENONnT all taking their first steps into the neutrino fog. The neutrino fog is an irreducible background to WIMP search experiments since neutrino interactions with xenon closely mimic dark matter interactions. Though the neutrino fog is still a searchable parameter space for WIMPs, the inclusion of neutrino backgrounds causes the sensitivity to WIMPs to degrade and parameter degeneracies become important. While the remaining space between the current limits and the neutrino fog may appear small, it will take many years of dedicated novel analysis techniques to cover. Additionally, the low-mass WIMP provides a large amount of parameter space that is yet to be probed. The mass limits on the left-hand side of each line are determined by the threshold set during the analysis. In principle, these detectors are sensitive to individual electrons ionised from the atoms in the tonnes of liquid xenon present. The thresholds are therefore often limited by persistent low-energy backgrounds. The purpose of this thesis is to introduce novel analysis techniques designed to improve the discovery potential of LZ and other experiments to low-mass WIMPs and enter uncharted territory for liquid xenon TPCs.

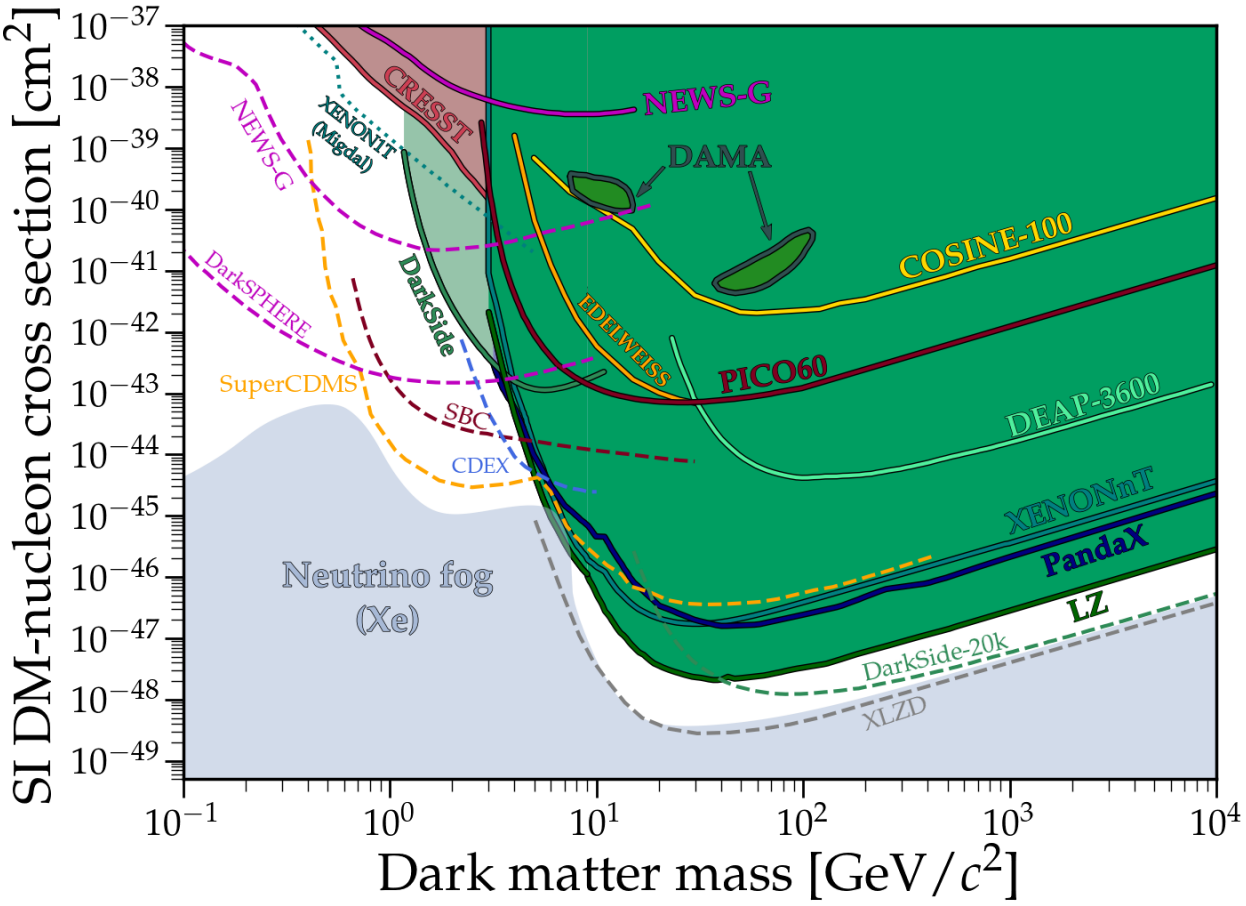


Figure 1.6: DM-nucleon spin-independent cross-section limits (solid lines) versus dark matter mass set by the experiments labelled. The parameter space excluded by these experiments is shaded above the lines. The neutrino fog is shown for a xenon target in the shaded grey at the bottom. Projected limits are shown for various experiments with the dashed lines. White space indicates parameter space which is yet to be explored. Adapted from [35] to include the latest LZ results down to 3 GeV [1].

2 | The LUX-ZEPLIN Experiment

The LUX-ZEPLIN (LZ) experiment is a direct dark matter detection experiment operating $1.5\,\mathrm{km}$ underground at the Sanford Underground Research Facility (SURF) in South Dakota, USA. It currently sets the strongest limits for SI WIMP-nucleon scattering cross-sections for WIMP masses above $5\,\mathrm{GeV/c^2}$, shown in Figure 1.6, with the strongest SI exclusion of $2.2\times10^{-48}\,\mathrm{cm^2}$ at the 90% confidence level, and a best median sensitivity of $5.1\times10^{-48}\,\mathrm{cm^2}$ for a $40\,\mathrm{GeV/c^2}$ mass WIMP [36]. This result is produced from 280 live-days of run time. LZ aims to accumulate 1000 live-days of run time with a projected sensitivity of $1.6\times10^{-48}\,\mathrm{cm^2}$ for the SI WIMP-nucleon interaction of a $40\,\mathrm{GeV/c^2}$ WIMP. The spin-dependent WIMP-nucleon scattering is possible within LZ when scattering occurs with a ^{129}Xe and ^{131}Xe nuclei. Both have an unpaired neutron which gives sensitivity to WIMP-neutron scattering cross-sections. In principle, this signal and methodology remains the same as the SI interactions with the main difference being just the signal model which is passed into the PLR test.

LZ also boasts sensitivity to neutrino detection, publishing the first evidence of ^{8}B coherent elastic neutrino-nucleus scattering (CEνNS) interactions observed with xenon at 4.5σ significance [1]. Additionally, LZ is capable of performing other rare event searches such as the neutrinoless double beta decay ($0\nu\beta\beta$) of the ^{136}Xe isotope, with a projected sensitivity of 1.06×10^{26} year half-life for the decay process by the end of the 1000 live-day run [37]. This constraint is similar to other experiments dedicated to this search [38, 39].

Such stringent limits are achievable as a result of decades of technological advancements and background mitigation techniques honing the state-of-the-art detection methods to what they are today. Rare-event searches such as these require extremely low backgrounds to observe the ~ 5 WIMP events expected per year in the detector. Placing the detector underground at SURF provides a $4300\,\mathrm{m}$ water-equivalent overburden to shield from cosmic rays. In addition, veto detectors and xenon self-shielding in the TPC ensure that the detector centre remains an exceptionally low-background environment.

2.1 The LZ Detector

The heart of LZ is a liquid/gas dual-phase xenon TPC containing $7\,\mathrm{t}$ of liquid xenon encompassed within two veto detectors, depicted in Figure 2.1. The TPC is $1.456\,\mathrm{m}$ in diameter and height [40]. Specifically, this is the distance between the cathode high-voltage grid mesh wires near the bottom of the detector and the gate high-voltage grid mesh wires $5\,\mathrm{mm}$ below the liquid surface near the top of the detector. This is the maximum drift length through the liquid xenon. $8\,\mathrm{mm}$ above the liquid surface lie the anode grid wires in the gaseous xenon. These three grids control the electric field present in the liquid and gaseous xenon. For the latest WIMP search results, the cathode and gate produce a $97\,\mathrm{V\,cm^{-1}}$ field in the bulk liquid xenon, drifting the ionised electrons up towards the liquid surface. The gate and anode produce a $6.5\,\mathrm{kV\,cm^{-1}}$ field in the gaseous xenon. The top and bottom of the TPC are instrumented with a combined total of 494 Hamamatsu R11410-22 3-inch photomultiplier tubes (PMTs), 253 in the top array and 241 in the bottom array, sensitive to VUV light produced by xenon scintillation. The walls of the TPC are constructed from slotted segments of polytetrafluoroethylene (PTFE) that house the field-shaping rings designed to maintain consistent field lines throughout the detector volume. The TPC is held within the ultra-radiopure titanium inner cryostat vessel (ICV), which itself is enclosed inside the vacuum-insulated outer cryostat vessel (OCV). These allow for temperature control of the xenon, which is circulated at $175\,\mathrm{K}$. The xenon is circulated through a heated zirconium getter to maintain high purity.

The space between the TPC walls and the ICV contains an additional $\sim 2\,\mathrm{t}$ of liquid xenon, which acts as an active veto capable of tagging gamma and neutron backgrounds originating from the radiogenic detector materials. The skin region is instrumented with an additional 113 PMTs to observe interactions in this volume. Beyond the OCV resides the Outer Detector (OD) which consists of 9 acrylic vessels, highlighted in green in Figure 2.1, containing $17.5\,\mathrm{t}$ of Gd-loaded liquid scintillator designed to detect and veto neutrons which have been thermalised in the surrounding $5.92\,\mathrm{m}$ tall and $7.62\,\mathrm{m}$ wide water tank. The water tank contains $228\,\mathrm{t}$ of ultrapure deionised water, which provides passive shielding against radiation from the surrounding rock cavern and moderates passing neutrons, allowing for efficient neutron capture in the Gd-loaded liquid scintillator. 120 8-inch dome PMTs detect the scintillation light produced by these neutron-capture events. The combined veto efficiency of these systems is $>88\%$.

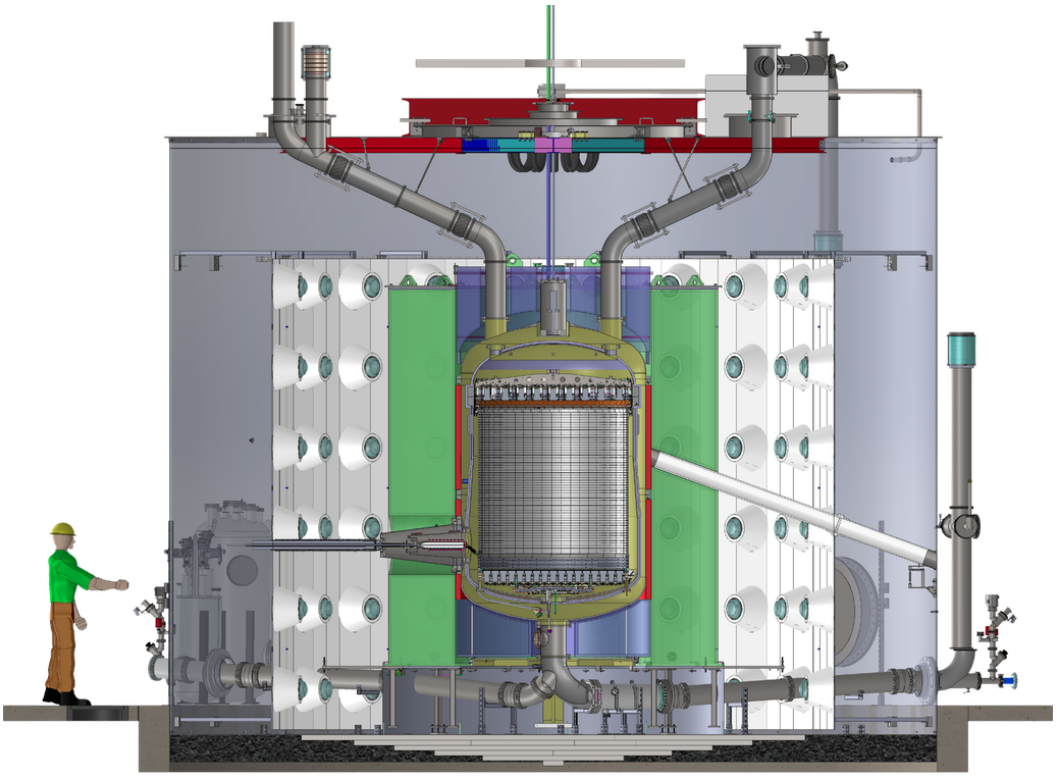


Figure 2.1: Technical rendering of a cross-section of the LZ detector in the Davis Cavern at SURF. The TPC is displayed in the centre of the figure, with the skin veto immediately surrounding it. The green regions show the Gd-loaded liquid scintillator acrylic tanks. The white region is the water tank instrumented with 120 PMT domes.

2.2 Signal Model

Scattering interactions in the liquid xenon in the TPC produce prompt scintillation light and ionised electrons. Prompt scintillation photons are produced when an incoming particle transfers energy to the xenon atom, causing it to form excimer states (Xe_2^*) which de-excite and release VUV photons of wavelength $\sim$175 nm. When energy is transferred by a direct excitation from some particle χ, these excimers decay in the following process,

$$
\begin{aligned}
Xe + \chi &\rightarrow Xe^* + \chi, \\
Xe^* + Xe &\rightarrow Xe_2^\sigma, \\
Xe_2^\sigma + Xe &\rightarrow Xe_2^* + Xe, \\
Xe_2^* &\rightarrow Xe + Xe + \gamma,
\end{aligned}
$$

where σ superscripts denote possible singlet ${}^1\Sigma_u^+$ and triplet ${}^3\Sigma_u^+$ excited states which decay with global time averages of the singlet decay time, $\tau_s = 3.1 \pm 0.7\ ns$ and triplet decay time $\tau_t = 24 \pm 1\ ns$ in liquid xenon [41]. Alternatively, the collision can directly ionise the xenon atom, which may then undergo electron-ion recombination and release additional photons in the following manner

$$
\begin{aligned}
Xe + \chi &\rightarrow Xe^+ + e^- + \chi, \\
Xe^+ + Xe + Xe &\rightarrow Xe_2^+ + Xe, \\
Xe_2^+ + e^- &\rightarrow Xe^{**} + Xe, \\
Xe^{**} + Xe &\rightarrow Xe^* + Xe + \text{heat}, \\
Xe^* + Xe + Xe &\rightarrow Xe_2^* + Xe + \text{heat}, \\
Xe_2^* &\rightarrow Xe + Xe + \gamma.
\end{aligned}
$$

The electron-ion recombination mechanism is suppressed in LZ by the electric field, which separates the xenon ion and the ionised electron, preventing them from recombining. This suppresses the electron-ion recombination but does not eliminate it entirely, impacting the light-to-charge ratio of the interaction by reducing the number of ionised electrons that escape recombination and increasing the number of photons emitted. The photons released are detected by the PMT arrays with a detection efficiency of $\sim$10% and form the first signal pulse in an event; an *S1* pulse.

Electrons ionised from the xenon atoms during the collision drift upwards in the field produced by the cathode and gate in the liquid xenon. Once these electrons pass the gate, they enter the extraction region where they are accelerated by a stronger field, $\sim 3.5\,\mathrm{kV\,cm^{-1}}$, in order to have sufficient energy to leave the liquid surface and pass into the gaseous xenon. Electrons that make it into the gas are accelerated towards the anode grid wires and collide with the gaseous xenon atoms in transit. This produces secondary scintillation through the same direct excitation of xenon atoms as previously shown. The light yield of this electroluminescence process is amplified due to the strong electric field applied. These photons are also detected by the PMT arrays and form the secondary signal pulse in the event, the *S2* pulse.

The combination of these two pulses encodes all information necessary for analysis within LZ. S2 pulses consist of sufficiently many photons that the transverse (x, y) position of the initial collision

can be reconstructed using the hit pattern of the top array PMTs. With a known drift field, and thus drift speed, the time between the S1 and S2 pulses allows for position reconstruction in the z direction, enabling full 3D reconstruction of the event position. Additionally, the ratio of the size of the S1 and S2 pulses allows for NR/ER discrimination and energy reconstruction, both of which are detailed in Section 2.3.

An example event is shown in Figure 2.2, showing the schematic exponentially modified Gaussian shape of the S1 pulse and the Gaussian shape of the S2 pulse. The shapes of both the S1 and S2 pulses are dependent on the interaction type, energy and position of the interaction. A full description of the microphysics pertaining to the S2 pulse is given in Chapters 4 and 5.

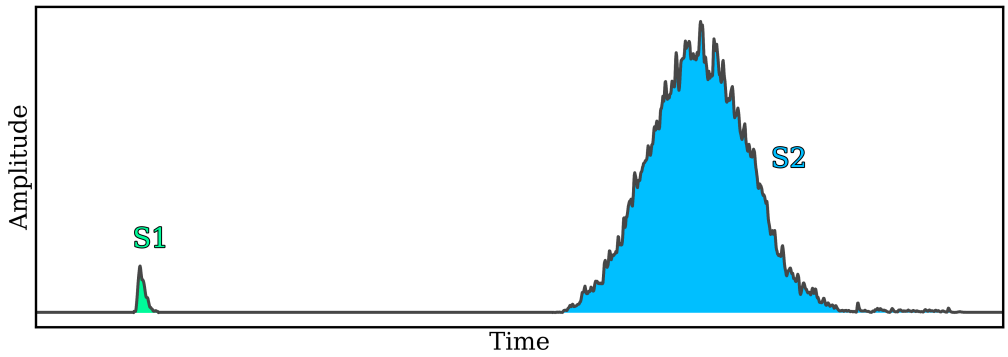


Figure 2.2: An example event containing a single S1 and S2 waveform annotated as such.

LZ analysis is performed on Reduced Quantities (RQs) extracted from these waveforms, such as the pulse area and width. These are extracted with the LZ Analysis Package (LZAP), detailed in Chapter 3. The pulse areas are recovered by integrating the area under the pulse and are calibrated to measure the number of photons detected (phd) in the PMTs. Geometric dependencies are corrected for with 3D maps which normalise the sizes of the pulses relative to the centre of the detector. Herein, $S1_c$ and $S2_c$ denote the spatially corrected S1 and S2 pulse areas measured in phd.

The event displayed in Figure 2.2 is an example of a clean single scatter (SS) event. Single scatter events are defined as events with a single prominent S1 and a single prominent S2 pulse within the event window, indicating a single point-like energy deposit in the liquid xenon. A prominent pulse is significantly larger in size and width than other S1s or S2s that may occur within the same event window. This helps retain SS signal events even if there is some spontaneous electron emission

from the grid wires, producing an additional, smaller S2, for example. This distinction will become important in Chapter 6 when discussing fragmented S2s. The SS is the expected topology for a WIMP particle due to its inherently low probability of interacting with Standard Model particles. However, other sources (e.g., high-energy gammas and neutrons) can scatter multiple times within the xenon. Multiple scatter (MS) events are classified as events containing a single prominent S1 and multiple prominent S2 pulses. Only one S1 is allowed in an MS event, since if multiple S1s are present, it is impossible to determine whether the event is produced by a single incident particle scattering multiple times in the detector or by multiple incident particles each scattering once in the detector. Events with multiple S1s and S2s are labelled 'Pileup Scatters'. Since LZ triggers on the S2 pulse, it is able to observe events with no S1 present, named S2-Only (S2O) events. Predominantly, LZ utilises the SS events to search for WIMPs and other rare events, but the MS and S2O datasets may be utilised in future for accompanying searches. Details of the S2O searches are given in Chapter 7.

2.3 Calibrations

Extensive detector calibrations characterise the detector response at all energies [42]. For LZ, this involves using a combination of internal and external calibration sources to initiate both ER and NR responses, which are characterised for analysis. LZ also conducts monthly PMT calibrations using internal LED lights, which monitor the response and health of the PMTs throughout the lifetime of the detector. In particular, we monitor the gain for PMT response calibration and the afterpulsing rates to assess the health of the PMTs. The results of these calibrations are directly utilised by LZAP to apply temporal corrections in situ to the pulse areas observed by each PMT.

Once the PMTs are accurately calibrated, the ER and NR responses must follow. There are 3 methods of source deployment within LZ. Firstly, we may inject the source directly into the xenon circulation system. This can only be performed for sources which can be removed effectively or decay with a very short half-life so as not to contaminate the purity of the xenon. Secondly, the detector is outfitted with 3 calibration source deployment (CSD) tubes spaced 120° apart which allow sources to be placed next to the detector ($12\,\mathrm{cm}$ away) in the inter-space between the ICV and OCV. The CSD tubes include a stepper motor which can lower the source from above to any height along the TPC. Finally, for dedicated deuterium-deuterium (DD) neutron calibrations, the

DD conduit is shown in Figure 2.1 as the diagonal white tube on the right-hand side. A non-exhaustive list of useful calibration sources and their deployment methods is given in Table 2.1.

	Source	Particle species used	Deployment	Energy (keV)	Half-life
ER	^{3}H	β	Injection	0-18.6	12.3 y
	^{83m}Kr	IC,AE,X-ray,γ	Injection	32.1 and 9.4	1.83 h
	^{131m}Xe	γ,X-ray	Injection	163.9	11.8 d
	^{220}Rn	α, β, γ	Injection	Various	55.6 s
	^{228}Th	γ	CSD	2615	1.91 y
NR	DD	n	DD conduit	2450	-
	^{241}AmLi	(α, n)	CSD	(5638,0-1500)	433 y
	^{241}AmBe	(α, n)	CSD	(5638,0-11 $\times$ 10^3)	433 y

Table 2.1: Common calibration sources deployed in LZ separated into ER and NR responses. IC refers to internal capture, and AE refers to Auger electrons. Note that the energies listed for each source are the energies relevant for LZ and are not exhaustive. Adapted from [42].

Monoenergetic sources are particularly useful for calibrating key detector parameters such as the photon and electron gains, g_1 and g_2. These are defined by

$$S1_c = g_1 n_\gamma, \qquad S2_c = g_2 n_e, \tag{2.1}$$

where n_γ is the number of scintillation photons produced, and n_e is the number of ionisation electrons produced. For LZ $g_1 \simeq 0.1\ phd/\gamma$ and $g_2 \simeq 33\ phd/e^-$. The energy of the interactions can be reconstructed from the measured $S1_c$ and $S2_c$ sizes with

$$E = \frac{W}{\mathcal{L}_i(E)} \left(\frac{S1_c}{g_1} + \frac{S2_c}{g_2} \right), \tag{2.2}$$

where $W = 13.7\,\mathrm{eV}$ is the average work done to produce a scintillation photon or ionisation electron from xenon [43, 44] and $\mathcal{L}_i(E)$ is the energy-dependent Lindhard quenching factor for interaction type $i \in \{ER, NR\}$. For ER recoils, $\mathcal{L}_{ER}(E) \simeq 1$, but for NR recoils this is an energy-dependent factor with $\mathcal{L}_{NR}(10\ keV) = 0.17$ [45]. With a series of known monoenergetic sources, one can produce a Doke plot to fit the values of g_1 and g_2, as shown in Figure 2.3. Hence, in principle, full recoil energy reconstruction can be applied to any single scatter event with these fitted parameters.

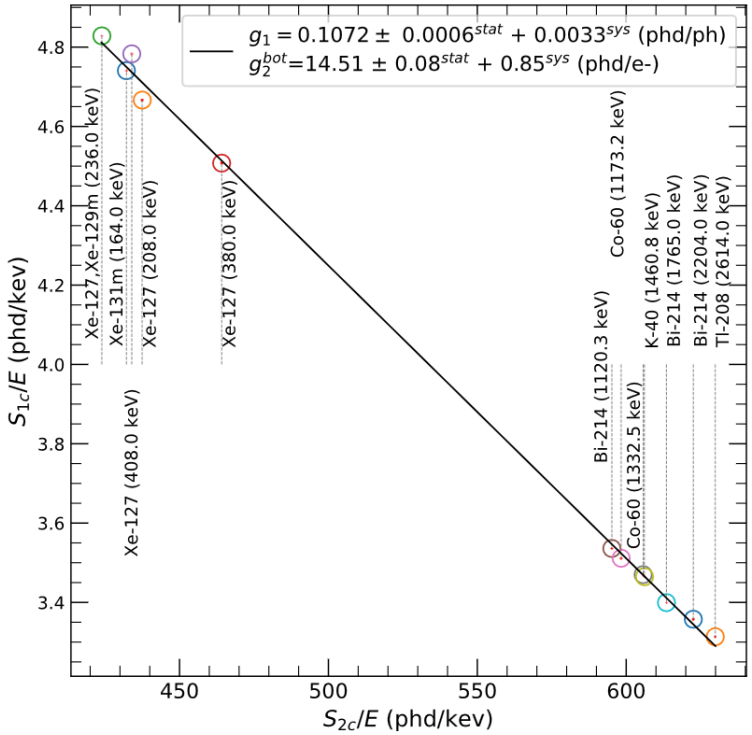


Figure 2.3: Doke plot of $S1_c$ per unit energy against $S2_c$ per unit energy to extract the values of detector parameters g_1 and g_2^{bot} for the WS2022 detector configuration [46]. g_2^{bot} refers to the electron gain as measured from the bottom PMT array. This is performed to circumvent PMT saturation occurring in the top array. Generally, $g_2 = 3g_2^{bot}$. Figure from [47].

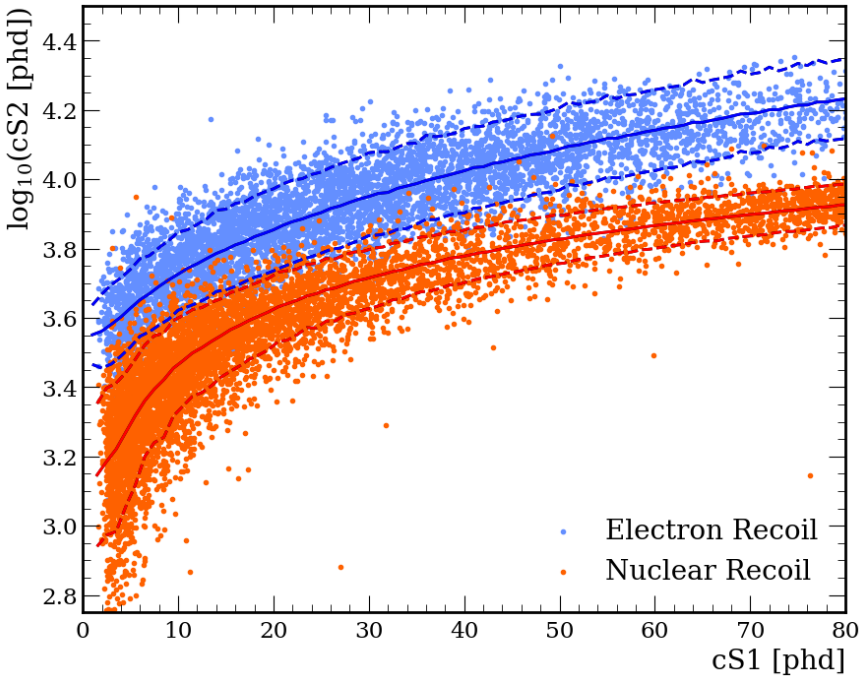


Figure 2.4: ER and NR bands in $S1$-$log(S2)$ space, as is custom for LZ, produced by tritium and DD calibrations, respectively. There is 99.9% discrimination of ER and NR events below the $40\,\mathrm{GeV/c^2}$ WIMP median.

For the WIMP-search 2024 (WS2024) result, the WIMP region was calibrated using the tritium (^{3}H) source for the ER response and DD for the NR response. The ER and NR bands are shown in Figure 2.4 with 99.9% discrimination between NR and ER events below the $40\,\mathrm{GeV/c^2}$ WIMP median. Physically, the ratio of the S1/S2 is used to discriminate between the ER and NR interactions with the understanding that the energy deposited will partition between light and charge. Since ER interactions happen directly with the electrons, they will produce a larger ionisation cloud and produce larger S2s for the same size S1. The ER and NR bands are used in simulations to produce the signal model PDFs for various searches, as well as allowing for light yield, L_y, and charge yield, Q_y, measurements which are provided to the Noble Element Simulation Technique (NEST) software. NEST is heavily utilised for simulations and will be discussed in detail in Chapter 3.

2.4 Backgrounds

A strict understanding of the background sources present in the detector is critical to establishing any potential discovery power. The majority of backgrounds within LZ can be separated into the radiogenic materials of the detector itself, the gamma radiation from the rocks in the cavern and irreducible physical backgrounds. The material components from which the detector is constructed were all radio-assayed and thus the radiogenic backgrounds are well understood [40, 48]. The dominant decay chains present in these materials are those of ^{238}U and ^{232}Th along with ^{40}K and ^{60}Co. Within these decay chains, there exist many γ-decays which populate the ER band, but also many (α, n) and spontaneous fission reactions which can mimic NR signatures. The strong self-shielding properties of xenon mostly constrain these backgrounds to occur near the detector walls, motivating a simple fiducial volume to ensure these backgrounds constitute $<$1% of the expected background events in the final dataset.

For the WS2024 result, the background model (pre- and post-fit) is given in Table 2.2. This table demonstrates the strength of LZ's background modelling with good agreement between expectation and the final fit. It is also clear that the largest contribution in this region of interest (ROI) resides in the ER band, with several β and γ sources contributing 97.8% of all backgrounds, dominated by the contribution from ^{214}Pb βs.

When approaching the low-energy regime, which will be outlined in more detail in Section 2.7,

Source	Pre-fit Expectation	Fit Result
^{214}Pb βs	743 ± 88	733 ± 34
^{85}Kr + ^{39}Ar βs + det. γs	162 ± 22	161 ± 21
Solar ν ER	102 ± 6	102 ± 6
^{212}Pb + ^{218}Po βs	62.7 ± 7.5	63.7 ± 7.4
Tritium + ^{14}C βs	58.3 ± 3.3	59.7 ± 3.3
^{136}Xe $2\nu\beta\beta$	55.6 ± 8.3	55.9 ± 8.2
^{124}Xe DEC	19.4 ± 2.5	20.4 ± 2.4
^{127}Xe + ^{125}Xe EC	3.2 ± 0.6	2.7 ± 0.6
Accidental coincidences	2.8 ± 0.6	2.6 ± 0.6
Atm. ν NR	0.12 ± 0.02	0.12 ± 0.02
^{8}B + *hep* ν NR	0.06 ± 0.01	0.06 ± 0.01
Detector neutrons	$0.0^{+0.2}$	$0.0^{+0.2}$
Total	1210 ± 91	1202 ± 41

Table 2.2: Background table for the relevant backgrounds of the WS2024 result, including the pre-fit expectation and post-fit result. The detector neutrons pre-fit expectation is derived from a sample of veto-tagged events and is not explicitly used in the final combined fit. Table from [36].

the NR band lies sufficiently far below the ER band that the ER contribution to the backgrounds is negligible. Instead, the dominant background becomes accidental coincidences. These are events in which an isolated S1 and a separate isolated S2 are coincident in time such that they resemble an SS event. A suite of data-quality selection criteria is used to remove these events by leveraging the underlying physical processes necessary to form S1 and S2 pulses. With respect to each other, the S2 shape may be inconsistent with the inferred drift time of the S1, for example. However, these analysis cuts are not 100% efficient, and thus a dedicated model for these events is devised using events with drift times longer than the maximum drift time of the detector ($\sim$1050 µs for WS2024) as a proxy for the distribution of these events. This is then modelled and extrapolated to physical drift times to estimate the number of accidentals present in the final dataset. This is the main background for low-mass WIMP and ^{8}B CEνNS searches.

2.5 Bias Mitigation

When performing searches of this nature, it is important to remain vigilant against introducing bias in the search results. Within the field of dark matter detection experiments, two methods of bias mitigation are frequently utilised: blinding and salting. Blinding, as is performed in collider experiments such as ATLAS and CMS, involves restricting analysers' access to a small percentage

of the total dataset or to a dataset outside the signal region. It is with this dataset that the analysers must finalise the framework of their analysis, which will be applied to the entire dataset. With this approach, analysers are less able to tune the framework to appear more sensitive than it truly is. Blinding requires that there is sufficient data such that analysis can be performed on the subset with a reasonable degree of accuracy.

Dark matter detectors are often statistics-limited and thus favour salting as a bias mitigation technique. The salting technique involves injecting synthesised signal events into the experiment's raw data stream. This allows analysers access to the entire dataset with the caveat that they do not know whether each event is real or salt. The salt events are constructed by sampling an S1 pulse from one calibration event and an S2 from another and stitching them together to replicate a single signal-like event. Extensive campaigns are conducted to demonstrate that the salt events cannot be easily uncovered during analysis. Once the analysis framework has been completed, it is frozen, and the salt events are removed from the dataset before the final statistical inference results are determined.

2.6 WS2022 + WS2024 Results

LZ's strongest limit to date was set using a total exposure of 4.2 tonne-years. The final dataset, including signal and background contours, is shown in Figure 2.5. The region of interest for this search is defined with the bounds of $3 \leq S1_c \leq 80\ phd$ and $2.8 \leq log_{10}(S2_c) \leq 4.4$.

From this final dataset, from the result of a Profile Likelihood Ratio (PLR) test, LZ reports the world-leading SI WIMP-nucleon cross-section ($\sigma_{SI}^{N\chi}$) for WIMP masses $\geq 9\,\mathrm{GeV/c^2}$, shown in Figure 2.6, with the strongest exclusion at $2.2 \times 10^{-48}\,\mathrm{cm^2}$ for a $40\,\mathrm{GeV/c^2}$ WIMP.

2.7 ^{8}B CEνNS/Low Mass WIMP Results

Extending to lower recoil energies, the latest low-mass WIMP results are made in the region bound by $2 \leq S1_c \leq 15\ phd$ and $3.5 \leq S2 \leq 14.5$ electrons. Note the transition here from measuring the S2 size in terms of electrons rather than phd. This choice is made because we approach the limit where counting individual detected electrons is possible, which is unavailable at higher energies due to Gaussian smearing. In this region, the nuclear recoil spectrum of a $6\,\mathrm{GeV/c^2}$

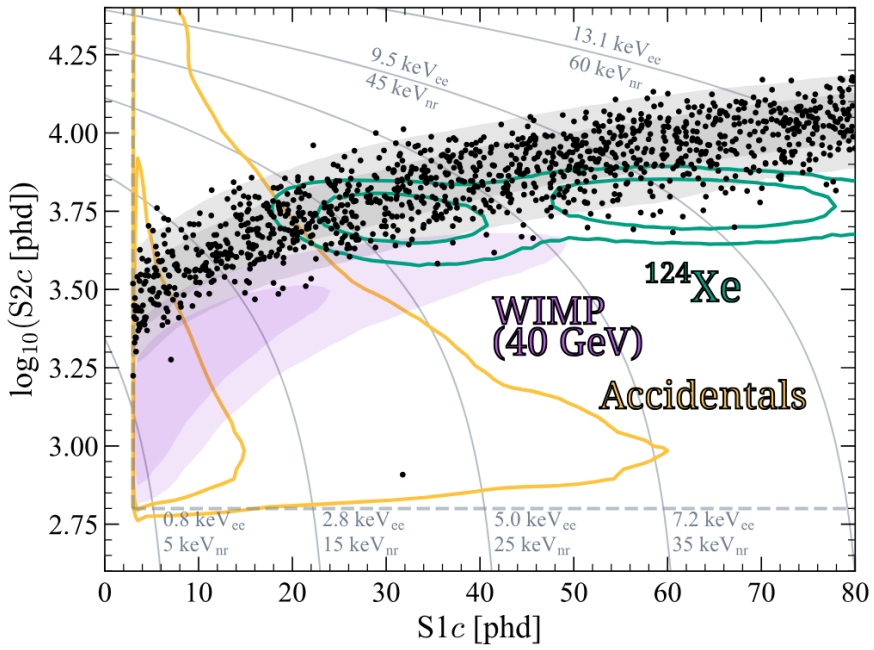


Figure 2.5: Final dataset (black points) from the WS2024 result in $S1$-$log(S2)$ space. Also shown are the contours of the $40\,\mathrm{GeV}$ WIMP signature as well as the accidentals and ^{124}Xe double electron capture backgrounds. The ER band is shaded in grey. Recoil energy lines are given in the solid grey lines.

WIMP and a ^{8}B CEνNS interaction are near-degenerate. Additionally, the accidentals contour and the WIMP contour occupy the same space in the S1-S2 regime. This makes both the ^{8}B CEνNS and accidentals challenging backgrounds for this search. With the final dataset, shown in Figure 2.7, LZ reports new limits for the SI WIMP-nucleon cross section for WIMP masses from 3–9 $\mathrm{GeV/c^2}$, providing world-leading $\sigma_{SI}^{N\chi}$ limits for WIMP masses above $5\,\mathrm{GeV/c^2}$ as well as the first observed evidence of ^{8}B CEνNS interactions on xenon with 4.5σ significance.

This analysis marks a significant improvement in decreasing the threshold of LZ and increasing sensitivity to light dark matter masses, overcoming many challenges posed by new detector effects and background modelling. Chapters 6 and 7 detail some approaches utilised in this effort and outline future techniques that may be used to continue to push the boundaries of what LZ is capable of.

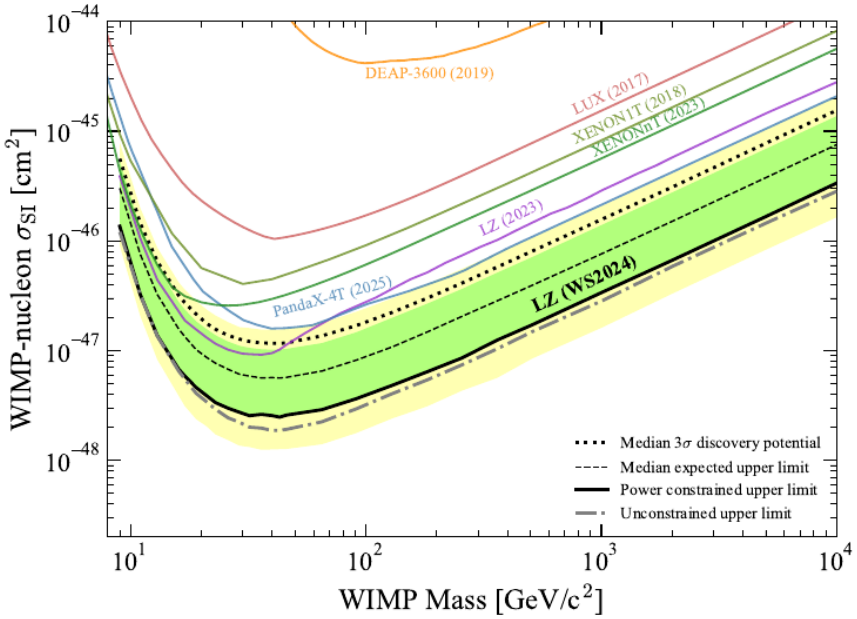


Figure 2.6: 90% confidence level SI WIMP-nucleon cross-section limit as a function of WIMP mass from the WS2024 dataset in Figure 2.5. The solid black line shows the -1σ power-constrained limit, while the dashed grey line shows the unconstrained limit. The green and yellow shaded regions show the range of expected upper limits from 68% and 95% of background-only experiments, while the dashed black line shows the median expectation, obtained with post-fit background estimates. Figure from [36].

2.8 Towards XLZD

Dark matter searches of this kind are predominantly limited in their sensitivity by the exposure they can accumulate over time. This can be increased in two ways: run the experiment for a longer time or build an experiment with a larger target mass. LZ is set to run through to 2028 with a projected sensitivity of $1.6 \times 10^{-48}\,\mathrm{cm}^2$ for a $40\,\mathrm{GeV/c}^2$ WIMP, but this limit will not cover the entire WIMP region, which motivates the necessity to either extend the run time or build a larger experiment. However, LZ is approaching the limit at which additional exposure from an increased run time will contribute very little to improving sensitivity.

Instead, the collaborators from LZ, XENONnT and DARWIN have proposed a next-generation experiment, XLZD, with an increased proposed xenon active target mass of 60–80 t [49]. The increased sensitivity has a projected 3σ evidence potential reaching approximately $3 \times 10^{-49}\,\mathrm{cm}^2$ for a $40\,\mathrm{GeV/c}^2$ WIMP. XLZD is projected to advance into the neutrino fog, shown in Figure 1.6. This allows the experiment to be designed not only as a dark matter detection experiment but as

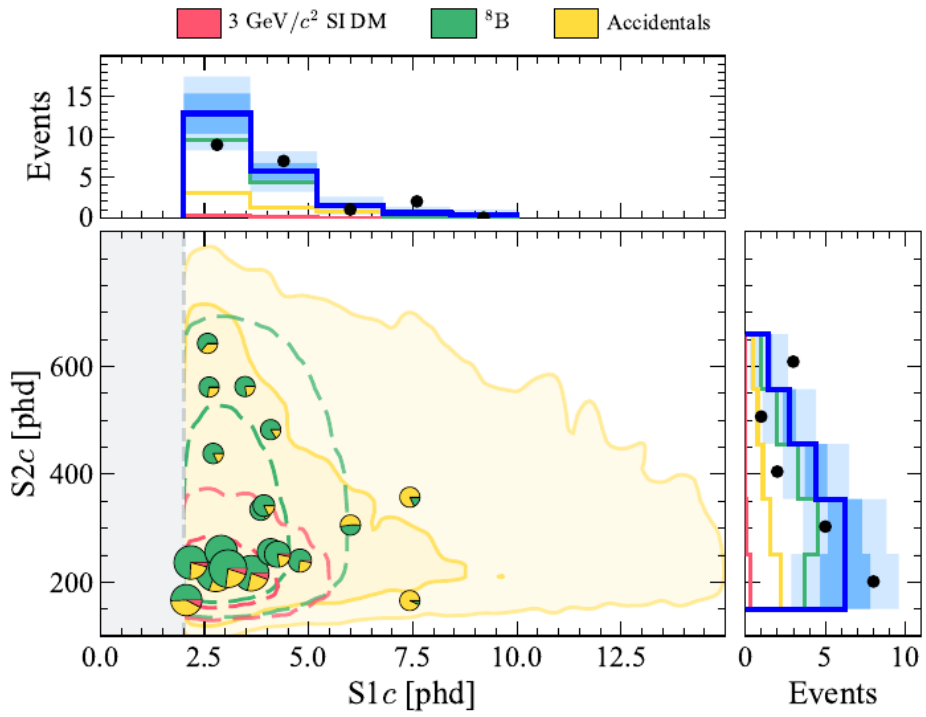


Figure 2.7: Final dataset used for both the ^{8}B and low-mass WIMP searches. 1σ and 2σ contours are shown for the accidentals background (yellow), ^{8}B neutrino CEνNS signal (green) and a $3\,\mathrm{GeV/c^2}$ WIMP (pink). Each data point is represented by a pie chart indicating the fit contributions of each event to each model. The relative size of each pie chart is proportional to its contribution to the $3\,\mathrm{GeV/c^2}$ WIMP profile. The data and best fit models are shown as projections in both $S1_c$ and $S2_c$. Figure from [1].

a xenon observatory with capabilities for WIMP search, neutrino measurements and searches for $0\nu\beta\beta$ decays, making this an extremely powerful scientific tool for future generations.

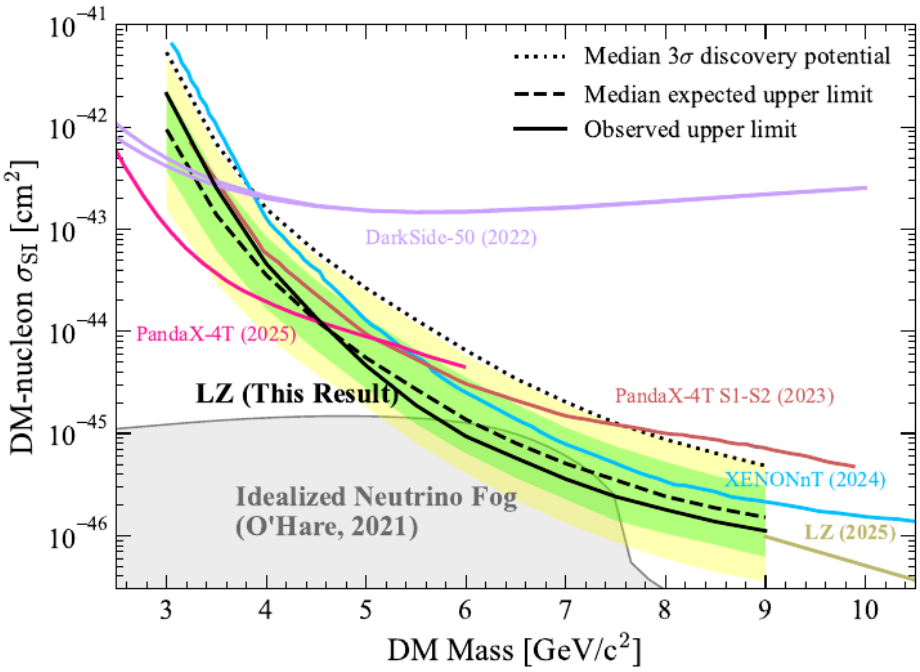


Figure 2.8: Low-mass WIMP upper limit (90% CL) observed from the dataset shown in Figure 2.7 with an exposure of 5.7 tonne-years (black line). The green and yellow bands show the upper limits of background-only for 68% and 95% of the experiments, respectively. The dashed black line shows the median expected upper limit, while the dotted black line shows the median 3σ discovery potential, using the post-fit model. Other experimental limits are also shown. The shaded grey region depicts the neutrino fog. Figure from [1].

3 | LZ Simulation Packages

Simulation techniques are fundamental to validating our understanding of physics within the LZ detector. The LZ simulation chain operates in two modes, often referred to as the fast chain and the full chain. The results presented in this thesis will focus on the improvements to the full chain simulation and the studies these improvements make possible. For completeness, we include a brief description of the difference between the fast and full simulation chains of LZ [50].

Simply put, the fast chain uses LZ's GEANT4 wrapper, BACCARAT (Basically A Component-Centric Analogue Response to AnyThing), to simulate particle interactions within the detector and save all energy deposits, which are then propagated through the LZLAMA package. LZLAMA is a fully parametric modelling of all post-interaction physics within the detector and is favoured for its speed. However, what LZLAMA gains in speed it loses in rigour. It is capable of producing large statistical distributions, $\mathcal{O}(10^8)$ events, in a matter of hours and is often used for producing PDFs of signal and background models. However, LZLAMA does not construct the pulse waveforms or additional effects (e.g. dark counts, PMT afterpulsing) in the event window. Since it lacks the real waveforms and PMT effects, there is often information available in the data that is not available in the fast chain simulation. This is where the full chain simulation is required.

The full chain utilises the full capability of BACCARAT to model the prominent detector processes, up to and including the photon propagation to the PMT windows, and will be discussed further in Section 3.1.1. The resulting simulated data is then passed through the Detector Electronics Response (DER) software, which simulates all effects that occur in the PMTs, electronics, and data acquisition chain. This is explained in detail in Section 3.1.2. The DER then produces a 'raw' data file with the same structure as that produced by the real data acquisition chain, which can be processed by the LZ Analysis Package (LZAP). This processing is identical to that of the real data and produces all possible variables and, crucially, the waveforms. The main drawback of the full chain simulation is that it requires considerable computing power due to its complexity and hence should be used for smaller scale productions, $\mathcal{O}(10^2 - 10^4)$ events, mainly to investigate effects directly from waveforms or from RQs that are not available in the fast chain output.

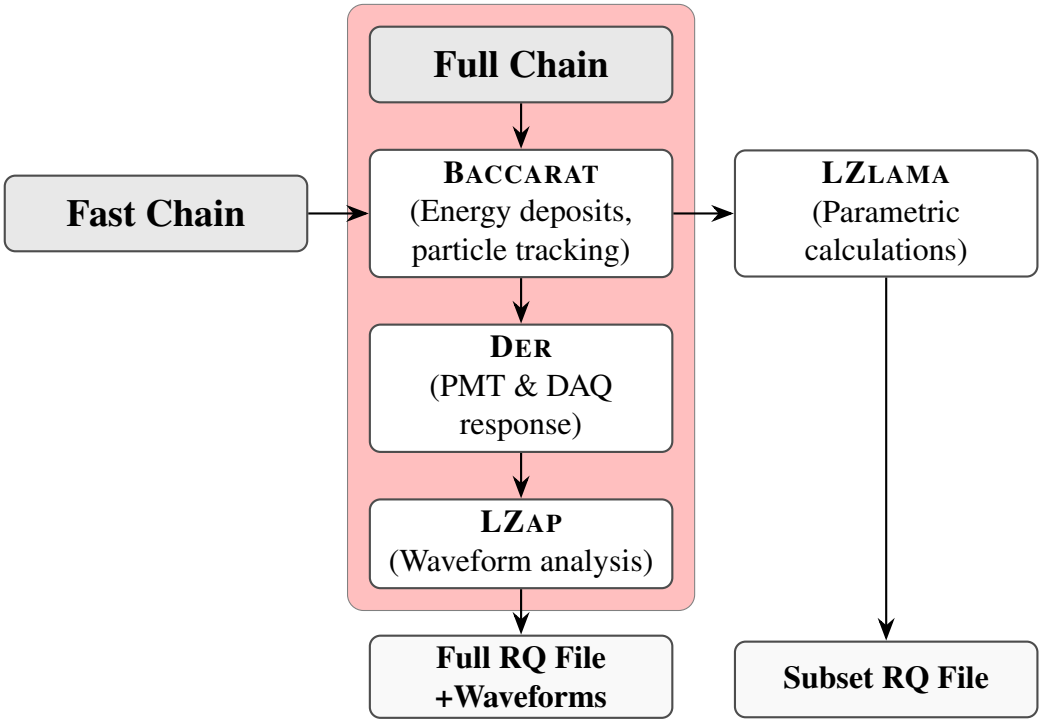


Figure 3.1: Simplified schematic showing the flow of both the full and fast chain methodologies in LZ simulations. BACCARAT is the cornerstone of both paths. The outputs of the simulation paths differ in the relative capability and detail within the full chain vs the fast chain.

3.1 Full Chain Simulation

The full chain simulation was developed alongside the commissioning of the detector to produce mock data to investigate the expected behaviour of LZ. This relied on modelling all detector effects known from the LUX experiment [51–54] as well as other dual-phase xenon TPCs. Once the LZ science run had commenced, discrepancies were discovered between the real and simulated data. Figure 4.2 demonstrates this discrepancy for a sample of single electron (SE) pulses: individual electrons extracted from the liquid surface.

Confronted with these discrepancies, the optimal tuning method became to overwrite the modelling of detector effects, such as drift speeds and photon yields, with fixed values or distributions measured from calibration data for that detector configuration. Naturally, this provided more accurate comparisons to the distributions observed in data. However, this also introduced many hard-coded values specific to that detector configuration and not necessarily physically motivated; therefore, any changes to the detector conditions require new measurements and the retuning of all values. This also removes the predictive power of the simulation to changes in the detector. For this rea-

son, there is strong motivation to move back from hard-coded distributions to physical models for each effect.

In this section, we will discuss the general setup of BACCARAT and explain the capabilities of the software which are utilised by the full chain simulation. We will also point out the positions where updates were made throughout this work; these are detailed in Chapters 4 and 5. Note that we omit the simulation of the LZ skin and outer detectors and instead focus entirely on the TPC signals.

3.1.1 BACCARAT

The cornerstone of all simulation pathways in LZ is BACCARAT. This software package uses the GEANT4 libraries and functionality to provide a basis for compiling additional physical models that are specifically relevant to TPC technologies, or to LZ alone in some cases.

3.1.1.1 GEANT4

The BACCARAT framework is a wrapper which makes the GEANT4 framework more easily accessible for LZ's needs. BACCARAT uses GEANT4 to implement the following:

- Detector geometry
- Detector materials
- Primary generation of events
- Propagation and tracking of particles
- Physics processes governing particle interactions

3.1.1.2 Physics List

GEANT4 requires a physics list which provides all of the fundamental information regarding the underlying physics of the simulation. This includes information such as the branching ratios, decay times and cross-sections of interactions, to name a few. A custom physics list is compiled for LZ's needs, which is defined in [50].

3.1.1.3 Event Process

To follow the timeline of a single event simulated within BACCARAT, we start in Figure 3.2 with the macro. This file contains all the relevant settings for this particular run of the simulation (e.g. particle source, particle energy). It is preferable here that any variables that are expected to change run by run should be given macro commands that can be edited per run of the simulation, as opposed to some hard-coded physics values that should be set by default in the simulation code. For example, one might specify the particle source in the macro but would expect the simulation to automatically know the cross-sections for this particle's interactions, which are stored in the pre-defined physics list.

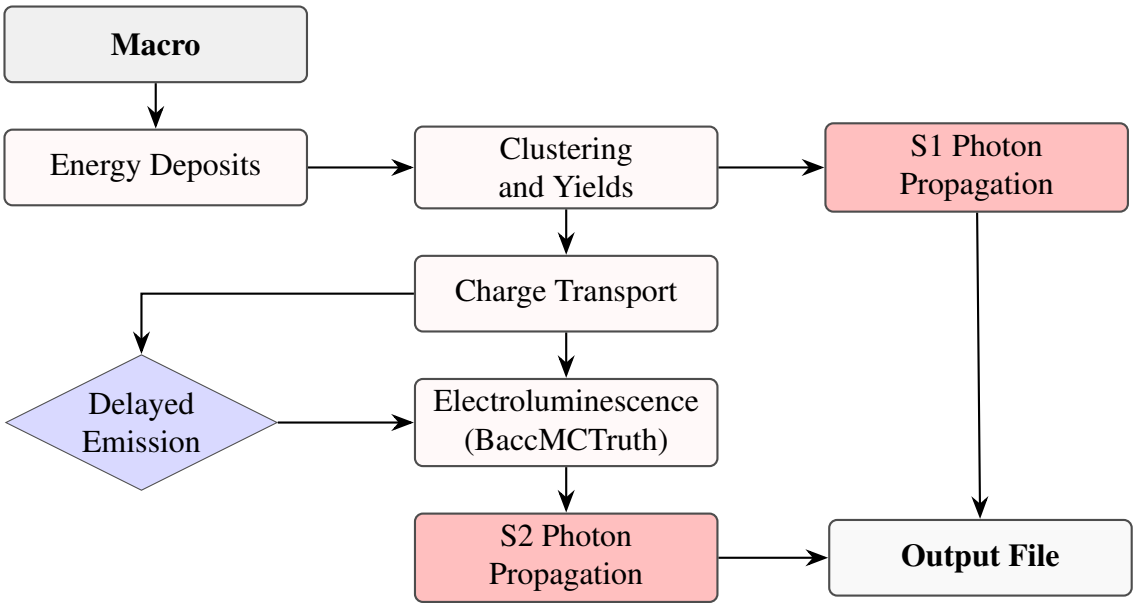


Figure 3.2: Simplified flow chart describing the stages of processing through BACCARAT.

The macro commands are processed and provide GEANT4 with the instructions to propagate the primary particles into the detector geometry and simulate all interactions that occur in each event. From each of these interactions, the positions, times and energy depositions are recorded. The amount of energy deposited in each interaction determines the light and charge yields (the number of photons and electrons emitted) that then contribute to the S1 and S2 signals. However, since the spatial resolution of the LZ detector, $\mathcal{O}(\mathrm{mm})$, is much larger than that of the GEANT4 stepping action, which has no effective lower limit, we can simplify the calculation of the yields by clustering energy depositions that are congregated in a single site. The criteria of this clustering are detailed in Section 3.1.1.4. Clustering in this way simplifies the calculation of the light and charge

yield since these are measured in experiment from the total energy deposited as opposed to each deposition.

3.1.1.4 Clustering

There are two clustering algorithms available in BACCARAT, clustering by lineage or by density. Clustering by density uses the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm, which identifies clusters in an N-dimensional space by ensuring all points in a cluster are within a fixed radius from at least one other point in the cluster [55, 56]. The cluster radius is physically motivated by a combination of the particle track structure and detector resolution. This algorithm is the default for LZLAMA but has been shown to depend on the energy of the depositions to accurately cluster interactions. Most notably, it is difficult to define the correct radius to scan over for clusters of varying densities at different energy scales, where the higher-energy particles produce longer track lengths and particle showers.

Alternatively, the default in the full chain is to cluster events by lineage; this is where all energy depositions that originate from the same particle are clustered together irrespective of density. New lineages are created when

- a track is a primary particle,
- a gamma photon has travelled farther than $9\,\mathrm{mm}$ (motivated by the mean gamma path length),
- or a radioactive decay occurs, giving a daughter particle of a different type.

Clustering correctly is important for two reasons. Firstly, and most importantly, the clustering defines what is observed in simulation as a single or multiple scatter. As previously described, WIMP-nucleon interactions will be SS events; hence, all MS events are immediately excluded from our analysis. If energy deposits are clustered too finely, the simulation will generate more MS events than expected; conversely, clustering too coarsely will misclassify real MS events as SS, corrupting the signal model. Secondly, the amount of energy accumulated in each cluster directly affects the light and charge yields of the interaction, and hence changing the clustering algorithm directly impacts the sizes of the S1s and S2s simulated. For this work, we focus mostly on the low-energy regime, which has been well calibrated and validated across multiple experiments in the field such that no changes were required to produce accurate events. However, at high energies,

evidence suggests that both the clustering algorithm and the yield models need to be tuned to replicate the distributions observed in LZ.

3.1.1.5 Particle Propagation

Once the clusters are determined, the yields are calculated using the Noble Element Simulation Technique (NEST) software, described briefly in Section 3.2. At this stage, the processing of the S1 light and S2 electrons can be disentangled. For the S1s, the photons are propagated from the event position around the LZ detector using GEANT4's stepping function until they reach a PMT face. The optical parameters for this light propagation had been tuned prior to the start of this work to match the S1 light collection efficiency [57]. For high-energy events, photon propagation becomes the bottleneck of speed. To avoid a similar bottleneck when transporting electrons, maps and parametrised functions are used. These include drift times, diffusion, electron lifetime, and wall-attachment effects. The drift times of the electrons are sampled from drift maps which are produced using the LZ Electric Fields (LZEF) package described in Section 3.3. Subsequently, all other timing or yield effects are applied through models which have been derived or tuned as part of this work and will be detailed in Chapter 5.

It can also be seen in Figure 3.2 that, as part of the charge transport stage, there is an option for delayed emission of electrons. This is a simplified term for a multitude of effects, most of which exist in the simulation but are in the process of being tuned. These include effects like electron trains: following a significantly large deposition in the detector, there is a flurry of electrons which reach the surface and produce a spree of small S2s that reduce acceptance for subsequent events. These electron trains are well modelled and have designated, highly efficient cuts to remove such events; hence, simulating these becomes a lower priority for the collaboration. Alternatively, it can refer to photoionisation, in which photons may strike some element of the inner TPC and produce subsequent electrons some time after a pulse. While photoionisation events can occur at any energy regime, they are less likely to occur at low energies since fewer photons are produced. They do, however, become significant for higher-energy searches such as the neutrinoless double beta decay ($0\nu\beta\beta$) and as such should be simulated well but will not be discussed in detail in this thesis.

Once the electrons reach the liquid surface, they are killed in BACCARAT and all effects happening in the gas are performed using a Monte Carlo simulation script which we call BaccMCTruth.

3.1.1.6 BaccMCTruth

The BaccMCTruth script handles the output stage of the simulation, taking all the truth information and packaging it up into a single ROOT file to be passed forward to the DER or analysed directly. It also provides the final steps of the simulation of the S2 pulse, electroluminescence. There is a multitude of physical effects occurring in the gaseous xenon for these electrons, which are discussed in detail in Chapter 4. For now, we will briefly describe the order of operations that result in a single output file ready for the DER processing.

Firstly, the electric field in the extraction region must be well modelled, as it governs all physical effects there. To do this, we account for the two dielectric materials (liquid and gaseous xenon) and the effects of electrostatic attraction between the two mesh grids at high voltage, which alters the spacing between the grids in some areas, thereby increasing the field strength. Once the field strength is determined, we can establish whether the electrons are extracted into the gas or are trapped at the liquid surface. If they are extracted, we then determine their drift time through the gas to the anode and also the electroluminescence light yield along the track. During the electroluminescence process, the excited xenon atoms occupy two different states: singlet and triplet, with different decay time constants. The difference in photon emission time is accounted for during this process. Propagating these photons individually, in a similar fashion to the S1 photons, would be very computationally expensive; instead, we opt to use another map which contains the probabilities of photons reaching PMTs from each point in space from the gaseous xenon. This map is produced using one large-scale optical simulation from which the probabilities are extracted. It therefore relies on the material parameters present in the geometry which are discussed further in Section 4.8. This improves computation speed and saves significant computational resources. Finally, the BACCARAT stage ends with the output ROOT file containing all of the truth information from the event in addition to the information of which photons hit which PMTs and at what time. This is everything needed to process the signal through the electronics chain with the DER.

3.1.2 Detector Electronics Response (DER)

With the photon hit patterns and times determined, the DER process can begin. This software package is designed to be entirely modular, with each module taking input from the previous one, mirroring the real LZ data acquisition system, FADR [58].

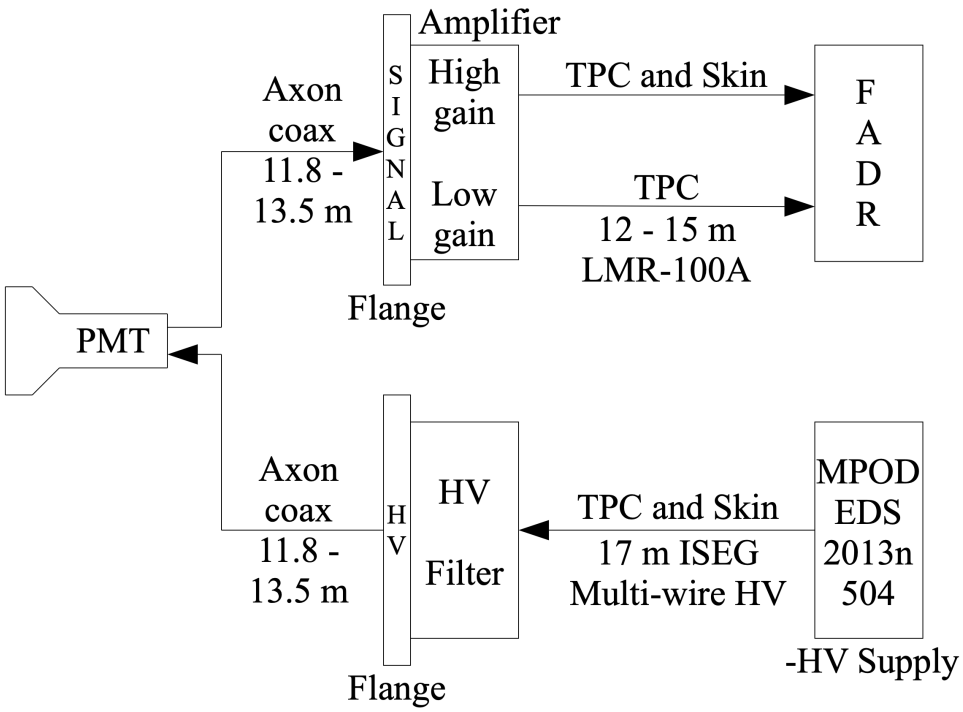


Figure 3.3: Simplified schematic of the LZ data acquisition pipeline from the PMT through the amplifier and through to FADR. The FADR process includes digitisation of the voltage signal, data sparsification, triggering, and event building [58]. The DER only needs to map the top chain from this figure, which is mirrored in the simple flow diagram in Figure 3.4.

Figure 3.3 shows the general data acquisition chain for a signal from a single PMT and the flow path for it to be stored as an event. The DER methodology maps this same pathway and models each step, altering the observed signal and ultimately producing a file identical to the raw data files to be processed by LZAP in the future. The further-simplified schematic of the DER flow is shown in Figure 3.4.

3.1.2.1 SPE Modelling

Firstly, we introduce how the initial response is simulated for a single photon striking the PMT window, following the modular process adopted by the DER. Recall that the BACCARAT output file conveys that a given PMT has been hit by a photon at the photocathode; hence, we must determine whether a photoelectron is released from the photocathode into the PMT. This is done using the values for the quantum efficiency (QE) of all of the Hamamatsu R11410-22 PMT photocathodes as quoted by Hamamatsu [59]. However, the values quoted by Hamamatsu are for room-temperature operating conditions, whereas LZ PMTs operate at cryogenic temperatures of $\sim$175 K [40]. For

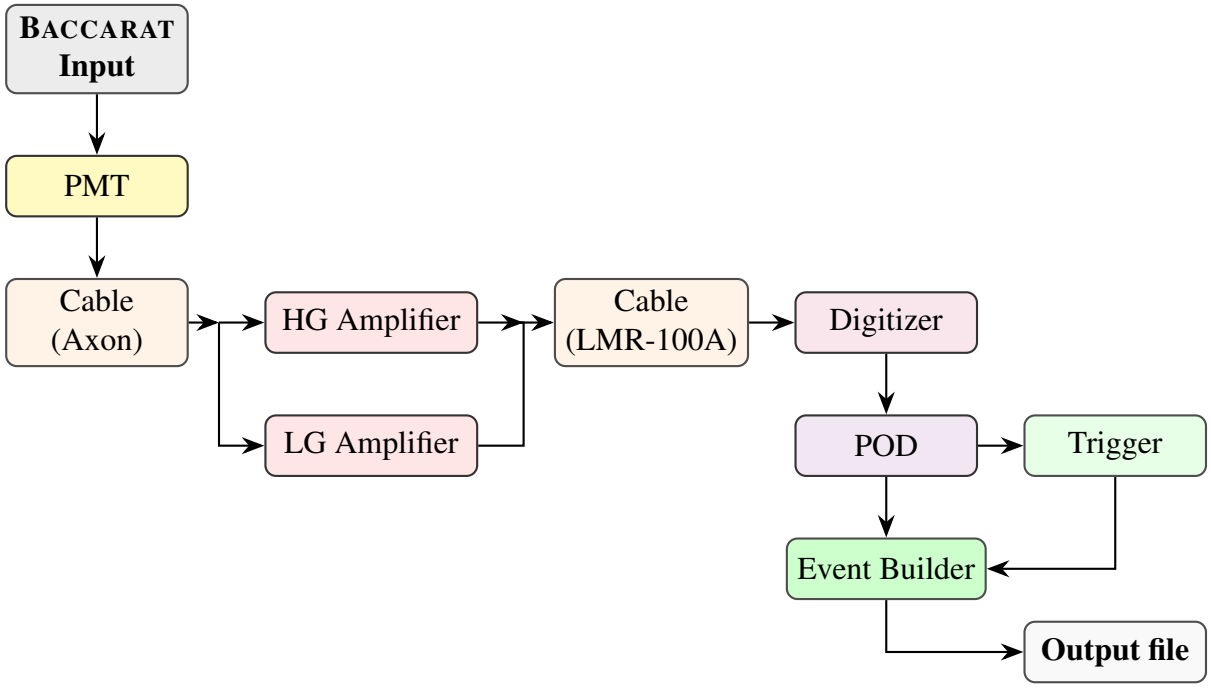


Figure 3.4: Schematic showing the general steps taken through the DER process from initial PMT hits from BACCARAT through to the event output file. This output file is then ready to be passed into LZAP.

this reason, calibration curves are used to translate between warm QE measurements and cold QE measurements by $QE_{cold} = 0.73QE_{warm} + 0.07$ [59]. The QE for every PMT in our detector is shown in Figure 3.5, which averages $\sim 30\%$. Hence, a simple Monte Carlo can determine whether this photon produces a photoelectron or not.

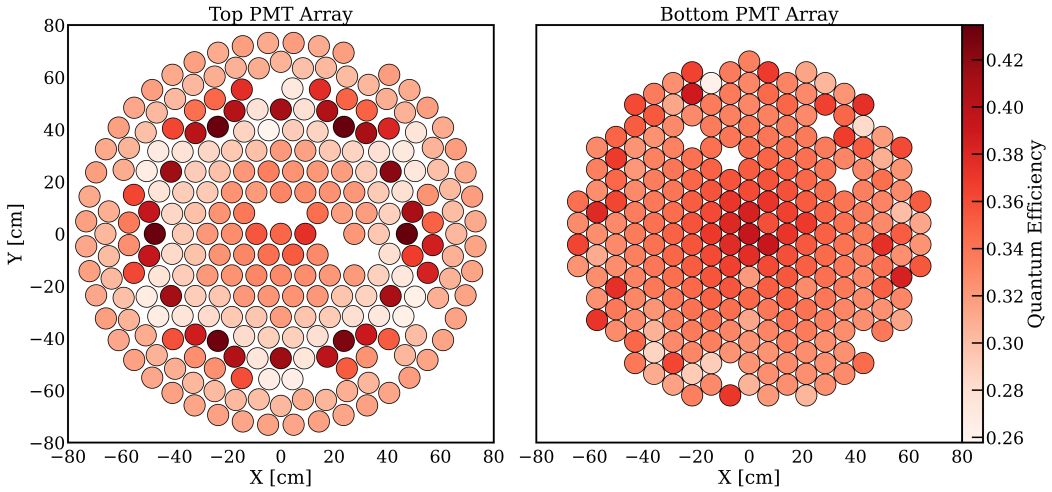


Figure 3.5: Heat map displaying the quantum efficiency of all PMTs in the top and bottom array of the LZ TPC. There are visible gaps in these arrays; they indicate where PMTs have been disabled due to high afterpulsing rates, which would otherwise degrade data quality. Both arrays follow the colour bar on the right side of the plot.

The photoelectrons released from the photocathode are then accelerated toward the dynodes, where collisions between the incident photoelectrons and the dynodes cause more electrons to be released into the electron avalanche. This avalanche eventually reaches the anode and produces a detectable signal for a single photon. The number of photoelectrons produced in the PMT is determined by measuring the gain of the PMT, which is strongly dependent on the voltage, as well as other secondary effects which are detailed in Section 3.1.2.2. For simplicity, the base PMT pulse shape is modelled following the schematic diagram in Figure 3.6 to reach the following equation for the single photon response,

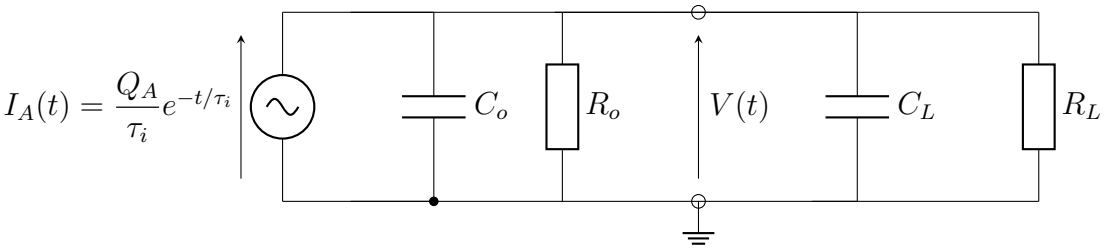


Figure 3.6: Circuit diagram for the output voltage from the R11410-22 Hamamatsu PMT base. The expression for $V(t)$ is given in Equation 3.1.

$$V(t) = \frac{Q_A R}{\tau_f - \tau_r} \left(e^{-\frac{t}{\tau_f}} - e^{-\frac{t}{\tau_r}} \right) \tag{3.1}$$

where Q_A is the total charge of all electrons reaching the anode, R is the total resistance of the circuit in Figure 3.6 and τ_r and τ_f are the rise and fall times of the pulse, respectively. We note here that the value of Q_A is linearly dependent on the PMT gain, G, which is a measurable quantity in our data from the monthly LED calibrations. We also note that, for simplicity in simulation, we assume that τ_r and τ_f are identical for all PMTs and are fixed to $\tau_f = 4.76\,\text{ns}$ and $\tau_r = 3.81\,\text{ns}$, as determined by a test stand setup prior to LZ running [59].

3.1.2.2 Secondary PMT Effects

Numerous effects may occur within a PMT that can affect the signal observed directly from it. In LZ, we identify seven secondary PMT effects, detailed below.

Double Photoelectron Effect (DPE)

When a VUV photon from liquid xenon (LXe) scintillation is incident on the PMT photocathode, there is a possibility that the photocathode can release two photoelectrons from its surface instead

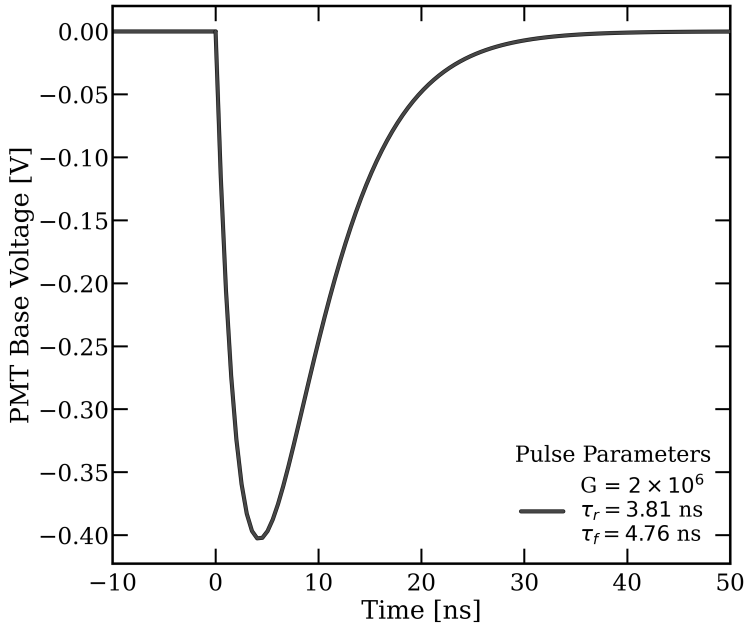


Figure 3.7: Voltage output from the PMT base for a single photon incident on the PMT face. The nominal values of all parameters used in the simulation are provided for this pulse, showing the relevant timing scales for the specified rise and fall times. The voltage scale shown motivates the necessity for amplification in the signal chain.

of just one. This leads to a notable increase in the size of the single photon pulse. For this reason, it is necessary to measure the probability of DPE occurring in each PMT. On average, the LZ TPC PMTs have a 21% probability of producing DPE pulses. This value comes from the quoted measurements from Hamamatsu, which are then calibrated to cold conditions, similarly to QE, using the formula $DPE_{cold} = 0.95 DPE_{warm} + 0.033$ from [59]. This information is stored in the PMT parameter spreadsheet, along with the QE values for all PMTs. When simulating each photon hit on the PMT surface, it is simple to draw from a probability distribution using Monte Carlo techniques to add an additional photoelectron released from the PMT face.

First Dynode Hits

An alternative process that can occur in PMTs is one in which the incident photon passes through the photocathode and directly strikes the first dynode, causing a photoelectron to be released from the dynode surface rather than the photocathode. This reduces the overall gain of the pulse because the gain from accelerating the photoelectron to the first dynode is omitted, resulting in pulses smaller than expected. This effect was observed during testing and was fit in Figure 3.8 with a probability of 10%, which is applied across all PMTs in the simulation. This effect can be observed

during the same process used to measure the DPE fraction and appears as the peak closest to $0\,\mathrm{nV\,s}$ in Figure 3.8.

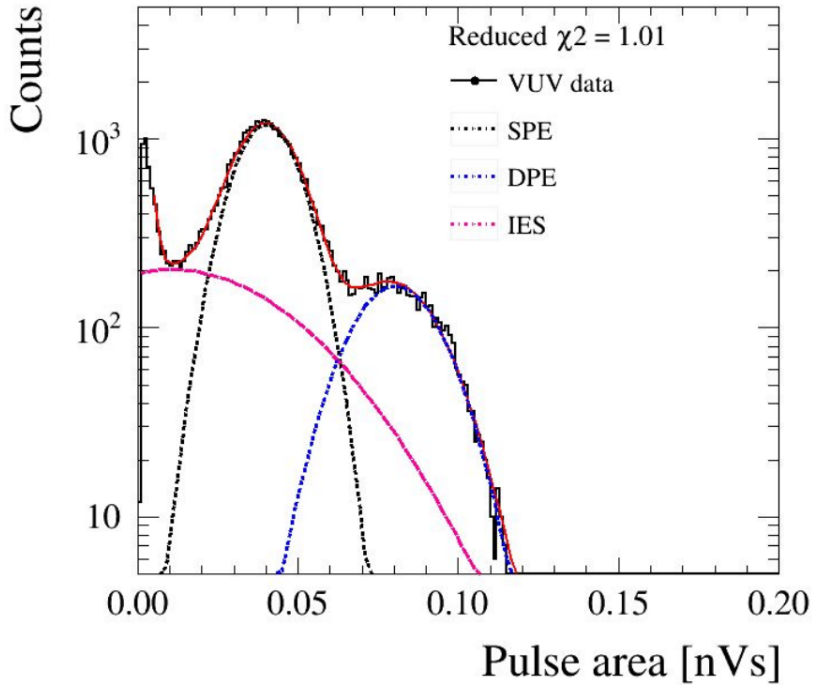


Figure 3.8: Calibration fits of a high gain R11410-22 PMT from the Imperial College London test stand [59]. This PMT was set to a gain of 3.5×10^6, the expected gain of the LZ PMTs before commissioning. The plot shows the separation of single-photoelectron (SPE), double-photoelectron (DPE), and inelastic-scatter (IES) pulses. The additional peak in pulse area at $0\,\mathrm{nV\,s}$ is a result of first dynode hits.

Second Dynode Collection

Another way in which some undersized pulses are produced is by second dynode collection: the photoelectron is released from the photocathode but is accelerated directly towards the second dynode, again bypassing the gain from the first dynode. This process is noted as a possibility but was not observed during commissioning with any statistical significance, and hence we choose to ignore it in simulations. This is to be expected since the likelihood of this occurring is significantly small in practice that it can be ignored.

Inelastic Scattering

We now take the case of inelastic scatters whereby the photoelectron backscatters off the first dynode and retains only a fraction of its initial energy. Thus, when it is accelerated either back into the first dynode or towards the second, it has less energy and will produce fewer secondary

electrons, leading to an undersized pulse as well as some transit time delay. This can also be seen in Figure 3.8 with a much wider profile than the other effects since the amount of energy lost in this process is dependent on how far from the dynode the photoelectron travels after scattering. This was characterised during commissioning to have a global average of 20% probability of occurring in each PMT and is applied as a flat probability in the DER.

Elastic Scattering

Alternatively, if the photoelectron scatters elastically from the first dynode, it retains all energy and will decelerate back towards the cathode, stopping just before it. Then the photoelectron will begin to accelerate back towards the first dynode and the process will continue as expected. This process does not affect the size of the pulses produced but will affect the transit time for the signal to pass through the PMT, similar to inelastic scattering. For this reason, the photoelectron transit times from the photocathode to the anode are saved in addition to those from the first and second dynodes. These are stored in the PMT parameter spreadsheet such that the timing distributions from these effects can be accurately simulated. Elastic scattering was estimated to have an average probability of 10%, which is applied to all PMTs.

Afterpulsing

Afterpulsing occurs when a gaseous impurity enters the PMT, usually due to PMT degradation over time. In this instance, as the electron avalanche travels through the PMT, it may ionise residual gas molecules, causing the ions to be accelerated towards the cathode and the ionised electrons to contribute to the avalanche. Often, as the ion drifts towards the cathode, it may strike dynodes or the photocathode, producing secondary electrons that generate a delayed pulse. The time delay and the size of the additional pulse depend on the ion's mass and charge. During commissioning, these values were calculated for the ion species we would expect to observe during LZ operation, as informed by LUX [51]. However, as part of the work outlined in Chapter 4, these values were measured and updated from LED calibration data to resolve discrepancies in the position reconstruction algorithm caused by an overestimation of these afterpulsing rates and sizes.

Dark Counts

For all PMTs, there is some probability for spurious emission of photoelectrons from the photocathode as a result of thermionic emission, even in the total absence of light. These electrons

appear as regular SPE pulses in the data and can become a background if they occur simultaneously in multiple PMTs. To accurately simulate these effects, we sample their production from a Poisson distribution determined by their rate. The dark rates are provided to LZ by Hamamatsu. Once sampled, they are treated by the simulation as regular photoelectrons, only with the inability to display the DPE effect since this is not possible for dark counts.

3.1.2.3 Data Acquisition

Throughout the data acquisition stages shown in Figure 3.4, four main transfer functions are used to model the filtering of the signal passing through each stage. These transfer functions are shown in Table 3.1. Using these filters we can model the stages of passing through cables, which are

Filter Type	Transfer Function: H(ω)
Low Pass (LP)	$\frac{V_{out}}{V_{in}} = \frac{1}{1+i\omega\tau}$
High Pass (HP)	$\frac{V_{out}}{V_{in}} = \frac{i\omega t}{1+i\omega\tau}$
Pole Zero (PZ)	$\frac{V_{out}}{V_{in}} = \frac{1+i\omega\tau_1}{1+i\omega\tau_2}$
Sallen-Key (SK)	$\frac{V_{out}}{V_{in}} = 2 \times LP$

Table 3.1: Transfer functions used by the DER in modelling steps through stages of the data acquisition flow. In these functions, ω refers to the cut-off frequency of the filter and τ refers to the RC time constant of the circuit schematic. These values will differ between each application stage in the DAQ pipeline.

individually simulated based on the cabling at each stage as shown in Figure 3.3, and the low gain ($G \times PZ \times PZ \times SK$) and high gain ($G \times PZ \times LP \times LP \times HP \times PZ \times PZ \times SK$) amplifiers. Since none of this was altered as part of this work, further details are omitted describing the stages up to the digitiser; instead, the effect of each stage on an initial SPE pulse is demonstrated in Figure 3.9.

In the digitiser, and subsequently during Pulse Only Digitisation (POD), the data is transformed into a state ready for analysis. For the digitiser, the input from amplifiers is now subsampled on a $10\,\mathrm{ns}$ basis as opposed to the $1\,\mathrm{ns}$ steps used prior. This is also where the pulse amplitude is digitised from volts to Analogue-to-Digital Converter (ADC) counts. The digitisers used in LZ are 14 bit ADCs with $2\,\mathrm{V}$ range giving the linear transformation between +1 V = +8191 ADCC

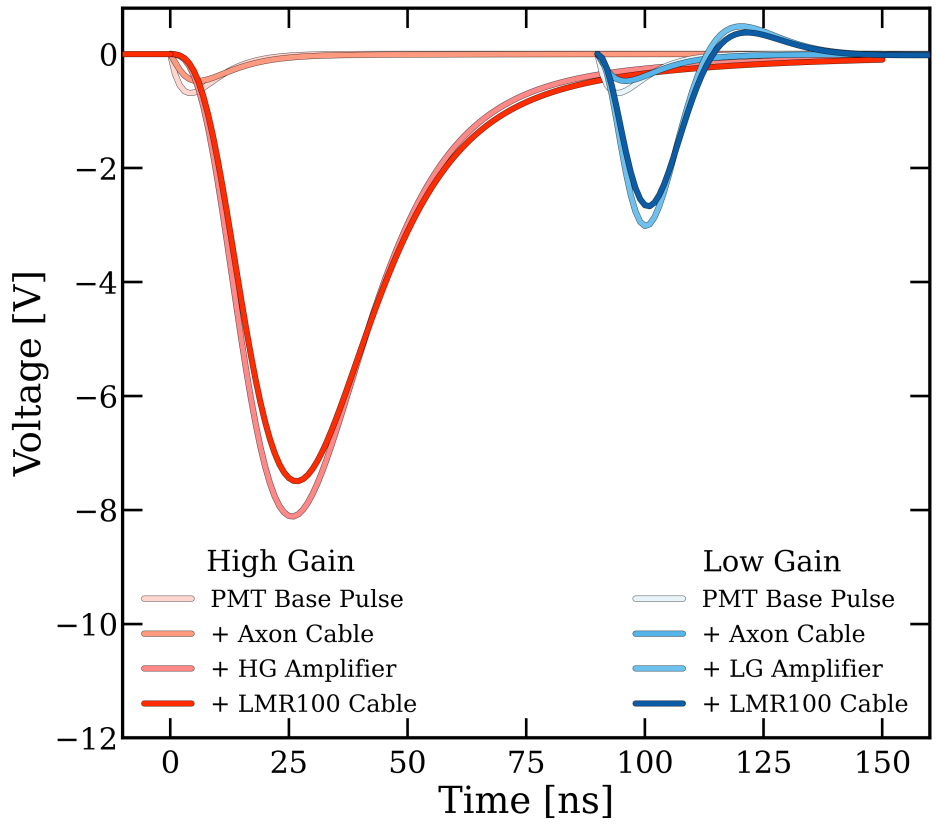


Figure 3.9: Alterations made by passing the base PMT pulse through each stage of the data acquisition chain. The full characteristics of each change are approximated by applying a series of transfer functions described in Table 3.1. Note that in this plot, the low-gain response has been shifted by $80\,\mathrm{ns}$ for readability.

and $-1\,\mathrm{V} = -8191\,\mathrm{ADCC}$. In simulation, we artificially place the baseline amplitude and baseline noise as a Gaussian distribution around the baseline. During this work, the baseline and noise were adjusted to match the average of the raw LZ data; however, because of the analysis steps in LZAP, this has no observable effect on the output and is performed solely for the accuracy of the simulation.

The digitised signal is then passed through the POD algorithm. This reduces data storage in LZ by a factor of 50 in practice and also in simulation [60]. This is achievable using a conventional trigger to record data only when the sample is above the POD threshold in ADCC. For example, a TPC high gain PMT channel has a $25\,\mathrm{ADCC}$ trigger threshold. If three consecutive samples in a given PMT are observed to be above threshold, the data is stored in a POD; this ends when there are three consecutive samples below threshold. While this reduces the amount of data saved from the simulation, we retain the option to access it if necessary for validation and tuning.

Finally, for an event to be recorded, we opt to trigger on S2 pulses since they are generally much larger than S1 pulses and hence can be more easily identified above the noise. Whenever at least 3 high gain PMT channels pass the S2 threshold of $13\,500$ ADCC (as opposed to the much smaller 25 ADCC for each POD trigger) the event is triggered upon and a time window 2.2 ms before and 2.5 ms after the trigger is recorded for the event. In simulation, whether the event is triggered is irrelevant to whether it is recorded. Nevertheless, this does provide a good test bed for implementing new trigger algorithms and measuring trigger efficiencies.

The DER output is recorded in a file identical to the data equivalent, with the addition of the `DetectorMCTruth` tree, which contains all relevant truth information from the simulated event, passed through from BACCARAT with the addition of some relevant effects from the DER (e.g. photons detected, dark counts).

3.1.3 LZAP

All data recorded by the LZ experiment is processed using LZAP. For brevity, the intricate details of how LZAP functions will not be discussed here. Instead, we state the general process by which LZAP processes event data and produces the reduced quantities (RQs) necessary for analysis.

Beginning from the raw data, the baseline of each PMT is measured by averaging the 32 samples before the trigger of a POD, which is then subtracted from all samples in the event window to reduce the baseline to a global zero. The Pulse Finder module then identifies each pulse in the event window and extracts its parameters. This involves measurements of total pulse areas and widths along with individual PMT channel areas and widths, to name a few. Pulse widths are defined using Area Fraction Times (AFTs), with time stamps recorded during integration to indicate the times at which predetermined percentiles of the total area coverage occur. Widths can then be extracted by taking differences between percentile positions. An example is depicted in Figure 3.10. Then the Pulse Classifier module is applied to each pulse, using its parameters to classify them into categories: S1, S2, single photoelectron (SPE), multiple photoelectron (MPE), single electron (SE) or Other [61]. Since an S2 pulse is the amalgamation of multiple electrons in time, S2s are separated from SEs based on the pulse area. This distinction is useful for monitoring detector stability in the extraction region. The 'Other' classification of pulses consists of any pulse which does not confidently fit into any of the aforementioned categories.

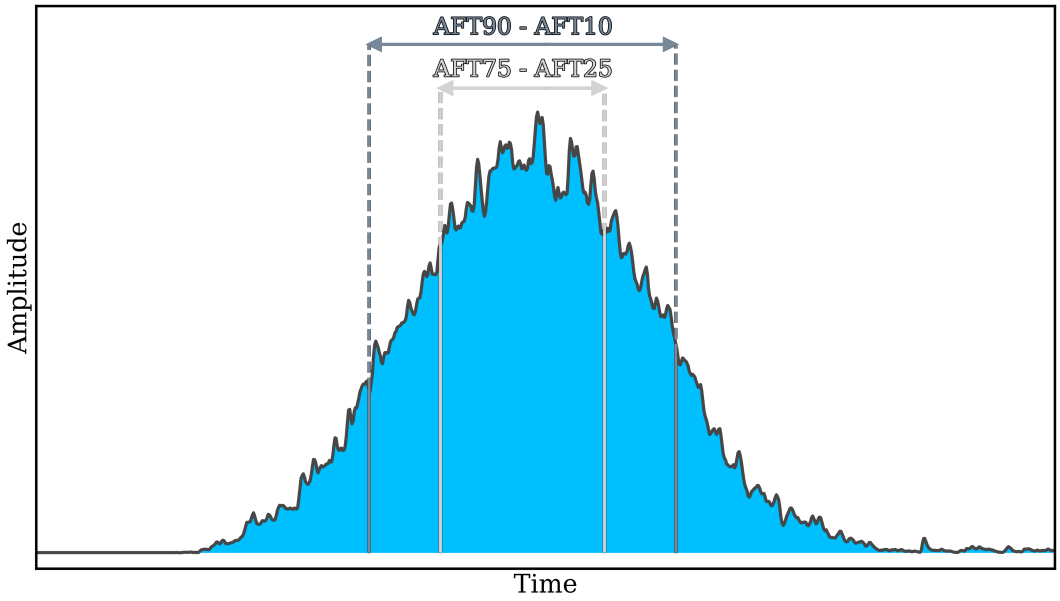


Figure 3.10: An example S2 waveform showing the integrated area (blue) used to define the S2 pulse area. Two common width metrics using AFT parameters are shown. An identical process is performed for all other pulse types.

Once all pulses in the event window have been identified, the Interaction Finder identifies event topologies; specifically, we classify events as SS or MS. Any events classified as SS are potential signal events. For this reason, they must be selected cleanly and be comparable to other events within the detector by applying geometric corrections, as discussed in Chapter 2. Along with the geometry dependences, varying detector conditions can alter the size and shapes of these signals. To account for this, we read from the conditions database, which stores the detector environment variables (e.g. temperature and pressure) for each run. This is done by reading the timestamp of the raw data file and locating the detector configuration in the database. This poses a problem for the simulation, which does not allow time-dependent changes. To overcome this, we apply a custom timestamp to the raw file in the DER that corresponds to the period of detector stability we wish to replicate.

By default, waveforms are not saved in all LZ data. However, it is an option within LZAP to save either the waveform summed over all the channels or the individual channel waveforms to the output file. This increases the storage size of the data; however, the waveforms encode much more of the microphysics occurring than just the RQs, which can be utilised in many machine learning algorithms. Three examples of which are detailed in Chapter 7.

3.2 Noble Element Simulation Technique (NEST)

The NEST software is a stand-alone package widely used by the dark matter community to provide empirical functions that map observable quantities to the initial energy depositions in a dual-phase TPC [62]. These empirical functions are generated by fitting general functional forms to measurements taken by experiments within the field. With new measurements constantly being published, these models inherently become more well-defined. However, these NEST models are global averages of a convolution of difficult-to-separate effects; hence, it is necessary to re-evaluate the model's applicability for each experiment. Additionally, since not all regimes of parameter space have been probed by experiments yet (for example, TPCs with $< 20\,\mathrm{V\,cm^{-1}}$ drift fields), NEST extrapolates the models, which in some cases proves ineffective. LZ uses NEST to provide many models within the full chain simulation, all of which are validated during the process of updating simulations in this thesis. Most notably:

- Light and charge yields
- Gaseous xenon drift speeds
- Extraction efficiency

3.3 LZ Electric Field (LZEF)

For electric field simulations, LZ uses LZEF to produce 2D (r, z) axisymmetric electrostatic maps of the TPC volume. These maps are necessary for simulating inhomogeneities within the detector that may arise from charge transport properties on different dielectric surfaces, such as the PTFE and grid wires, or from field leakage due to the transparency of the gate ($\sim 97\%$) or cathode ($\sim 92\%$) [63]. Simulating these fields is preferable to assuming a perfectly uniform field, since the field directly affects S1 and S2 properties such as light yield, electron drift speeds, and diffusion. Local deformities are corrected for during LZAP processing; hence, it is necessary to simulate them whenever possible. Many improvements were made to this software by Sparshita Dey in parallel with the tuning of the full chain; a complete description can be found in [63].

3.3.1 Field Maps

To produce electrostatic field maps, LZEF takes a Finite Element Analysis approach that uses meshing to simulate the necessary fields in each mesh element and then combines them to produce a total field map [64]. The resulting map contains the following information for each position in the liquid xenon bulk of the detector:

- Radius, R [cm]
- Height from cathode, Z [cm]
- Radial Electric Field Strength, E_R [V cm^{-1}]
- Z-component of Electric Field Strength, E_Z [V cm^{-1}]

This map is then read in by BACCARAT and sampled when performing any calculations involving the bulk electric field strength, which is critical for accurate light and charge yields.

3.3.2 Drift Maps

Similarly to the electric field maps, BACCARAT also utilises drift maps to speed up calculations and provide more accuracy than BACCARAT itself is capable of. Drift maps are constructed by sampling the individual electron trajectories from solutions to the field maps, thereby solving for the electron streamlines. These maps contain the following relevant information:

- Radius, R [mm]
- Height from cathode, Z [mm]
- Drift Time [µs]
- Drift Time Variance, σ_{dt} [µs]
- Wall Attachment Probability

Note that, along with the drift time of the electrons, these maps store information regarding the wall attachment probability and diffusion of the electrons. The attachment probability refers to the probability that an electron drifts towards the PTFE walls of the detector and is captured by the dielectric surface. LZ has observed an incremental build-up of charge on the PTFE walls over time,

making this an important addition to the electric field modelling since it directly influences the fiducial volume designation. Additionally, the diffusion of electrons drifting to the liquid surface is included in the Drift Time Variance, σ_{dt}. However, this is determined by the statistical variation in drift times, which is found to be inaccurate. The corrections made to this are discussed thoroughly in Chapter 5.

4 | Simulating Single Electrons

Accurately simulating S2 pulses requires a thorough understanding of the physical processes occurring throughout the production chain of an S2 waveform, from initial energy deposit through to signal digitisation. The full chain simulation is usefully modular, allowing for each stage to be individually investigated and improved upon. However, due to the large number of parameters, there is a large amount of degeneracy between various stages along the simulation chain. For example, when tuning the S2 pulse area to calibration data, a decrease in S2 area can be achieved by adjusting charge yields, electron lifetime, extraction efficiency, reflectivity parameters, PMT quantum efficiencies, PMT gains, or DPE fractions. Hence, the process of improving the simulation becomes a cyclical one of implementing corrections to converge on a final model that most accurately agrees with all RQ distributions across all energy regimes and the relevant field measurements for each model.

To reduce the degeneracy of the modelling, a tuning methodology was devised to first ensure that SE pulses are well simulated before handling a full S2. Since S2s are comprised of multiple single electron pulses in time, ensuring an accurately simulated SE waveform is crucial for producing accurate S2 waveforms. Additionally, this removes all physical effects below and at the liquid surface, focusing entirely on the gaseous electroluminescence region, which is conveniently modelled in the stand-alone BaccMCTruth script.

The modelling of the physical effects in the gaseous xenon can be reasonably separated into quanta and timing effects within BaccMCTruth. For accurate pulse widths, the timing effects are dominated by the electron drift speed, scintillation timing, photon propagation times, and the distribution of photon starting points along the electron track. For accurate pulse areas, the quanta effects are dominated by the electroluminescence yield and light collection efficiency. The electric field model in the gas governs the majority of these effects, all except the scintillation timing and light collection efficiency, in fact. Hence, the electric field model in the BaccMCTruth script is the most important model to ensure accuracy. External to BaccMCTruth, the PMT response must also be accurate since this contributes to both the pulse widths and pulse areas, though the effects are much stronger for the pulse areas, as is shown in Section 4.6.

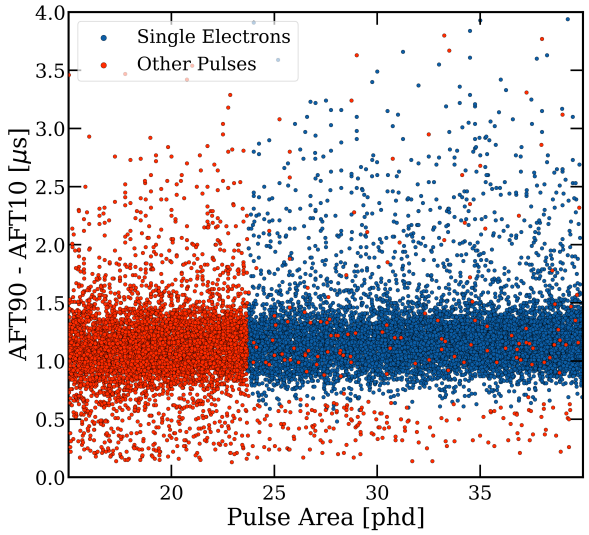
4.1 Single Electron Data Selection

To effectively tune the simulation output to agree with data, we must first establish the set of pulses we are attempting to replicate. LZAP automatically labels SE pulses based on their S2-likeness and pulse area. The mean SE size for the WS2022 detector configuration is $50\,\mathrm{phd}$ and the WS2024 detector configuration is $44.5\,\mathrm{phd}$. For context, the mean AFT75 - AFT25 width of the SE pulses is $706\,\mathrm{ns}$ in WS2022 and $758\,\mathrm{ns}$ in WS2024, though no width information is used for data selection. Unless stated otherwise, we adopt the WS2024 mean values for this thesis. Initial data selection methods rely on scanning over some stable time span of data for all SE pulses and use this to formulate the relevant RQ distributions. The problem with this is that LZAP is not 100% accurate in identifying SE pulses, and as such we find that we attempt to tune to distributions containing some contaminants instead of pure SE pulses. This can be seen in the pulse area distribution of the SEs, which we expect to be approximately normally distributed yet has an apparent shoulder at the lower area threshold, as shown in Figure 4.1c. The selection criteria of SEs can be improved by including information from other RQs. We use the understanding that the SE signals are produced at the top of the detector to ensure all pulses are physically motivated in top-bottom asymmetry (TBA) space. TBA is defined as

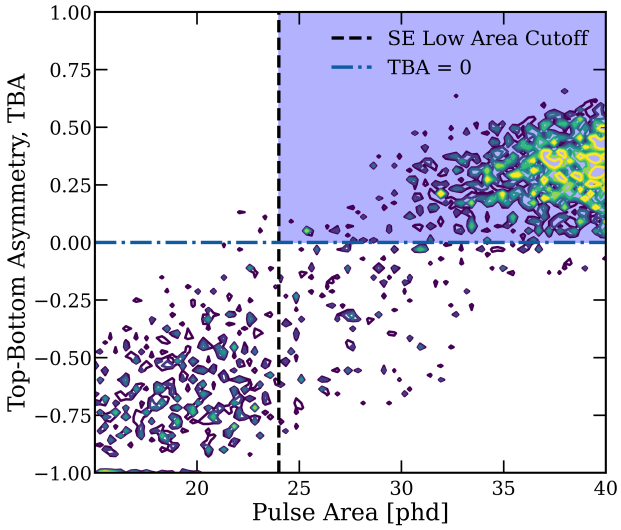
$$TBA = \frac{A_{top} - A_{bottom}}{A_{top} + A_{bottom}}, \tag{4.1}$$

where A_{top} and A_{bottom} are the pulse area observed by only the top or bottom array of PMTs, respectively. By definition, $-1 \leq TBA \leq 1$. Positive values indicate that the light source is closer to the top array than the bottom array, and negative values indicate the inverse. All light from SEs is emitted very close to the top array, so we expect TBA > 0 for SE pulses. While it is true that some statistical fluctuations may cause SE pulses to fall below this limit, enforcing this limit increases the purity of SE selection by 17%.

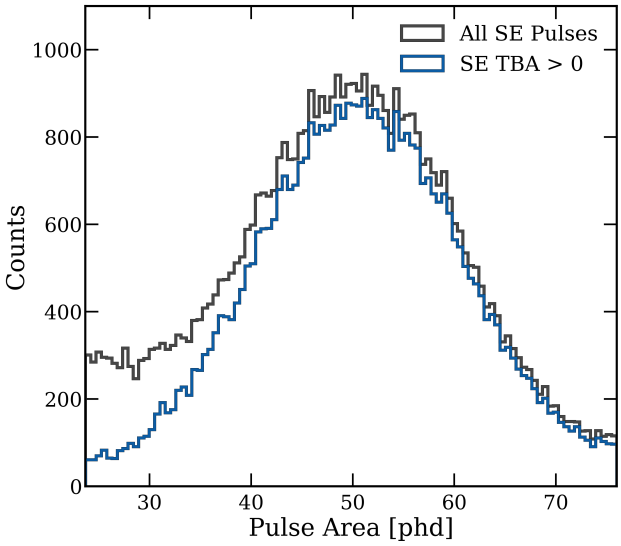
From inspecting the pulses with a pulse area of $10\,\mathrm{phd}$ to $40\,\mathrm{phd}$, we can see in Figure 4.1a the continuum between pulses classified as Other pulses and SE pulses, with a hard cut-off at $24\,\mathrm{phd}$. In TBA space, shown in Figure 4.1b, the transition between SE and Other is more easily resolvable, and hence we show that a cut in TBA space is a reasonable method to separate such pulses and increase the purity of the SE selection. This extra classification step for SEs will eventually



(a) Continuous transition between Other pulses and SE pulses seen in pulse width vs pulse area space. This shows that there is some leakage of Other pulses that, in this parameter space, resemble the SE pulses.



(b) The TBA space shows the separation between SE and Other pulses and demonstrates the cut placed in this parameter space to improve selection of pure SEs shown in Figure 4.1c.

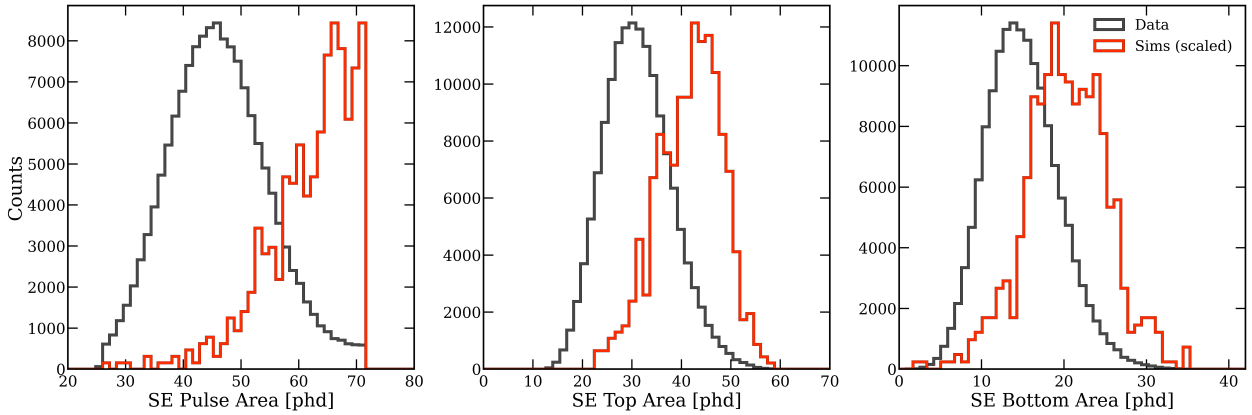


(c) Impact of applying TBA > 0 cut on the SE pulse area distribution. This removes the lower shoulder contamination from the Other pulses and provides a cleaner SE distribution to tune to.

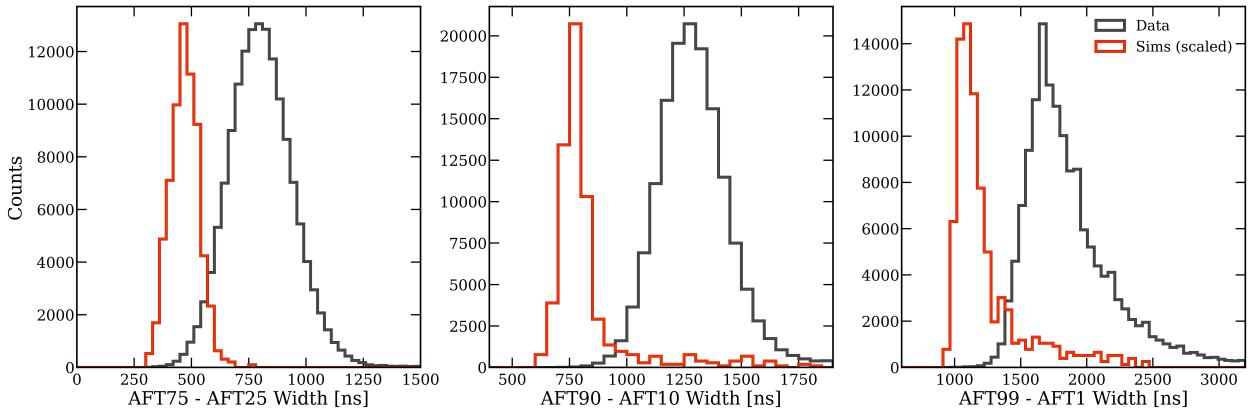
Figure 4.1: Separating the SE pulse selection from the Other pulses by applying a TBA cut. The distribution observed here is from the WS2022 science run and hence differs slightly from the WS2024 results shown in this chapter due to the higher electric field.

propagate into future versions of LZAP to improve pulse classification.

Comparing the SE pulse areas and widths to a simulated dataset reveals the magnitude of the discrepancies mentioned in the previous chapter, shown in Figure 4.2. This motivates the necessity to improve the simulation of SE pulses.



(a) Comparison of SE pulse area distributions in simulation and data for total (left), top (middle) and bottom (right) PMT arrays. The simulation overestimates the pulse area of SEs in both PMT arrays. The sharp truncation on the simulation in the left plot is caused by a hard cut on the possible SE size (72 phd in LZAP).



(b) Comparison of three different width metrics for SE pulses from simulation and data. Multiple width metrics must be validated to ensure waveform shapes are well simulated and include relevant tail effects.

Figure 4.2: Original (BACCARAT-V6.3.0, 2022) state of the simulation compared to a real selection of SE pulses in data. This comparison motivates the need for this tuning; discrepancies in SEs are amplified in S2s.

4.2 Electric Field Modelling

The physical processes occurring in the extraction region are strongly field-dependent. This includes the initial extraction of electrons from the liquid surface into the gas, accelerating the electrons to an energy above the electroluminescence threshold in gaseous xenon, and determining the electrons' drift speed through the gas, which in turn determines the widths of the SE and S2 pulses.

The threshold for electroluminescence (EL) in gaseous xenon is reported to be $\sim 1\,\mathrm{kV\,cm^{-1}\,bar^{-1}}$ at 293 K [65, 66]. LZ operates at $\sim$175 K with a slightly increasing temperature gradient towards the top of the detector, which is held at 178 K; hence, this EL threshold needs to be scaled to the

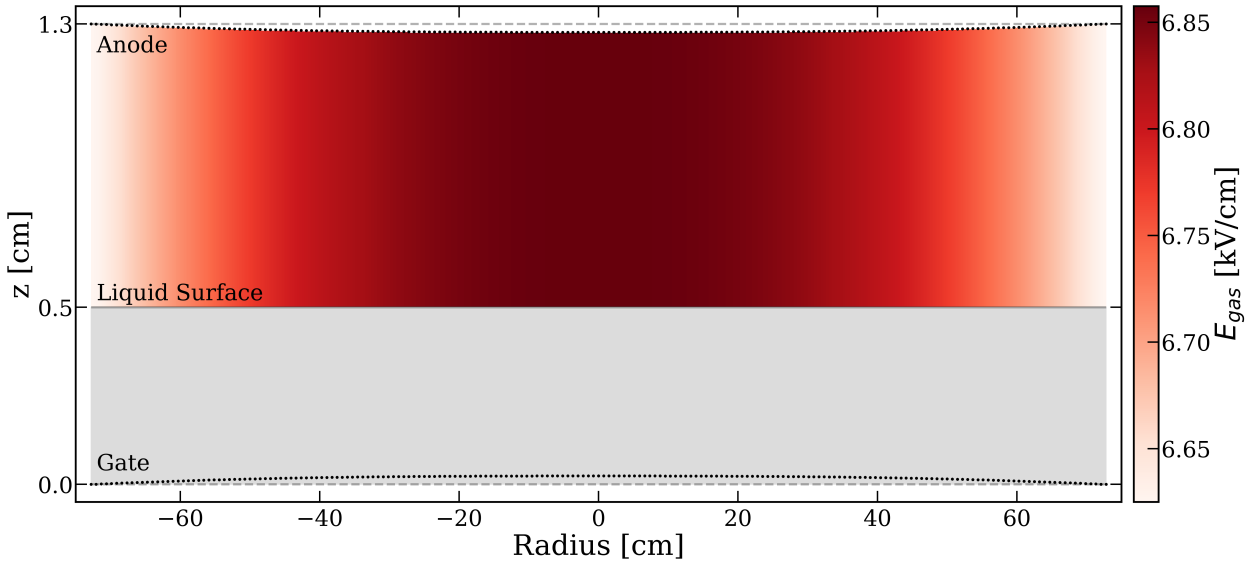


Figure 4.3: Simplified schematic of the LZ extraction region showing the scale of the liquid-gas xenon interface as well as the gate and anode grids and their relative deflection. The figure also shows the electric field as modelled in the gas region for the WS2024 detector configuration, showing the radial dependence induced by the small deflection of the grid wires.

correct temperature. EL yield is indirectly linearly dependent on gas temperature through the number density of xenon atoms. Therefore, at $178\,\mathrm{K}$ we find the conversion factor of $293/178 = 1.65$. This means LZ requires fields upward of $1\ (kV\ cm^{-1}\ bar^{-1}) \times 1.86\ (bar) \times 1.65 = 3.1\,\mathrm{kV\,cm^{-1}}$ to produce any EL light in the detector. The LZ electroluminescence region operates at field strengths upwards of $6.6\,\mathrm{kV\,cm^{-1}}$ as can be observed in Figure 4.3; above the EL threshold at this temperature and pressure.

4.2.1 Grid Deflection

Woven stainless steel wire grids generate the electric field in the extraction region. The extraction region is contained between the gate grid, $5\,\mathrm{mm}$ below the liquid surface, and the anode grid, $8\,\mathrm{mm}$ above the liquid surface. The voltage applied to each of these grids varies with the detector configuration, but for the WS2024 science run, the gate was biased to -$3.5\,\mathrm{kV}$ and the anode to +$4\,\mathrm{kV}$. Each wire in the grid is tensioned at the edge of the detector and, as a woven mesh structure, both grids are more resilient to the electrostatic forces of attraction between the two grids. Nevertheless, some amount of grid deflection is observed. The forces of attraction between the grids pull them towards each other, reducing the distance between them and strengthening the electric field. This introduces a radial dependence in the field, whereby only the outer edges of

the detector remain with a nominal $13\,\mathrm{mm}$ between the gate and anode, since these are fixed at the detector edges. In the centre of the detector, the grids are deflected towards each other due to the increased distance from the fastening points at the edges of the detector. The effect of this on the electric field is depicted in Figure 4.3.

This grid deflection was known to occur in prior experiments, such as LUX [51], and was therefore considered when designing the LZ grid wires. A campaign was carried out to measure and model the deflection of the grids as a function of radius and field strength (under the assumption of no deflection) [67]. The result of this work is Equation 4.2,

$$\Delta z(r,E) = A(E)\sqrt{\left(1+\cos\left(\frac{r\pi}{R}\right)\right)\left(B(E)^2+r^2\right)}\left(e^{(r-C(E))}+1\right)^{-1} \tag{4.2}$$

where $\Delta z(r,E)$ is the deflection from the nominal position and A, B and C are polynomial functions of E, the electric field strength, under the assumption of no deflection. The radial position of the deflection is given by r, and hence this equation assumes the deflection applied to each of the grids is symmetric around the TPC of total radius R. This assumption holds true in data since the grid wires are uniformly spaced and aligned, except for a small region of our detector where the gate and anode grid wires are not perfectly aligned as they should be [68]. To date, this has had no observable impact on the size or shape of SE pulses in this region, but we do see a specific topology of S2 pulses with so-called 'stingers' attached. Stingers are formed when the electron in the gas passes through the gaps in the mesh anode rather than accelerating towards the anode wire. The electric field above the anode is not strong enough to produce EL, and so no light is emitted for some time while the electron is pulled back toward the anode. As the electron approaches the anode wires, the field strength increases, and EL photons are emitted again in a brief burst, often classified as a small S1 pulse following the S2. This is an effect that we can model separately and, if necessary, incorporate into the full chain in the future. Still, for now it does not pose any problems for current searches, and hence we are able to produce an axisymmetric field in the gaseous xenon.

In the primary stages of this analysis, we observed that the deflection model was estimating a maximum deflection in the centre of the TPC of $0.18\,\mathrm{mm}$, correlating to an increase in the electric field from $6.62\,\mathrm{kV\,cm^{-1}}$ at the walls of the detector to $8.86\,\mathrm{kV\,cm^{-1}}$ at the centre. This level

of non-uniformity would lead to extremely obvious radial variations in the size and shapes of S2 pulses, which are not observed. After investigating the implementation of this model, two main errors were discovered. Firstly, we investigated the electric field inputs to the model. When applying this model, the deflections of the anode and gate are approximated as equal and opposite. Deflection of both grids is caused by the electric field in which the grids are placed; given that the dielectric constant for LXe is almost double that of GXe, we would expect them to experience different amounts of deflection. The reason we can therefore approximate them to be equal is due to the inclusion of buoyancy forces acting on the gate grid wires, pushing them upwards toward the liquid surface and bridging the gap between the two levels of deflection [67]. It is therefore clear that the electric field input to the model should be the field in the gas because the only contributor here is the electrostatic attraction since the gravitational forces on both wires are negligible in comparison.

When first investigating this implementation, we found that the field input was $E = V/d$ and had assumed parallel plates with a vacuum dielectric. This indicates the electric field input was under-predicted by $0.86\,\mathrm{kV\,cm^{-1}}$. Secondly, we discovered that the model units were inconsistent and hence the resulting deflection was an order of magnitude too large. Correcting for these errors decreases the modelled maximum deflection at the centre of the TPC from $0.18\,\mathrm{mm}$ to $0.03\,\mathrm{mm}$ and significantly improves the field uniformity, shown in Figure 4.4.

4.2.2 Electric Field

Deriving a full 3-dimensional electric field map, like that applied in the liquid xenon, is much harder for the extraction region due to the increasing fields very close to the grid wires, particularly in the small gaps clamping them, where field solutions tend to diverge. For this reason, a simplified approach is taken by assuming a uniform field at every position in the extraction region, accounting at each radial position for the grids' deflection. The additional complication to apply on top of the parallel plate equation is that there are two different media, namely liquid and gaseous xenon, with different dielectric constants. Therefore, we can calculate the electric field strength in liquid xenon, without accounting for grid deflection, using Equation 4.3.

$$E_{liq}(V, d_{gas}, d_{GA}) = \frac{\alpha \cdot V}{(\varepsilon \cdot d_{gas} + (d_{GA} - d_{gas}))} \tag{4.3}$$

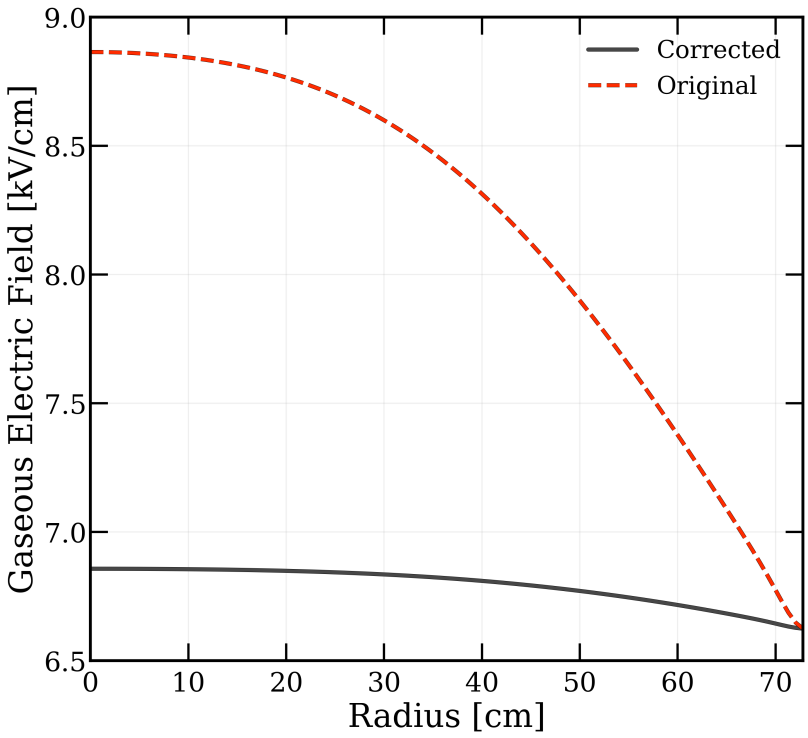


Figure 4.4: Difference in electric field uniformity when correcting the implementation of the deflection model in the extraction region.

where V =7.5 kV is the potential difference between the gate and anode for WS2024, $\varepsilon = \epsilon_L/\epsilon_G = 1.96$ is the ratio of dielectric constants in liquid and gaseous xenon, d_{gas} is the distance between the liquid surface and the anode, d_{GA} is the distance between the gate and anode and $\alpha = 0.932$ is the transparency factor of the field through the grids [67, 68]. To calculate the electric field strength of gaseous xenon, we must account for its dielectric constant.

$$E_{gas}(V, d_{gas}, d_{GA}) = \frac{\varepsilon \cdot \alpha \cdot V}{(\varepsilon \cdot d_{gas} + (d_{GA} - d_{gas}))} = \varepsilon E_{liq} \tag{4.4}$$

Using Equation 4.4, we calculate the input for the deflection model of the grids in Equation 4.2. This affects the distances d_{GA} and d_{gas}. It is important to note that the deflection from Equation 4.2 only applies to a single grid, since both the gate and anode deflect, the value of d_{GA} becomes $d_{GA} - 2\Delta z(r, E)$ and d_{gas} becomes $d_{gas} - \Delta z(r, E)$. These new values are then substituted back into E_{liq} and E_{gas} to compute the actual values for use in the simulation. Combining Equations 4.4, 4.3 and 4.2 results in the electric field map shown in Figure 4.3 where we note again the simplification at this stage of ignoring the increased field strength close to the grid wires, this will be addressed in Section 4.4.

4.3 Electron Drift Speed

In modelling the single electron, the quanta effects can be reasonably separated from the timing effects which contribute to the SE widths. For this reason, we separate this into first discussing the tuning of all timing effects, and then the tuning of quanta and light collection efficiency will follow. In reality, these tunings were performed in parallel with many iterations across various models; for simplicity, here we will separate them to clarify the presentation of the modelling degeneracies. From this point, until Section 4.7, it is assumed that the number of photons produced has been determined as it would have been at this point in simulation.

The electric field strength and gas density directly determine the speed at which electrons drift through the gas. We observe from Figure 4.2 that the widths of SEs are initially underestimated in simulation. The dominant timing effect required to simulate SE pulse widths accurately is the drift speed; hence, the observations in the figure suggest a significant problem with the implementation of this model. In BaccMCTruth, we use the NEST drift model given in Figure 4.5, which is a piecewise, global empirical fit to the results of MAGBOLTZ simulations, returning an average drift speed [69]. This averaging allows us to make a further simplification for the electron drifts by assuming they travel from the liquid surface to the anode in a straight line path. In practice, as shown by dedicated GARFIELD++ simulations for this analysis in Figure 4.6, the electrons exhibit Brownian motion due to numerous collisions with gaseous xenon atoms [70]. This model has been verified against multiple experimental measurements of the gaseous xenon drift speed [33, 46, 71, 72] to be accurate, and hence the error here must be in the implementation.

By cross-referencing the implementation in BACCARAT with the original open-source NEST code we discovered that the NEST model requires the field strength input in $\mathrm{V\,cm^{-1}}$ whereas the BACCARAT input was given in $\mathrm{kV\,cm^{-1}}$, leading to a factor of 10^3 difference and the incorrect piecewise function chosen to calculate speeds. Furthermore, the surrounding code requires the calculated speed to be given in terms of $\mathrm{mm\,ns^{-1}}$, whereas the NEST code outputs the speed in $\mathrm{mm\,\mu s^{-1}}$ which introduces a speed which is thought to be 10^3 times faster than reality, leading to the underestimation of the pulse width. These are simple fixes that yield immediate improvements in the widths simulated for these SE pulses. Prior to these corrections, the electron drift time through the $8\,\mathrm{mm}$ of gaseous xenon in an electric field of $6.85\,\mathrm{kV\,cm^{-1}}$ was predicted to be $226\,\mathrm{ns}$. With

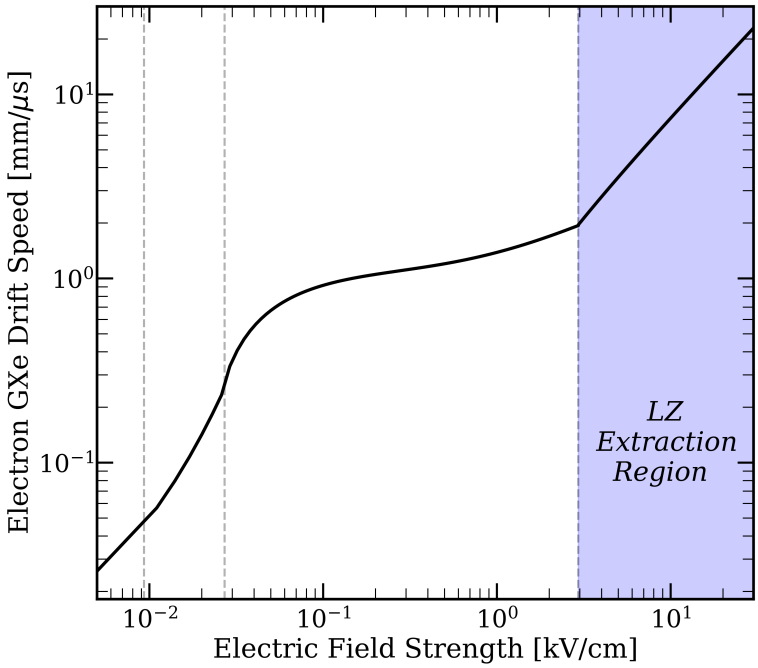


Figure 4.5: NEST drift speed model for electrons in gaseous xenon [41]. This is given in terms of the electric field by assuming a constant gas density and an ideal gas at LZ operating conditions.

these corrections implemented, we predict the drift time to be $1604.5\,\mathrm{ns}$, which is much closer to the widths measured in data, shown in Figure 4.2.

4.4 Photon Production Distribution

Since each electron-xenon interaction is not being individually simulated here, we must determine the position distribution of the origin of the EL photons. As mentioned previously, the electric field strength around the grid wires increases hyperbolically, and thus so does the EL yield. Initially, a uniform distribution of photons is assumed since the field is predominantly uniform along the drift direction; however, a more accurate model will follow a uniform distribution with an exponential increase near the grid wires. This additional exponential increase is necessary to ensure the tails of the SE pulses are accurately reproduced. The parameters for the improved distribution are determined using a dedicated GARFIELD++ simulation.

4.4.1 GARFIELD++ Simulation

The simplest simulation setup to construct this probability distribution is to produce a single cell within the extraction region, with the anode and gate wires oriented in one direction. We neglect the effects of intersections between woven grid wires here, since the field at these positions is difficult to simulate and the anode grid's area coverage is mostly due to individual wires. The single cell setup is shown in Figure 4.6.

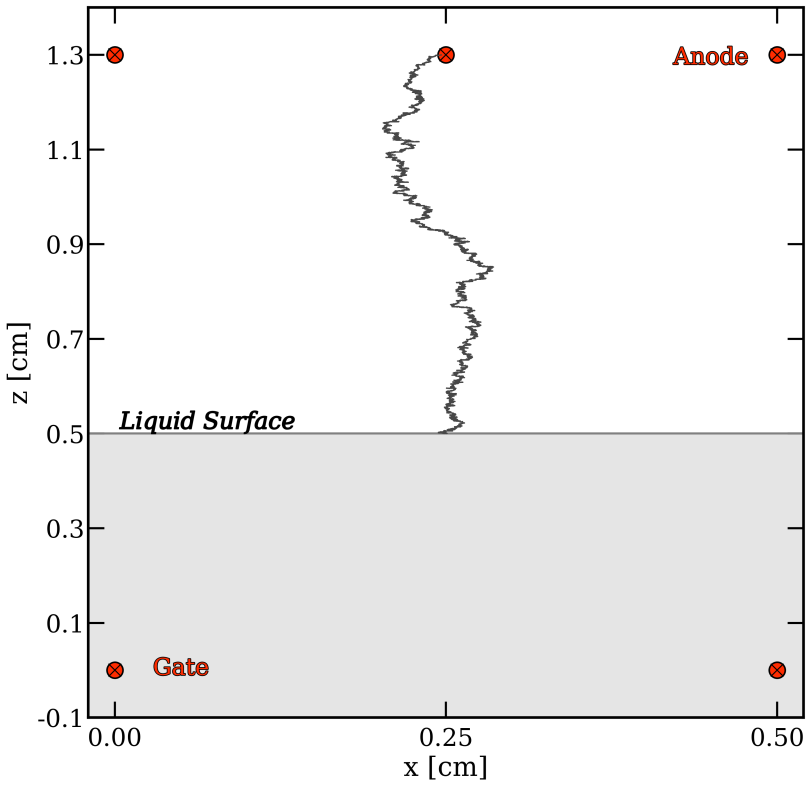


Figure 4.6: Individual cell generated in GARFIELD++ geometry. A simulated electron track from the liquid surface to the gas is also shown. Note that this cell is generated in a 2-dimensional cross-section and does not include the additional field strength from the wires in the x-direction.

By generating electrons from the centre of this cell and recording the positions of EL photons, a 1-dimensional distribution of the probability of EL photons produced as a function of z position is devised. This distribution is presented in Figure 4.7. The distribution is then fit to the functional form

$$f(z) = e^{A(z-B)} + C, \tag{4.5}$$

where A, B and C are constants to be found during the fit. A describes the gradient of the exponential rise in units of $1/distance$, B is the distance along the z-axis at which the exponential rise

begins to dominate, and C is the number of photons produced in all bins before the exponential increase.

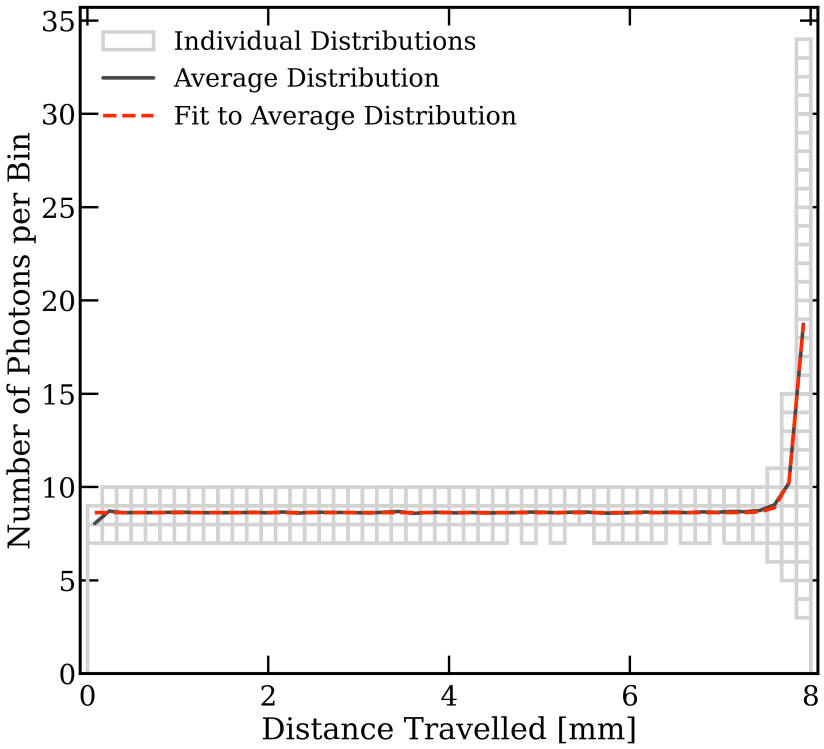


Figure 4.7: Fit to the observed photon production distribution along the path of a single electron through the gaseous xenon. The exponential rise is observed near the anode due to the increased field strength near the high-voltage wires.

Only the nominal distance ($1.3\,\mathrm{cm}$) between the anode and gate is tested using the GARFIELD++ simulation, so it must be generalised to account for grid deflection. To do this, we assert the exponential rise, controlled by A, should be independent of the distance between the gate and anode, so can remain as the fitted value for the entirety of the extraction region. B, however, must change as the maximum distance that the electron travels changes. The fit indicates that the exponential model takes over at 96% of the maximum drift distance; hence, this percentage is used to implement it across the entire extraction region. Finally, C will change with the number of photons produced and can be utilised here as the final free parameter used to normalise the distribution. Performing this normalisation yields

$$C = \frac{N_\gamma - e^{A(d_{max}-B)} + e^{-AB}}{A\, d_{max}},$$

where A, B and C were defined previously, N_γ is the total number of photons produced by the

electron and d_{max} is the maximum distance the electron can travel from the liquid surface to the anode wire, accounting for grid deflection. To generate our final probability distribution function for photon production, we take the functional form from Equation 4.5 and divide by the total number of photons generated to convert all values to probabilities, leading to the probability distribution

$$P(z, d_{max}, N_\gamma) = \frac{1}{N_\gamma}\left(e^{A(z-0.96d_{max})} + \frac{N_\gamma - e^{A(0.04d_{max})} + e^{-A(0.96d_{max})}}{A\, d_{max}}\right) \tag{4.6}$$

We now replace the uniform distribution model with the new fit from Figure 4.7. The optimal method for implementing this custom probability distribution is to use rejection sampling to determine the photon's starting position. By determining the initial position at which the photon is produced, the drift time of the electron to the initial position may now be calculated using the NEST function in Figure 4.5. This time is added to the photon emission time. The inclusion of this effect improves the modelling of SE tails. While it has a small impact on the width of the pulses, it has an observable effect on the shape of high-energy S2 pulses. Since this thesis focuses on the low-energy S2s, we omit a full validation of this effect and instead include this for completeness of the improved modelling of the SE and S2 pulses which will impact the waveform machine learning (ML) techniques detailed in Chapter 7.

4.5 Scintillation Timing

Once the gaseous xenon atoms are excited, they form singlet and triplet states. The decay times of these states are $\mathcal{O}(\mathrm{ns})$ and hence are a known subdominant effect relative to the SE pulse width, which is $\mathcal{O}(\mu\mathrm{s})$. Nevertheless, this should be modelled for improved accuracy of the simulation when fine-tuning detector effects. Following the literature, we adopt a double exponential decay timing profile for the singlet and triplet states,

$$P(t) = \frac{f_s}{\tau_s}e^{-\frac{t}{\tau_s}} + \frac{(1-f_s)}{\tau_t}e^{-\frac{t}{\tau_t}}, \tag{4.7}$$

where f_s is the fraction of singlet excimers and τ_s, τ_t refer to the decay times of the singlet and triplet states, respectively [73]. The global average scintillation time constants measured for gaseous xenon are $\tau_s = 5.18 \pm 1.55\ ns$ and $\tau_t = 100.1 \pm 7.9\ ns$, compared to $\tau_s = 3.1 \pm 0.7\ ns$

and $\tau_t = 24 \pm 1\ ns$ in liquid xenon [41]. The large difference in the triplet state is due to three-body collisions in the gas, which allow for energy transfer between xenon atoms, which stabilises the excimer state. In the Monte Carlo simulation, we must first determine the fraction of singlets and triplets produced after a collision. However, no direct measurements of these fractions exist in gaseous xenon at temperatures near the LZ operating temperature because of the difficulties involved with gas, notably the ability of collisions between excimers to transfer states by energy transfer. To overcome the lack of direct measurements, we must either assume some level of similarity to liquid or higher-temperature measurements, or manually tune this value in our simulations to assess its impact on the width of the SE pulses. In practice, we find that the fraction of singlets has a negligible impact on the SE widths and is therefore not a critical parameter in this simulation at this stage. It will become crucial if LZ ever decides to produce simulations of events originating in the gas, in which the singlet-to-triplet ratio will significantly affect the shape of the S1 pulses in this region. We therefore adopt the most recent room-temperature gaseous xenon singlet-to-triplet ratio measurements taken by the group at Imperial, from which we infer $f_s = 0.2^{+0.04}_{-0.05}$ and note that this is subject to improvements in the future [74].

Once the excimer state is determined, the decay time is sampled using

$$t = -\tau_i \, log(P),$$

where $i \in \{s, t\}$ and P is uniformly sampled between 0 and 1. In doing so, we establish a probability density function for an exponential decay using a uniform variate. At this stage, the full timing modelling up to emission of the photon is complete. The final stage is to add the photon propagation time, which is sampled from probability maps constructed and defined in Section 4.8. Once photon propagation times have been sampled, all timing effects have been modelled, with the dominant improvements coming from the improved implementation of the NEST drift speed model. The final widths require that the pulse area is also accurately simulated to agree with the calibration measurements. The final validation plots showing the improvement in width modelling are shown in Figure 4.17. We now return to the quanta and light collection efficiency in order to tune the SE pulse area.

4.6 SPE Modelling

At this stage, the main timing effects have been corrected, and the SE widths have been significantly improved from Figure 4.2 to Figure 4.17. In the process of tuning the SE pulse area, we identify the 3 dominant processes to model accurately:

1. Single photoelectron (SPE) pulse area
2. EL light yield
3. Light collection efficiency (LCE)

There is a direct degeneracy between the EL light yield and the LCE. However, the SPE area can be tuned independently in the DER by comparing to SPE pulses in the data. We will start with this tuning to ensure we are accurately simulating each photon pulse, before returning to BACCARAT to complete the yield and LCE tunings.

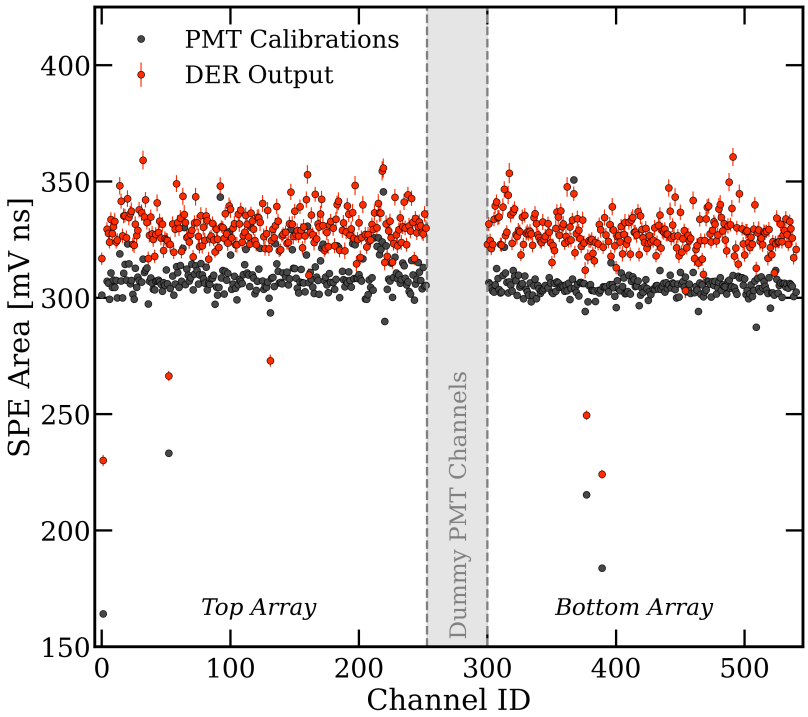


Figure 4.8: Comparison of SPE data from LED calibrations to the default DER settings for SPEs. There is a clear systematic difference between the modelling in simulation and the calibration data.

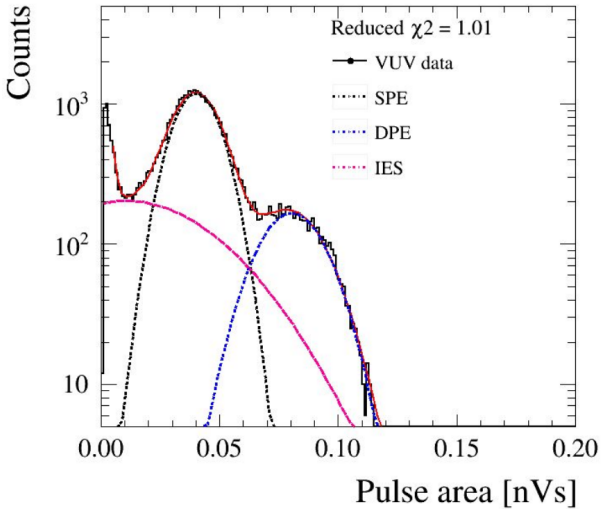
We must first define a data selection for comparison to the simulation. We select pulses which are classified by LZAP as an SPE during a stable run and only observe light in a single PMT for the

duration of the pulse. This SPE data was directly compared to the pulse areas produced from the DER for TPC PMT channels, shown in Figure 4.8. This figure indicates that the SPE pulse area is overestimated for all TPC PMTs by $\sim 35\,\mathrm{mV\,ns}$. This discrepancy must be resolved for any pulses to be accurately simulated with the DER.

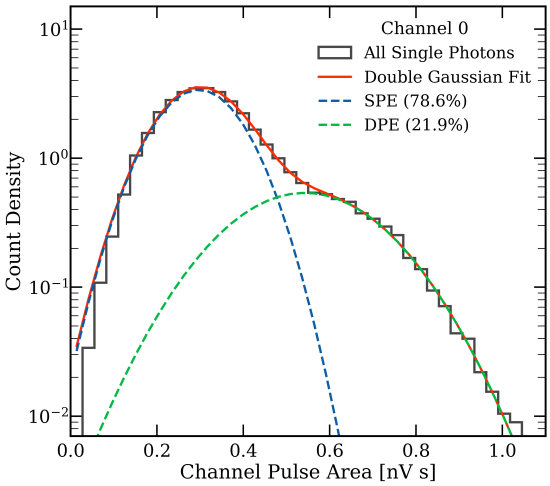
Figures 3.4 and 3.9 demonstrate that there are 5 stages from which this discrepancy could originate: PMT, axon cables, amplifiers, LMR-100 cables and digitisation. All stages after the PMT response are modelled using the circuit diagrams of each component or a simple transfer modelling, both of which we deem accurate. Thus the discrepancy arises from the PMT modelling itself, which is to be investigated in the following.

4.6.1 Secondary PMT Effects

The secondary PMT effects explained in Section 3.1.2.2 include processes that reduce the size of the SPE pulse in a PMT. Namely, first-dynode hits, second-dynode collection, and inelastic scatters. Their relative probabilities were determined by fitting to the Imperial characterisation data [59], which characterised all PMTs with their expected gain of 3.5×10^6. However, LZ PMTs are set to a nominal gain of 2×10^6, which directly affects the probability of secondary effects occurring within these PMTs. Using the same SPE data selection, the tests from the Imperial test stand were reproduced using LZ data, shown in Figure 4.9. The side-by-side comparison here indicates that the first dynode hits peak has likely been removed by a pedestal noise cut within the LZ processing chain, and that inelastic scatters cannot be reasonably represented with the same probabilities derived from the Imperial test stand. Since the IES Gaussian cannot be fitted to this data, we take its contribution to this SPE data selection to be negligible. Additionally, the S1 LCE parameters in BACCARAT had previously been tuned without these secondary effects included [57]. Combining these two reasons, we deem it plausible that these particular secondary effects need not be simulated in the DER, since there is no indication that they affect the final SPEs retained in the data, and we do not wish to undo the S1 LCE tuning if avoidable. Thus, while increasing the probabilities of these secondary effects will decrease the SPE discrepancy from Figure 4.8, there is insufficient evidence that these effects are physically responsible for this discrepancy. Hence, we move to the base PMT pulse modelling.



(a) Imperial test stand calibration of PMT effects. Figure from [75].



(b) Pulse area distribution of SPE pulses in PMT 0 (centre of the top array) in LZ.

Figure 4.9: Comparison of Imperial test stand characterisation of the PMT response with the observed response from a single channel in the top array of LZ. Note that the additional peak at $0\,\mathrm{nV\,s}$ in the Imperial test stand is not present in the LZ data. A triple Gaussian fit to the LZ data to encompass the IES contribution was tested and worsened the goodness of fit in all tested PMTs. Only the SPE and DPE peaks can be observed in the LZ data, with the DPE effect having a larger variance than the Imperial test stand.

4.6.2 PMT Gain

The PMT base pulse from Equation 3.1 is dependent on the gain and timing constants of the PMT. During this investigation, we noticed that the gain values present in the PMT parameter spreadsheet used by the DER were different from those measured by the LED calibrations. The simplest possible fix was to replace the gains in the DER with the measured gains from recent calibrations. If the modelling of the base PMT pulse shape is correct, we expect this to resolve this discrepancy entirely. In practice, this reduces the discrepancy but does not fix it. This reduces the average deviation in the data from $35\,\mathrm{mV\,ns}$ to $20\,\mathrm{mV\,ns}$. Since the remaining parameters controlling the base PMT pulse area are the rise and fall times, which were also determined from fits using the Imperial test stand, we now investigate whether there are any discrepancies in pulse shape.

4.6.3 SPE Pulse Shapes

The only other parameters in Equation 3.1 that can be tuned are the rise and fall times of the PMT base pulse. The TPC PMTs were chosen to have similar response across each array. Maintaining these PMTs at very similar gains, $\sim 2 \times 10^6$, they can be approximated to have the same behaviour in simulation. Hence, we initially assume they have identical rise and fall times for a single photoelectron pulse. An average SPE pulse for each PMT can be constructed from both the real and simulated SPE pulses. An example of a PMT comparison is shown in Figure 4.10, which provides critical evidence that these time constants are inaccurate and directly lead to incorrect SPE pulse areas.

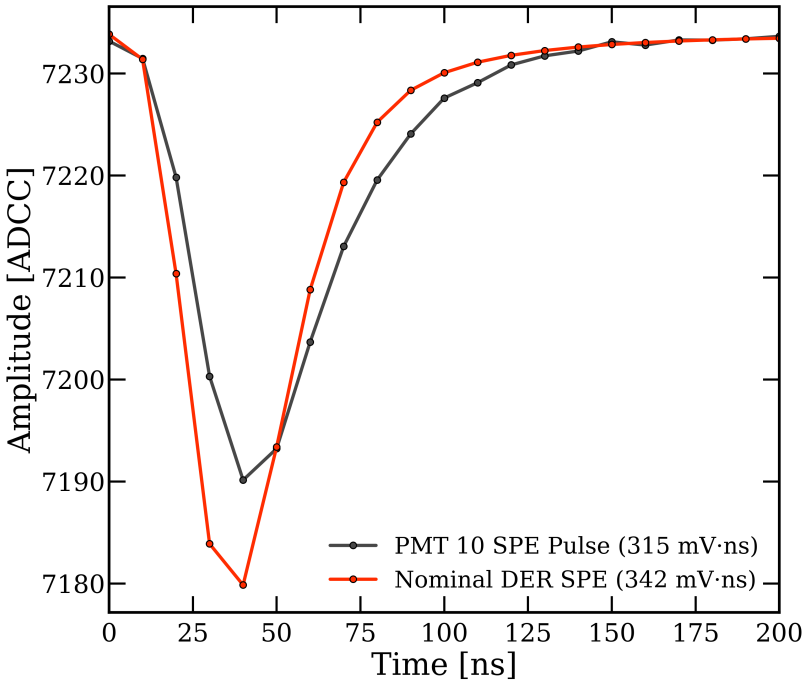


Figure 4.10: Comparison of the average SPE waveform observed in PMT 10 with the SPE pulse using the nominal timing characteristics of every PMT pulse in the TPC from the DER. The variation in SPE size is entirely driven by the gain of each PMT.

From the transfer functions shown in Table 3.1, we can see that all stages after the initial PMT base, including cables, amplification, and digitisation, exhibit timing effects that could also contribute to this discrepancy. However, the rise and fall times measured for the Imperial test data were at a higher gain, which directly affects the timing characteristics of the pulse. Additionally, the timing characteristics of the amplifiers and digitisers are derived directly from the internal circuitry, providing stronger motivation to assume that the transfer functions are correct and to tune the rise

and fall times of the base PMT response.

Since there are some outlier PMTs within the TPC which operate at different gains, we expect the PMTs to have some variation in their associated time constants. We first attempt to extract these constants for all TPC PMTs individually. From the selected SPE data, we produce an average SPE waveform for each PMT in the TPC. Care must be taken here to correctly handle the comparison of the data with results from LED calibrations because of the lack of the DPE effect in the calibrations. The LEDs emit visible light, which lacks sufficient energy to liberate two electrons from the photocathode. However, taking SPE pulses classified by LZAP from a random data sample will include this DPE effect. To address this when constructing the average waveform for each PMT, we apply a simple pulse-area boundary ($1.5\times$ the mean SPE pulse area) for the SPEs included in this template to maintain confidence in the sampled SPE pulses.

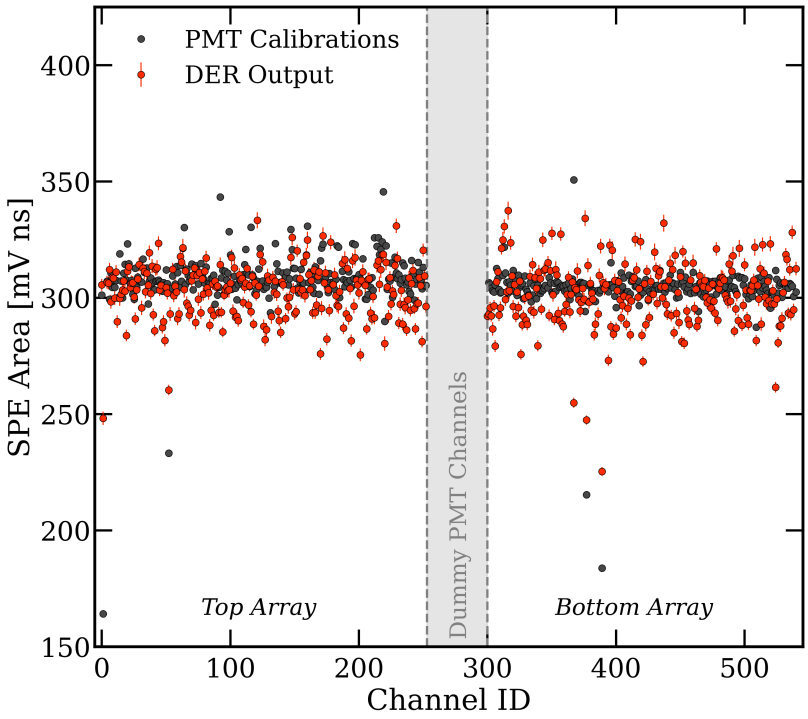


Figure 4.11: SPE pulse area comparisons for individually tuned timing constants for each of the PMTs. Average response has improved but has much larger variance than desired for this result.

The earliest stage of the electronics chain, which is saved to disk by LZ, is the digitiser's output. Hence, to fit time constants to the average SPE shapes from data, we first construct a function which takes in just the gain of the PMT, which is now taken directly from LED calibrations, and has the rise and fall times as fit parameters. This function then constructs the base PMT pulse from

Equation 3.1 and passes it through each of the electronics stages observed in Figure 3.4. We fit this function to the waveform templates generated for each PMT using a least-squares fit. The fit values are then passed back through the DER, altering the DER code to allow a new rise and fall time to be taken for each PMT. We also take care here to ensure the DPE effect is switched off in the DER so that these results are directly comparable to the LED calibration results. The result of this comparison is shown in Figure 4.11.

While the mean of this distribution is now much closer to the measured values, the variance of the pulse areas is wider than expected. This is due to some fits having larger errors because of difficulties in identifying the start of the average pulse waveform. The solution to this was to identify the global average rise and fall times for all PMTs and use these values to approximate all PMTs as equal. This results in τ_r = 0.48 ± 0.04 ns (previously 3.81 ns) and $\tau_f = 5.36 \pm 0.10$ ns (previously 4.76 ns) and yields the comparison to LED calibrations displayed in Figure 4.12 demonstrating the improved accuracy of the pulse area simulated for each PMT when compared to Figure 4.8. For completeness, a comparison of the new SPE shape is also given in Figure 4.13 to validate the widths of the new SPE pulse shape.

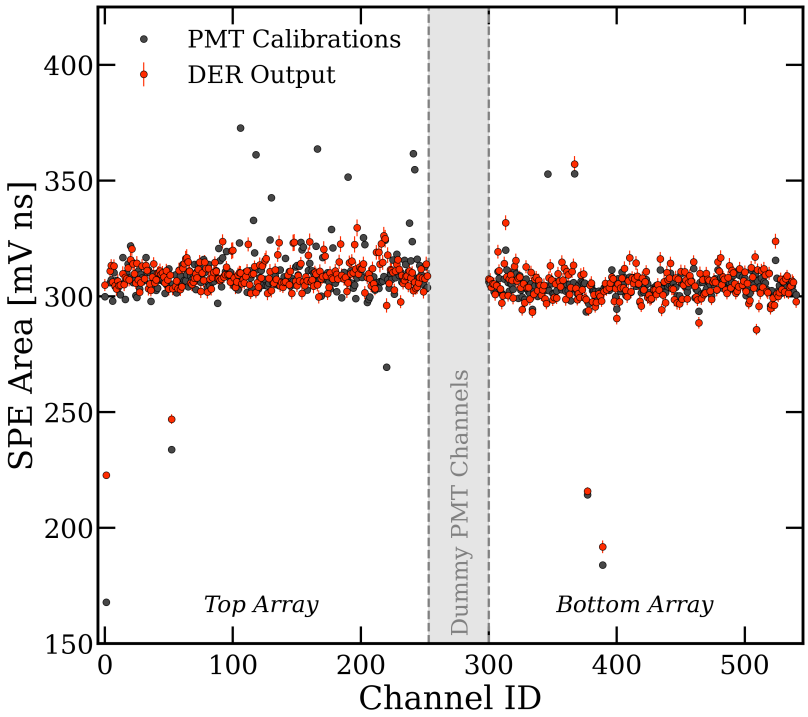


Figure 4.12: Final SPE pulse area comparisons with fixed $\tau_r = 0.48$ ns and τ_f =5.36 ns. Some large outliers remain for the lower gain PMTs, but the majority of TPC PMTs align well with the LED calibrations.

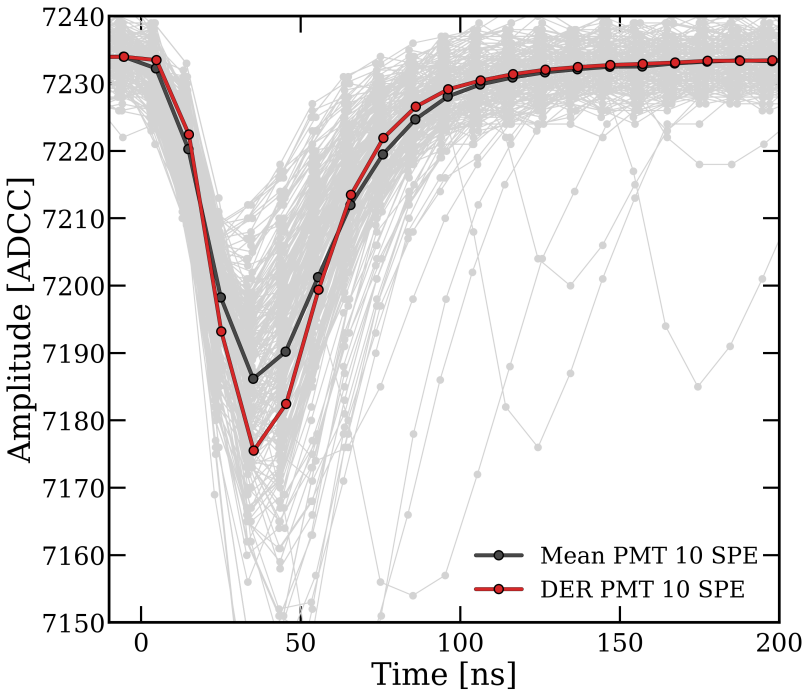


Figure 4.13: Comparison of tuned time constants for the same PMT 10 example as in Figure 4.10. The plot shows the full distribution of all SPEs contributing to the mean SPE pulse from data, showing some consistency with the simulation, despite the simulation having a larger pulse area.

We note here a discrepancy observed specifically for PMT 1 in Figure 4.12. The simulation significantly overestimates the SPE size in this PMT, which operates at a lower gain than most of the PMTs. This is because in the centre of the top array, shown in Figure 3.5, there are 4 disabled PMTs and thus, to not lose a large proportion of the central TPC volume, the gain is decreased in this PMT such that it can remain operational. This therefore breaks the expected flat modelling of all PMTs, which manifests itself as a small over-density in position reconstruction. During the reconstruction process, LZAP compares the expected SPE size from calibrations with the observed quantity of light to convert it into phd. Since the simulated PMT 1 will always have more light than expected per SPE, LZAP weights the reconstruction towards this PMT. Further investigations to resolve this discrepancy are left to future analyses; for now, we note that this has no observable effect on any other analysis RQs, except for the position reconstruction. Since this PMT is in the centre of the TPC, this has very little effect on the analysis.

4.7 Electroluminescence Yield

Once the SPE modelling accurately describes the expected PMT response, we return to BACCARAT to finalise the remaining yield tuning. During the drift of the electrons through the gas, we must determine the EL light yield and, subsequently, the GEANT4 material parameters required to match the LCE.

4.7.1 Araújo and Chepel

The electroluminescence light yield in gaseous xenon is well modelled and understood to be linear in E/n, the ratio of electric field to number density. For practicality, this can be represented in terms of E/P, the ratio of electric field to gas pressure; since $n \propto P$, the modelling remains linear. The Chepel and Araújo model was the accepted model for LZ at the beginning of this work [76].

$$N_{\gamma/e} = (0.137E_{gas} - 177P - 45.7)d_{gas}, \tag{4.8}$$

where $N_{\gamma/e}$ is the mean number of EL photons produced per electron extracted into the gas. All other symbols retain their defined meanings. From this modelling, we observe that an electron extracted in the centre of the TPC at a pressure of P =1.86 bar, with maximum grid deflection and thus strongest field strength, will produce $\sim$ 440 photons, decreasing to 430 at the edges of the detector. Recall that the WS2024 SE size has a mean of 44.5 phd. The SE size can then be represented as

$$SE = N_{\gamma/e}\,\varepsilon_{LC}\,\varepsilon_{Q}\,f_{DPE}, \tag{4.9}$$

where ε_{LC} is the LCE (specifically, the fraction of photons which reach PMT windows), ε_Q is the quantum efficiency of the PMT windows, and $f_{DPE} = 1.2$ is the mean DPE correction factor. Rearranging this gives an estimate of the required ε_{LC}, assuming the model in Equation 4.8 is correct, of $\varepsilon_{LC} = 44.5/(440 \times 0.3 \times 1.2) = 0.28$.

4.8 S2 Light Maps

The LCE for S2-like pulses in BACCARAT is handled in the S2 light maps. S1 photons are propagated using GEANT4 stepping actions, which proves to be computationally expensive and provides

a bottleneck for simulation speed. If S2 photons were to follow the same propagation methods, given the increased photon yield per electron (assumed for now to be $440\,\mathrm{photons}$), computations would become impractical. To prevent this, light maps are constructed by performing a single large-scale, $\mathcal{O}(10^9)$ photon simulation in which photons are uniformly produced in the extraction region gas volume. These photons are propagated around the entire TPC volume, with their paths and times recorded. The gas volume is then spatially binned, with each bin occupying a $5\,\mathrm{mm} \times 5\,\mathrm{mm} \times 1\,\mathrm{mm}$ cuboid. From this, the following maps are produced:

- **Hit Probability Maps:** These maps contain the probability per bin that a photon originating in that bin hits any of the PMT faces.
- **PMT Hit Maps:** For each PMT, a map is constructed containing the probability that a photon hits the given PMT in each of the spatial bins.
- **Time Maps:** These contain the probability that a photon hits a PMT from a radial distance, r, away in time, t, and hence we can sample the photon propagation times.

Photon propagation throughout the LZ geometry is controlled by material parameters such as reflectivities and absorption lengths. Material parameters for LXe physics were previously tuned to match S1 pulse areas, and thus changing LXe material parameters is avoided in this tuning since SEs are mostly collected by the top array and hence are predominantly subject to the GXe material parameters. In practice, the LXe parameters were tuned to S1 pulses while the SPE pulse areas in the DER were overestimated, inevitably leading to some uncertainty in these values. Retuning an optical model consistent with both the S1 and S2 sizes is left for future analyses since this thesis only requires S2 pulses to have accurate LCE. This does not pose a problem to any future results in this thesis, since the S2 light maps are self-contained, the tuned S1 LCE values remain wherever S1s are generated and the S2s are handled separately within the S2 light maps.

LZ can isolate the amount of light detected by the top and bottom arrays during analysis. It is therefore important that we tune these parameters in an array-dependent manner to conserve the top-bottom asymmetry. Tests were conducted to determine which parameters had the greatest effect on the SE LCE and TBA. These parameters and their default values are:

- GXe-PTFE Reflectivity (85%)
- GXe-Steel Reflectivity (20%)
- GXe-Kapton Reflectivity (20%)
- LXe-Steel Reflectivity (20%)

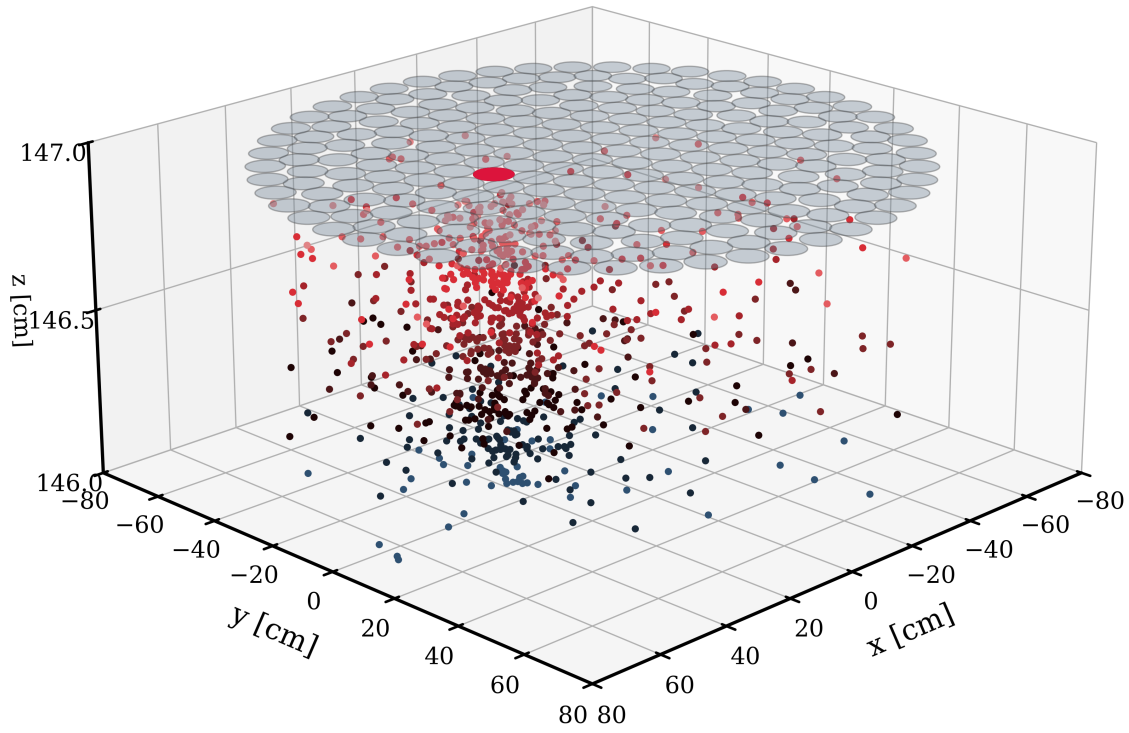


Figure 4.14: PMT hit probability map for PMT 30 on the top array of the TPC. Each data point shows the starting position of a photon which hits the PMT highlighted in red. The colour scale here is to allow the reader to visualise the depth. While there is a higher density of points below the PMT, the plot demonstrates that there is a broad distribution of potential starting positions.

Kapton refers to the small space around the edge of the detector's top flange that is outfitted with Kapton tape [40]. This is sometimes referred to within LZ as the 'Black Wall' since ideally it should have very low reflectivity. The GXe-Steel interface controls reflections from the anode grid, and the LXe-Steel reflectivity controls both the gate and cathode grids. The LXe-Steel reflectivity is chosen to be included here, despite trying to avoid the LXe parameters, since it is strongly correlated to the TBA of the pulse. With the cathode grid covering the bottom PMTs, decreasing their reflectivity directly increases the amount of light allowed to hit the bottom PMTs. A similar effect is true of the anode grids; however, the correlation here is weaker since there is only one grid to pass through towards the top PMT array. The impact of each parameter on the TBA was individually assessed while keeping the others at their nominal values. This demonstrated that the strongest correlation was from GXe-PTFE and LXe-Steel reflectivities. To reach the required 28% LCE, both parameters would have to be pushed above 100% reflectivity, which is unphysical. We therefore conclude that the EL yield model must be underestimating the light yield per electron.

4.8.1 Dey Model

Upon inspection, with more recent results published [77–79], the Chepel and Araújo model is found to be inaccurate mostly due to the over-generalisation of detector conditions. With many experiments since publishing new yield models, NEST has updated and improved its own light yield models. NEST predicts that LZ should produce ~ 613 photons per electron for the WS2024 configuration. Increasing the number of emitted photons allows the material parameters to remain within a physically reasonable range. However, implementing the NEST model in BaccMCTruth is more complex than the simple linear model fit proposed by S. Dey [63]. The Dey model combines measurements of light yields taken directly from LZ and other experiments with similar detector conditions and produces a fit, linear in E_{gas}/P, which produces a light yield of $N_{\gamma/e} = 574\,\text{photons}$, much closer to the NEST prediction than the Chepel and Araújo model. The Dey model is given as

$$N_{\gamma/e} = (a\frac{E_{gas}}{P} - b)\ Pd_{gas}, \tag{4.10}$$

where $a = 144.0 \pm 7.0\ photons\ e^{-1}\ kV^{-1}$ and $b = 136.0 \pm 11.0\ photons\ e^{-1}\ cm^{-1}\ bar^{-1}$ are constants determined from the fit. After implementing this model, the tuning of the material parameters is revisited.

To tune the four material parameters of interest, we first must measure the expected LCE from SE data. We again adopt the model for the SE pulse area given by Equation 4.9. This can be rearranged to calculate the expected $\varepsilon_{LC} = \langle SE \rangle / (N_{\gamma/e} \cdot \varepsilon_Q)$ where the DPE probability is absorbed into the measured $\langle SE \rangle$. However, this requires complete confidence in both the EL yield model and the PMT QE measurements. To reduce possible degeneracy while also testing the confidence of the QE measurements, we instead measure the ratio $\langle SE \rangle / N_{\gamma/e}$, which is then directly comparable to the product $\varepsilon_{LC} \cdot \varepsilon_Q$, both measured from the simulation.

A grid search of these parameters within the expected ranges reported in [40] is conducted and compared with the ratios measured in the data for the total SE size, as well as for the individual top and bottom arrays. The total SE comparison is shown in Figure 4.15. The optimal parameters are determined by the minimisation of all three tuning matrices to be

- GXe-PTFE Reflectivity = 60%
- GXe-Steel Reflectivity = 5%
- GXe-Kapton Reflectivity = 5%
- LXe-Steel Reflectivity = 5%

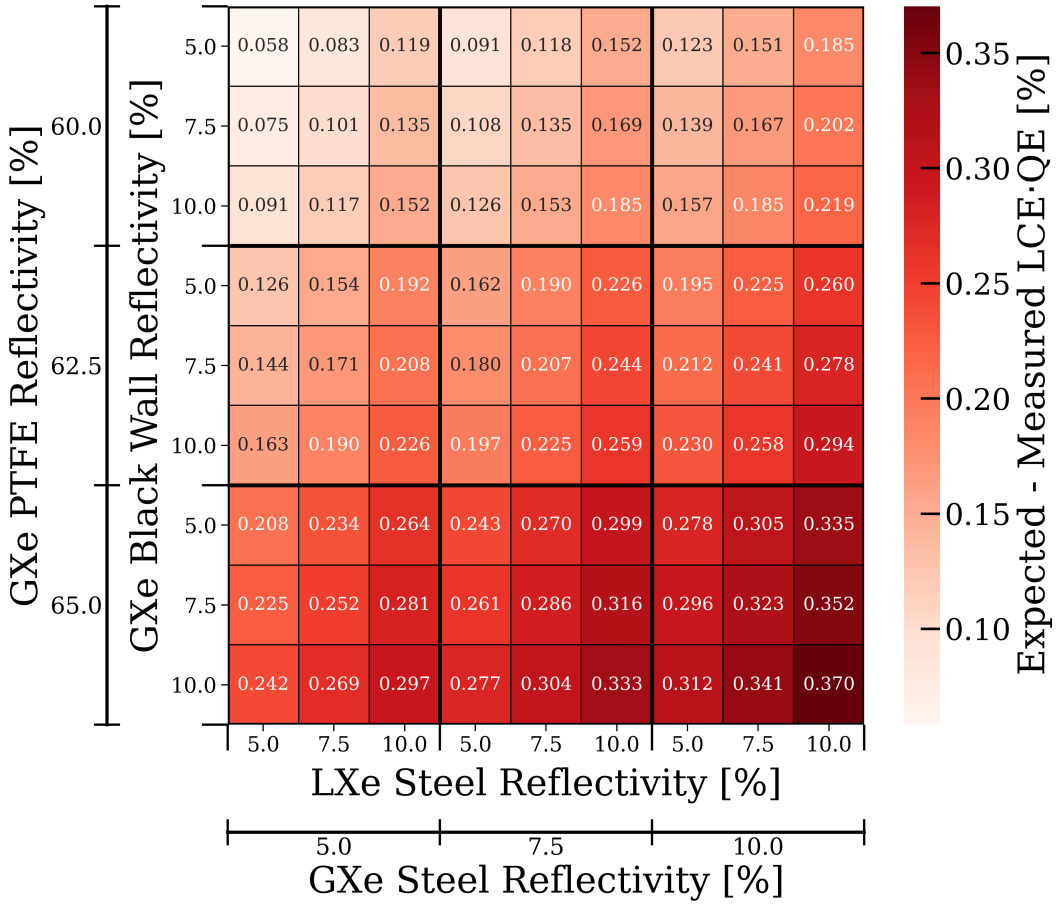


Figure 4.15: Grid search loss matrix showing the difference from the expected convolution of LCE·QE from the measured value from the simulation output. The variation in these values indicates the level of fine-tuning occurring here, and the optimal configuration is identified as the minimum value.

These values show very good agreement with the SE data and are consistent with the expected ranges reported in recent measurements. The GXe-PTFE reflectivity parameter was measured by Silva (2009) to vary from 47% - 66%, making this tuning consistent at 60% [80]. Kapton reflectivity has not been directly measured experimentally but is explicitly chosen in the experimental design due to its low reflectivity; hence, a tuning of 5% here is a reasonable assumption that could potentially decrease in future iterations [81]. Finally, the steel grid reflectivity in LZ is initially assumed to be between 0% and 40%, and a baseline of 20% is chosen since the grids were electropolished, increasing their reflectivity [40]. However, no direct measurements were taken during commissioning, so the true reflectivity is unknown. Nevertheless, reports by Bajpai (2024) are consistent with a measured reflectivity of 5.8% for steel wires in liquid xenon [82]. This result also indicates a large specular component, which can increase this reflectivity at larger angles, but

the most reasonable assumption remains at 20° and 5.8%, consistent with the result of this tuning. These tuned parameters should be interpreted not as measurements of these material parameters, but rather as effective optical parameters tuned to reproduce the LCE expected in the data, while potentially absorbing inaccuracies present in other models throughout the simulation. Further work is required to ensure a global optical model is formed for the simulation. Using the current tuned values, we construct light maps that accurately reproduce the SE pulse areas, as shown in Figure 4.16.

4.9 Tuning Results

The final step is to validate the simulated SE pulses against the SE dataset attained from Section 4.1. Since analysis at this stage in LZ is performed entirely using RQs rather than full waveforms, it is sufficient to demonstrate consistency with the RQ distributions. However, since Chapter 7 makes use of waveforms to pioneer ML methods that perform analysis using waveforms in addition to RQs, we also show here a comparison of the average SE waveform templates for simulation and data.

Firstly, Figure 4.16 shows the comparison of the array-dependent areas for the SE, which are now consistent with the data due to the improvements outlined in this chapter. Secondly, Figure 4.17 shows the agreement with the width of the SE pulse. Finally, Figure 4.18 demonstrates the consistency of the average SE waveform from simulation and data, providing a solid baseline upon which to build a full S2 modelling for full waveform analysis.

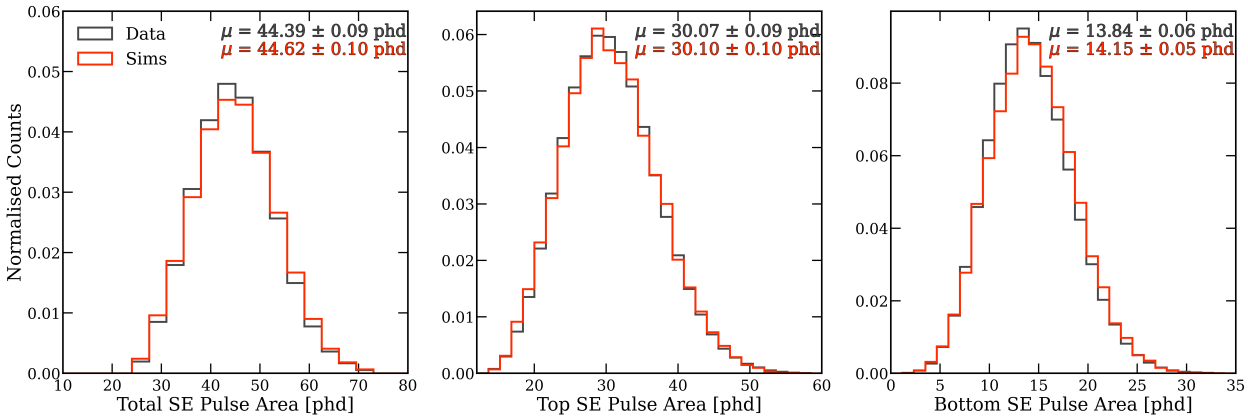


Figure 4.16: Comparison of the SE pulse areas observed split into the total SE (top + bottom) and each of the top and bottom PMT arrays. The mean SE pulse area is written in the top-right corner of each plot.

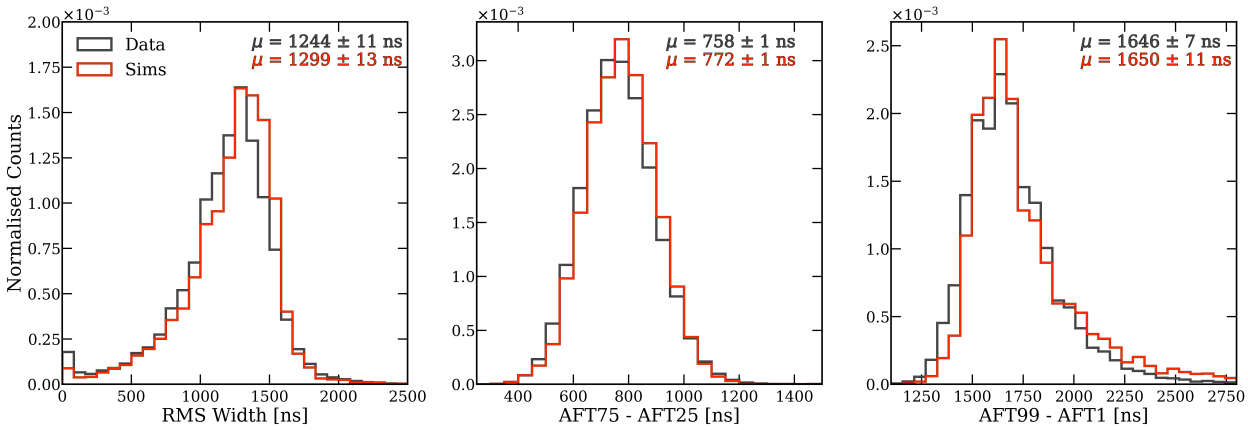


Figure 4.17: Comparison of different SE width metrics including RMS width, AFT75-AFT25 and AFT99-AFT1. Notice that the RMS and 99-1 widths are still slightly overestimated.

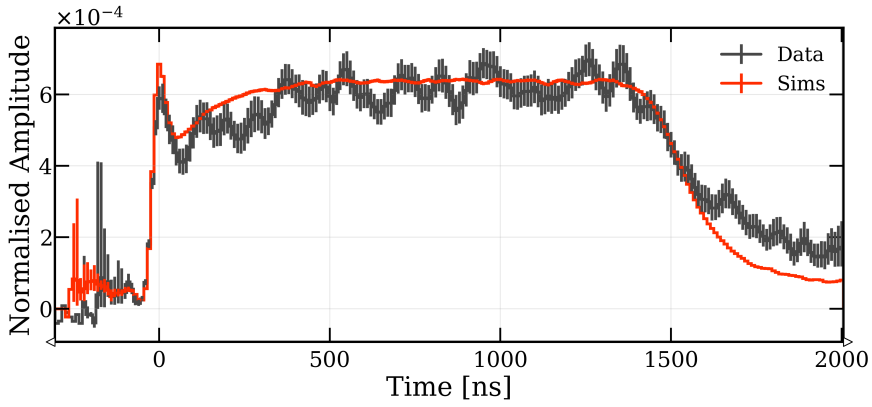


Figure 4.18: Comparison of the average SE pulse shape from simulation and data. The data SE waveforms here are a small sample ($\mathcal{O}(100)$ events) taken from a CH_3T calibration run, where all criteria derived in Section 4.1 are applied along with a limit of 10 SE pulses per event to eliminate electron train backgrounds. The average waveforms are aligned with 0 at AFT1. The tail effects in the data are more prominent than in the simulation, likely due to the inclusion of effects such as stingers, electron trains, and photoionisation in the data but not in the simulation.

Figure 4.17 displays a representation of the width of the SE pulses in the form of the root-mean-square (RMS) width, which indicates the spread of the pulse in time. This is chosen because the S2O analyses use the RMS width of S2 pulses as a proxy for depth in the detector due to diffusion. Given that these simulations are intended to aid the S2O search, it is important to validate this RQ. Of all distributions shown in this validation, the mean RMS width has the largest discrepancy with data, overestimated by $\sim 55\,\mathrm{ns}$. This discrepancy is thought to be caused by the accuracy with which LZAP defines the start and end points of a pulse. This would explain why the RMS is more sensitive than the two AFT widths displayed in the figure. Alternatively, the RMS may be more sensitive to optical modelling in the LXe. As mentioned previously, these parameters remained unchanged during this tuning, which could lead to inaccuracies in the photon path to the bottom PMT array. Attention should also be brought to the AFT99 - AFT1 comparison, which demonstrates the necessity of including the new photon distribution model defined by Equation 4.6 to simulate the tail in this distribution to higher SE widths.

The improvements presented in this chapter are crucial both for advancing the understanding of the LZ detector physics and for providing a foundation upon which we can begin to build S2 waveforms. Increasing the accuracy of the optical modelling parameters initiates the transition for LZ to use the waveform simulation not just as a sanity check but as a tool with the ability to significantly enhance background discrimination using S2s as will be demonstrated in Chapter 7. However, we must first ensure S2s are accurately modelled in the full chain simulation.

5 | Simulating S2 Pulses

Accurately simulating S2 pulses with the full chain becomes increasingly important for LZ as it approaches the low-energy thresholds, where backgrounds such as accidentals make high-quality calibrations difficult. Having a validated high-fidelity waveform simulation becomes a valuable asset to the analysis procedure. Furthermore, many effects discussed in this chapter can be decoupled to the individual electrons forming the S2. This implies, firstly, that accurately simulating SE pulses in Chapter 4 plays a crucial role in accurately simulating S2 pulses; and secondly, that many models implemented to form the S2s are energy independent. This means that accurately modelling low-energy calibration sources will inherently improve the modelling of high-energy sources, thereby enhancing the overall applicability of the full chain simulation as a tool for discovery.

All effects occurring in the gaseous xenon have been validated with SE pulses in Chapter 4. With these SE pulses tuned, the comparison of simulation and calibration tritium, CH_3T, S2s is shown in Figure 5.1. Hence, the focus in this chapter must now shift below the liquid surface. In this chapter we firstly investigate S2 electron yields, then their transport to the liquid surface including diffusion and electron-ion recombination effects, and finally their extraction into the gas phase. We also include a discussion of further developments to the DER afterpulsing model discovered during this work. Most importantly, all of these updates must be well validated before being trusted in analysis. This is achieved by comparison to calibration data, and more strictly by use of a classification neural network.

5.1 Clustering and Yields

In Chapter 3, it was explained that the initial BACCARAT stages rely on using GEANT4 to generate the primary particles, stated in the macro file, and propagate them through the LZ geometry, recording all energy deposits with their positions and times. These energy deposits are then clustered into physical interaction sites to calculate the light and charge yields produced in each interaction in the liquid xenon. Essentially, since GEANT4 operates at a much finer precision than the position resolution of LZ, $\mathcal{O}(\mathrm{mm})$, it is impractical to process each energy deposit. Instead, they are clustered, and each cluster is treated as a single energy-deposition point. This is important

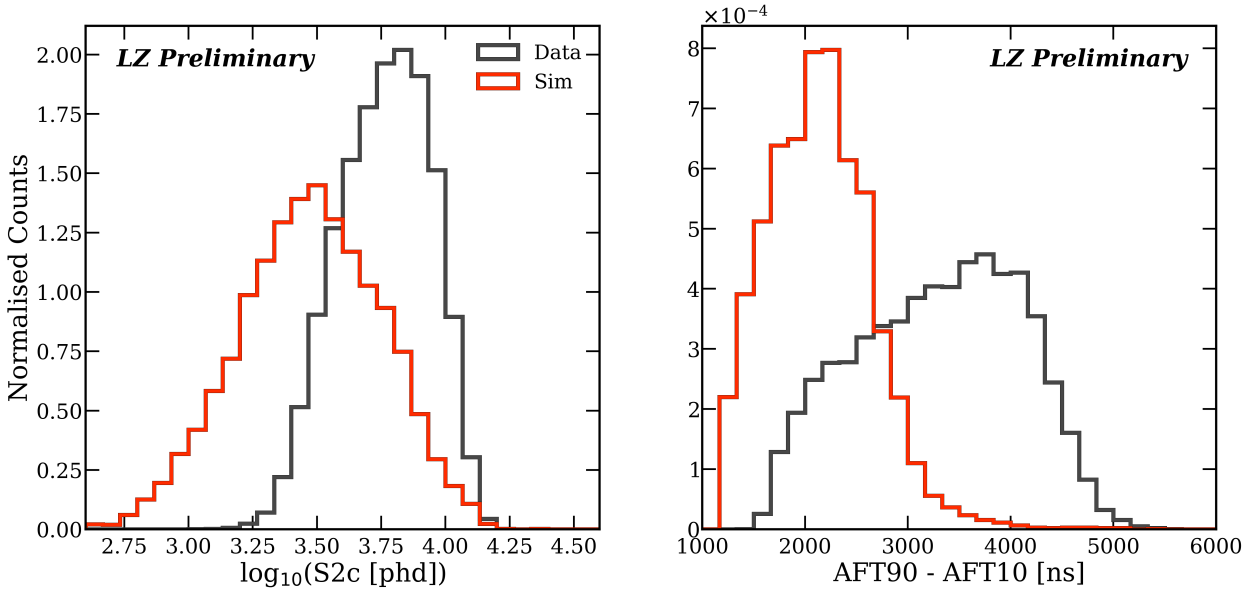


Figure 5.1: Direct comparison of the pulse area and AFT90 - AFT10 widths for simulated and calibration tritium, CH_3T, S2 pulses with a tuned SE response following Chapter 4.

when accounting for the difference between SS and MS events. Clustering deposits within a larger volume than necessary may merge multiple distinct interaction sites, producing a false SS event. Similarly, clustering within too small of a volume could lead to SS events being misclassified as MS events.

There are two clustering algorithms available in BACCARAT; clustering by lineage or by density (DBSCAN). These approaches are detailed in Chapter 3, where we state that the full chain simulation utilises lineage clustering as the default method, and the density-based clustering algorithm is used sparingly for high-energy interactions.

5.1.1 Light and Charge Yields

Once all clusters are defined using either of the aforementioned algorithms, they are then processed to quantify the light and charge yields of each cluster. The light yield, L_y, is the number of VUV scintillation photons emitted during the interaction per unit of energy deposited, while the charge yield, Q_y, denotes the number of ionised electrons emitted per unit of energy deposited. This calculation is performed using models from NEST. These depend on the interaction type; hence, each cluster must also contain information about the initial particle that deposited the energy. The four most common yield models, depicted in Figure 5.2, and their criteria are:

- **Beta**

  - Parent particle is an electron or a gamma. If the parent is a gamma the interaction process cannot be the photoelectric effect.

- **Gamma**

  - Parent particle is a gamma and the interaction process is the photoelectric effect.

- **Nuclear Recoil**

  - Parent particle is a neutron, or the child particle from the interaction is a xenon nucleus.

- **Ion**

  - Any decay process which produces an atomic nucleus as a daughter.

Each of these interaction types requires a different yield model since the initial interactions are fundamentally different. For NR interactions, the incident particle interacts directly with the nucleus. It transfers most of its energy to the xenon nucleus kinetically and thus has a smaller charge yield than the beta and gamma interactions, which will be ER interactions. In ER interactions, the energy of the incident particle is transferred directly to the electrons in atoms, causing them to be excited and ionised. Given that the partition of energy towards the electrons is larger for the ER events, they exhibit a larger charge yield at the same energy for a single interaction than NR interactions. Ion interactions are classified as decays in which the daughter particle is not an electron, for example, alpha decays. Ion interactions generally produce much more light than charge due to the increased electron-ion recombination for their point-like energy depositions.

In practice, L_y and Q_y are also field-dependent functions. Stronger fields reduce the probability of electron-ion recombination, thereby increasing the number of ionised electrons and decreasing the recombination contribution to the scintillation yield. Figure 5.2 demonstrates the yield models for these interaction types calculated for a fixed field of $97\,\mathrm{V\,cm^{-1}}$. Both the light and charge yield models are well validated in the WIMP search energy regime of $5\,\mathrm{keV}$ to $70\,\mathrm{keV}$, as this region is well calibrated across all dual-phase xenon TPCs [32, 36, 62].

During validation studies as the final stage of tuning S2 waveforms, large discrepancies were observed in the NR response of neutron calibration sources such as AmBe. Specifically, these simulations showed that the initial interactions with the neutrons produced the expected L_y and Q_y, but subsequent interactions of the gamma particles produced as secondaries produced much larger numbers of photons and electrons, which flooded the initial NR signal. This is thought to be due to

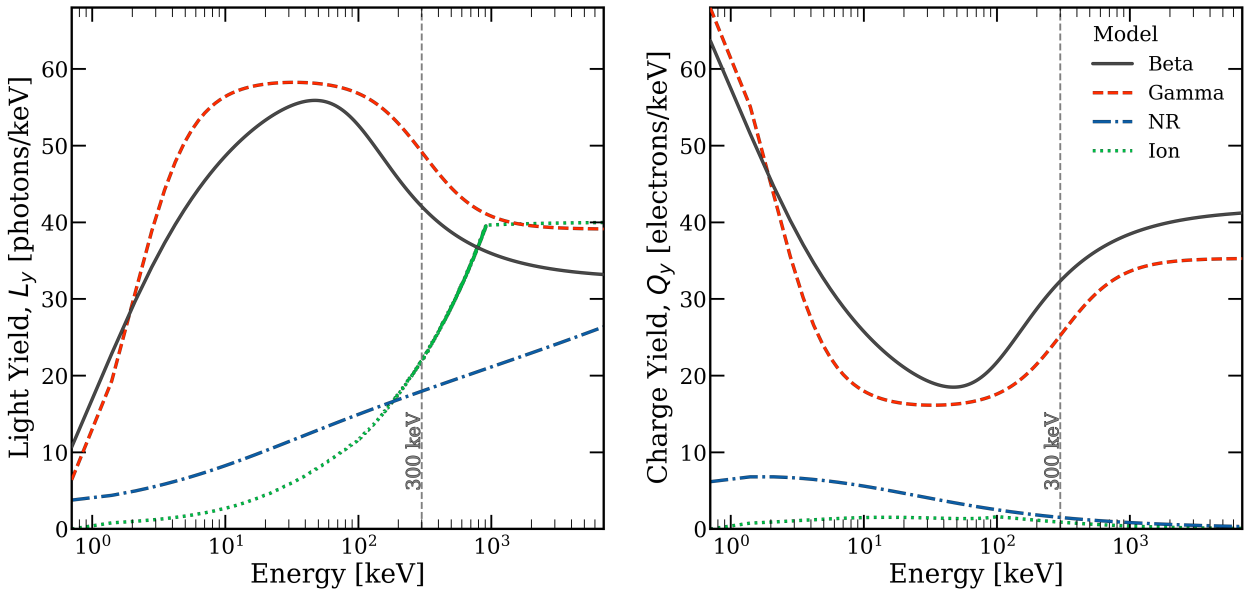


Figure 5.2: Light and charge yield models from NEST shown for the four different response models introduced in this thesis. The 300 keV line represents the upper bound of the NEST- tuned parameters for the NR response fixed by the endpoint of the AmBe neutron spectrum. Above this line, the NR model is an extrapolation. The beta model is tuned to a combination of the CH_3T and ^{214}Bi beta spectra extending up to ∼3 MeV.

an error in the implementation of the lineage clustering algorithm, rather than an issue in the NR modelling, since the yields from the fast chain simulation agree well with the data and do not rely on the lineage clustering algorithm. The exact nature of this issue requires further investigation but is likely caused by the interplay between the lineage clustering algorithm and the NEST gamma yield models at energies above ∼100 keV where gamma calibrations are scarce. To circumvent this issue, a new generator was developed named `ParticleBomb`, which is capable of producing a fixed number of photons or electrons from a random position in the detector, bypassing the need to validate this model for NR events for now and leaving a full analysis of the NR response to future analysers. The `ParticleBomb` generator is also useful for S2O searches, as will be discussed in Chapter 7. Validations of the beta response model show much better agreement with the data, as demonstrated throughout this chapter.

5.2 Charge Transport

Transporting the generated electrons to the liquid xenon surface via GEANT4's stepping action can be computationally expensive, as each step length depends on the physical processes occurring at

that step and has no specified minimum. For distances of $\mathcal{O}(\mathrm{m})$ in LZ, this could require very large quantities of steps. This would also require well-defined field simulations to mimic the detector conditions over very short intervals, which would increase computation time. Instead, LZ uses the field and drift maps from LZEF to transport the ionised electrons to the surface.

The drift maps are 3D axi-symmetric maps with predetermined drift times for each spatial bin in the detector determined by the electric field map. Drift times are sampled from this map depending on the initial position of the generated electrons. The diffusion model then adjusts this drift time.

5.2.1 Diffusion

Much like electrons drifting through the gas to the anode, those drifting through the liquid up to the surface are subject to many microscopic collisions that prevent their paths from being direct straight lines. Instead, these collisions cause the electrons to diffuse from their initial point-like position in the detector following the general diffusion equation,

$$\frac{\partial n(r,t)}{\partial t} = \nabla \cdot (D(n(r,t),r)\nabla n(r,t)), \tag{5.1}$$

where $n(r,t)$ is the electron density at point r in the medium at time t and $D(n(r,t),r)$ is the diffusion coefficient [83]. This equation describes the general diffusion of particles from a point source. However, the additional electric field in the TPC forces electrons to move in the z direction. Additionally, the assumption that the diffusion coefficient is independent of n and r holds for a uniform distribution of liquid xenon at constant temperature. Thus $\nabla \cdot D = 0$. Equation 5.1 can be constructed with the additional driving field force as

$$\frac{\partial n}{\partial t} = D\nabla^2 n - v_d \frac{\partial n}{\partial z}, \tag{5.2}$$

where v_d is the drift velocity in the liquid xenon. Equation 5.2 can then be applied to the LZ TPC, with the note that the diffusion in the TPC is anisotropic due to the applied field in the z direction. Therefore, the diffusion can be separated into the longitudinal (along the drift axis) and transverse (perpendicular to the drift axis) equations. Perpendicular to the drift axis, there is no applied field and hence $v_d = 0$ in this direction, leaving the pure diffusion equation as $\partial n/\partial t = D_T \nabla^2 n$ where D_T is the transverse diffusion coefficient. Conversely, the longitudinal diffusion coefficient D_L is

defined following the relationship in Equation 5.2.

Solving these equations shows that the electrons originating from a point position evolve following a Gaussian distribution with $\sigma^2_{L,T} = 2D_{L,T}t$, where σ denotes the standard deviation in space. The simulation requires both the spatial and temporal deviations that arise from this diffusion. The standard deviation in time, σ_t, is given by $\sigma_t = \sigma_{L,T}/v_d$. Therefore, simulating these diffusion effects can be performed by sampling from their given Gaussian distributions in both space and time for a known v_d. Hence, all that is required are measurements of D_L and D_T.

Transverse diffusion of the electrons leads to the spreading of the electron cloud in the (x, y) plane. A larger spread of electrons extracted from the surface can make it difficult to reconstruct the (x, y) position from the top array hit map. Though there are many measurements for D_T in liquid xenon [84–86], none exist at the field strength of LZ. The closest measurement in the literature is from W.T. Chen at Subatech, measuring $D_T \simeq 34\,\mathrm{cm^2\,s^{-1}}$ at $171\,\mathrm{K}$ and $500\,\mathrm{V\,cm^{-1}}$ [85]. In comparison, LZ operates at $175\,\mathrm{K}$ and has a drift field of $97\,\mathrm{V\,cm^{-1}}$. Extrapolating the trend observed by W.T. Chen's group, we can estimate that $D_T \simeq 30\,\mathrm{cm^2\,s^{-1}}$ and assess the observable impact of this factor on the data in LZ. As can be seen from Figure 5.3, the maximum expected spread for $D_T = 30\,\mathrm{cm^2\,s^{-1}}$ from the bottom of the detector is $2.5\,\mathrm{mm}$ which is much smaller than the $32\,\mathrm{mm}$ radius of the PMT faces used in the TPC [87]. For this reason, no amount of transverse diffusion within LZ will be observable by the detector and hence will have no impact on position reconstruction. Additionally, transverse diffusion has no significant impact on the timing of drifting electrons. For a maximum deviation of $2.5\,\mathrm{mm}$ after drifting the $1461\,\mathrm{mm}$ from the cathode to the liquid surface, an additional $\sim$1.53 ns is added to the $1050\,\mu\mathrm{s}$ drift time, a negligible contribution. Hence, a value of D_T can be taken as $30\,\mathrm{cm^2\,s^{-1}}$ with no considerable consequences.

Longitudinal diffusion, on the other hand, causes a spread of the electrons in the drift axis and hence in time. This increases the time during which EL photons are produced in the gas, thereby increasing the widths of the S2 pulses. The farther the electrons drift, the larger the diffusion impact becomes. This can be seen in the data in Figure 5.4. The total S2 time variance determines the shape of this relationship, σ^2_{S2}, which can be expressed as the convolution between the electron diffusion time and the SE width and hence can be measured from data by fitting the following

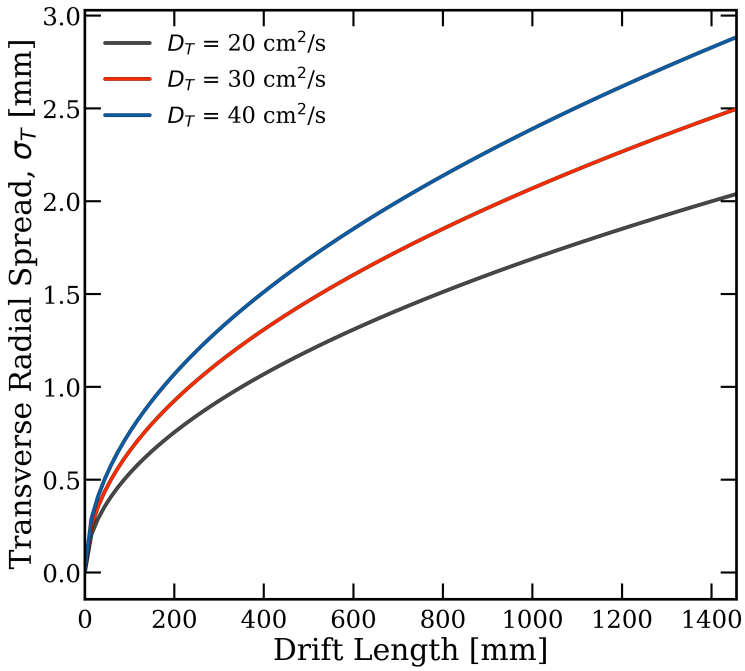


Figure 5.3: Transverse radial spread of an electron cloud as a function of the drift length through the TPC. Three example transverse diffusion coefficients are simulated ranging from 20–40 $\mathrm{cm}^2\,\mathrm{s}^{-1}$ to encompass the uncertainty on D_T.

function to this parameter space,

$$\sigma_{S2} = \sqrt{\sigma_D^2 + \sigma_{SE}^2} = \sqrt{\frac{2D_L t}{v_d^2} + \sigma_{SE}^2}, \tag{5.3}$$

where σ_D is the time variance due to diffusion and σ_{SE} is the time variance from the width of the SE pulse [88]. To translate Equation 5.3 to the observable parameters, AFTs, we must assert that all S2 pulses are Gaussian in nature, which holds by the Central Limit Theorem for large numbers of S2 photons. We may then observe the relationship

$$\xi = \int_{-\infty}^{AFT_\xi} \mathcal{G}(t, \mu = 0, \sigma_{S2})dt = \frac{1}{2}\mathrm{erf}\left(\frac{AFT_\xi}{\sqrt{2}\sigma_{S2}}\right) + \frac{1}{2}, \tag{5.4}$$

where ξ is the area fraction and AFT_ξ denotes the area fraction time at that ξ. This can be rearranged to provide an approximation for each AFT_ξ as,

$$AFT_\xi = \sqrt{2}\sigma_{S2}\ \mathrm{erf}^{-1}(2\xi - 1). \tag{5.5}$$

Equation 5.6 provides a direct relationship between the AFT widths and the longitudinal diffusion

coefficient, D_L, allowing for direct measurements of D_L from Figure 5.4. AFT widths can then be derived since

$$\begin{aligned} AFT_{\xi_1} - AFT_{\xi_2} &= \sqrt{2}\sigma_{S2}\text{erf}^{-1}(2\xi_1 - 1) - \sqrt{2}\sigma_{S2}\text{erf}^{-1}(2\xi_2 - 1), \\ &= 2\sqrt{2}\,|\text{erf}^{-1}(2\xi - 1)|\sqrt{\frac{2D_L t}{v_d^2} + \sigma_{SE}^2}, \end{aligned} \tag{5.6}$$

which holds for symmetric widths such as AFT75 - AFT25 due to the symmetry of the inverse error function.

Physically, it is understood that the number of electrons in the initial S2 electron cloud has no impact on the diffusion since the electromagnetic interactions will be negligible in comparison to the force applied by the electric field in the liquid xenon. Therefore, the diffusion coefficient is measurable from any sufficiently clean dataset, meaning it contains minimal abstract topologies that might bias the fit. In this chapter, we opt for a comparison with the tritiated methane, CH_3T, calibration source since it provides high-statistics, low-energy beta that can be easily replicated in the simulation. Initial measurements are made by fitting to this data. The result of this is shown in Figure 5.4 to yield $D_L = 24.28 \pm 0.06\ \text{cm}^2/\text{s}$.

With the longitudinal diffusion coefficient determined from calibration data, we now discuss how diffusion is applied in practice. As discussed previously, the drift time of each electron from its starting position is sampled from the drift maps derived using LZEF. The change in drift time due to diffusion is added to this value by sampling the Gaussian spreading such that the total time travelled, t, is calculated via

$$t = t_0 + \frac{1}{v_d}\mathcal{G}\left(0, \sqrt{2D_L t_0}\right), \tag{5.7}$$

where t_0 is the drift-map-sampled drift time and all other symbols have been defined previously.

The longitudinal diffusion is controlled by the electric field in the liquid xenon. Therefore, a field-dependent model can be devised to predict D_L across different field configurations, thereby preserving the simulations' predictive power. Iterations of these models within LZ have been constructed by collaborators, in addition to the NEST global average model, as shown in Figure 5.5. These models all overpredict the diffusion coefficient and, when tested, produced S2s with broader widths than are physically observed in data. The initial concern was regarding the implementation

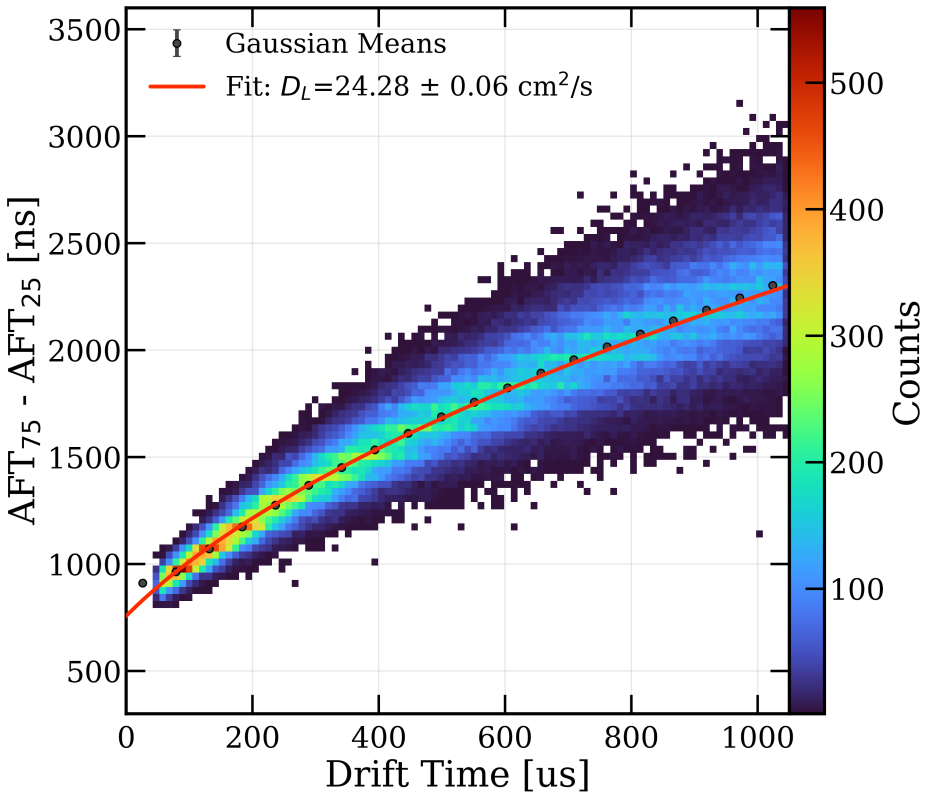


Figure 5.4: Measured fit from CH_3T calibration data. For this fit, only D_L is allowed to float and $v_d = 1.4\,\mathrm{mm\,\mu s^{-1}}$ and $\sigma_{SE} = 562\,\mathrm{ns}$ remain fixed. A drift time cut is applied, removing $0\,\mu\mathrm{s}$ to $60\,\mu\mathrm{s}$ to remove events occurring in the gaseous xenon since they do not obey the same diffusion relationship.

of the diffusion model in the code. To bypass this, a macro command was developed allowing the user to specify the desired value for D_L. Upon testing the nominal value of $D_L = 24.3\,\mathrm{cm^2\,s^{-1}}$, it was observed that this produced larger S2 widths than observed in data, depicted in Figure 5.7. A manual iterative tuning of this parameter found that $D_L = 22\,\mathrm{cm^2\,s^{-1}}$ best agreed with the CH_3T calibration data also shown in Figure 5.7. Furthermore, when measuring the diffusion coefficient from the output of the simulation with the method described previously, the result is consistent with the measured value of $D_L = 24\,\mathrm{cm^2\,s^{-1}}$, implying the implementation in the simulation is viable but the measurement method may be flawed, or there is a hidden effect elsewhere in the simulation biasing the widths of the S2s.

The principal assumption in the measurement method, which may not hold at low energy, is that the S2 pulse is entirely Gaussian. To investigate this, the fitting procedure is repeated for a collection of MeV gamma peaks from LZ background data, along with two low-energy gamma calibration sources, namely, ^{83m}Kr and ^{131m}Xe. The results of these measurements can be observed in Fig-

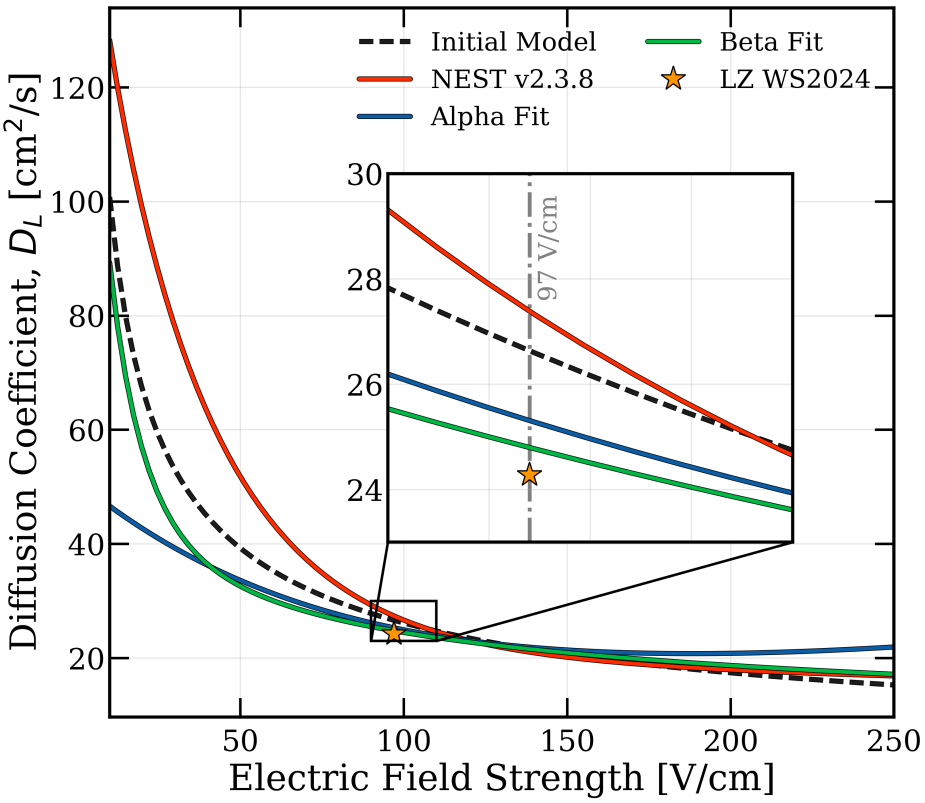


Figure 5.5: Various models for the longitudinal diffusion coefficient as a function of field strength in the liquid xenon. Initial Model refers to the model present in the simulation prior to this tuning; the alpha fit is a model derived from fitting to alpha peaks in the detector, and similarly, the beta fit refers to modelling performed on a beta spectrum. Both the alpha and beta measurements shown are derived by collaborators external to this thesis work. The NEST model is also provided here as a reference for a global fit to the latest measurements in the field. The LZ WS2024 result from Figure 5.4 is shown as a single data point derived by the previous fitting method from a CH_3T calibration.

ure 5.6 and indicate some energy dependence. This is due to the topology of the events. As energy approaches the MeV range, event energy depositions will have longer track lengths of $\sim 9\,\mathrm{mm}$, undermining the pointwise-spreading assumption on which the diffusion model is based. In the extreme example, with a gamma travelling directly downward through the xenon, an electron travelling an additional $9\,\mathrm{mm}$ can introduce a time delay of $6.43\,\mu\mathrm{s}$, comparable to the full width of S2s originating from the cathode. The track length increases with energy, effectively increasing the implied D_L value from these measurements. However, at low energies, the track length is point-like. These energies still disagree with the tuned value of $D_L = 22\,\mathrm{cm}^2\,\mathrm{s}^{-1}$, which implies there may also be some reconstruction bias from LZAP's ability to identify pulse start and end times – inherently biasing the width measurements. Alternatively, the switch from beta to gamma Q_y

response also directly affects the pulse width, with more electrons producing larger pulses. This can therefore be represented with $D_L(E) = D_{L,0} + f(E) + \delta_{reco}$ where $D_{L,0}$ is the true underlying diffusion coefficient, $f(E)$ is some energy dependent function which accounts for the topological differences in the events at different energies, and δ_{reco} is the bias introduced by reconstruction or yields effects.

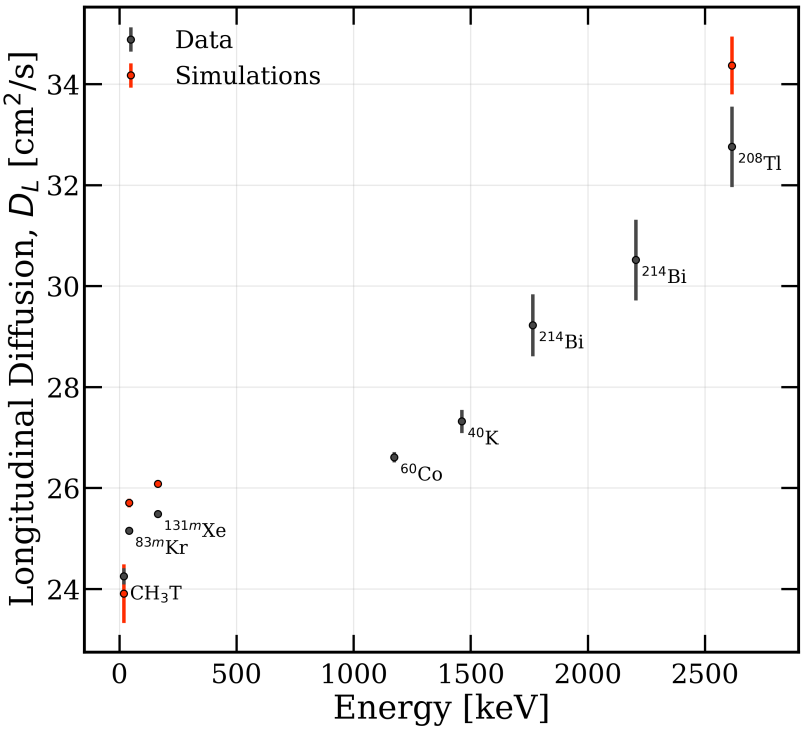


Figure 5.6: Measurements of the longitudinal diffusion coefficient from different energy calibration sources. The measurements from simulation are also shown with agreement with the general energy trend despite having a fixed input diffusion coefficient of $22\,\mathrm{cm^2\,s^{-1}}$.

To confirm this hypothesis, a handful of these peaks were simulated, which should also produce the same energy-dependent track-length effects and reconstruction bias. Figure 5.6 shows the level of agreement between simulation and data measurements. Despite all simulations having a fixed input of $D_L = 22\,\mathrm{cm^2\,s^{-1}}$, the simulation and data exhibit the same general trend across all energies. One way to minimise the energy-dependent effects is to opt for high-energy alpha decays, since they are still point-like and render the $f(E)$ term negligible. This method was carried out for alphas by a collaborator[1] and measured $D_L = 24\,\mathrm{cm^2\,s^{-1}}$, as observed for low-energy CH_3T decays. Both of these results, however, may be subject to some $\delta_{reco} \simeq 2\,\mathrm{cm^2\,s^{-1}}$ bias from the limits of the LZAP reconstruction. This correction factor is extremely difficult to measure without a

[1] Jacquet, E. *Thesis in progress.*

strong waveform simulation to tune upon. However, this does not entail a new measurement of the diffusion coefficient for LZ since it cannot be verified that δ_{reco} is not a bias provided somewhere in the simulation. Nevertheless, $D_L = 22\,\mathrm{cm^2\,s^{-1}}$ is taken as the effective simulation input to reproduce the accurate reconstructed widths and we note that this discrepancy remains unresolved.

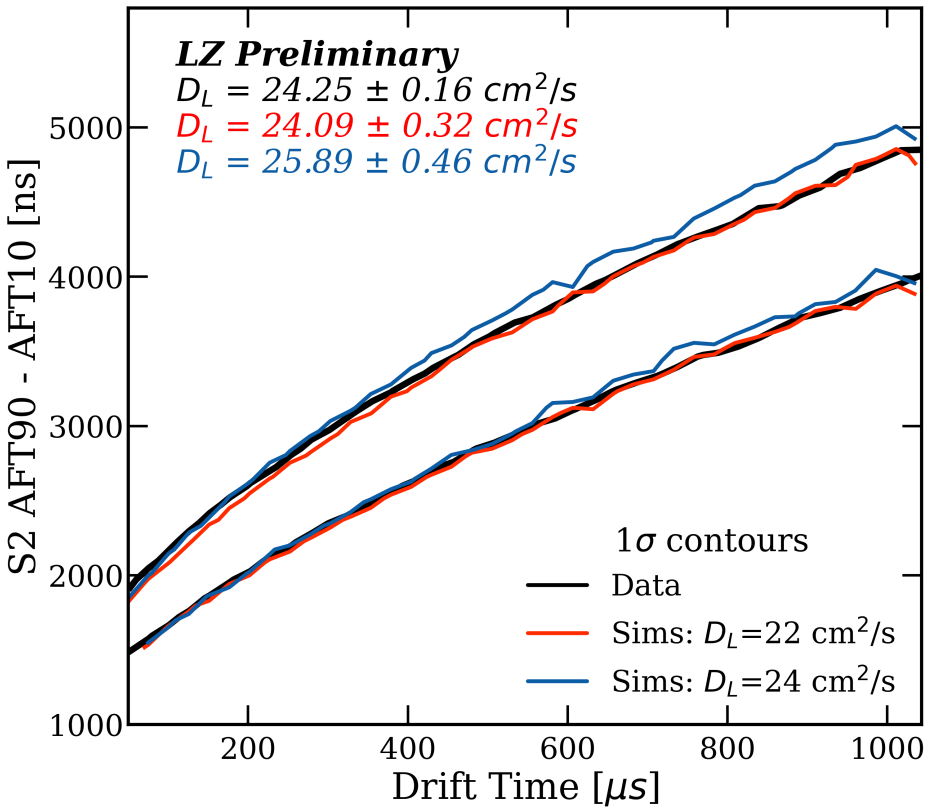


Figure 5.7: Comparison of diffusion properties from a CH_3T calibration source (black contours) and two runs of the simulation. The blue contours show the outcome of simulating with the measured coefficient from the data, indicating an overestimation of the widths. The red contours show the result obtained with the tuned coefficient. The diffusion coefficients measured from this plot are given at the top for each corresponding curve.

An identical check was made for the WS2022 detector configuration, which had a stronger electric field of $197\,\mathrm{V\,cm^{-1}}$, which determined that $D_L = 18\,\mathrm{cm^2\,s^{-1}}$ was the optimal value. Ensuring the diffusion output is correct is crucial for the studies in Chapter 6. The final comparison displaying the full agreement at all drift times for a CH_3T source is shown in Figure 5.7.

5.2.2 Electron Lifetime

Along the electron's drift path, electrons may be captured by electronegative impurities in liquid xenon. This effect is controlled in LZ by continuously removing the electronegative impurities

using a heated zirconium getter and by maintaining constant xenon circulation in the gas phase [31]. These impurities are required to have a concentration below $0.1\,\mathrm{ppb}$ to preserve signals from the cathode [40]. Reabsorption of the ionised electrons causes the S2 signal to be smaller than expected for a given energy deposition in the detector. We must monitor the electron lifetime – the mean time for which an electron drifts through the liquid xenon before being reabsorbed by impurities. The high purity of xenon in LZ allows for long electron lifetimes of $\sim 15\,\mathrm{ms}$, 15 times larger than the maximum drift time of the detector for the WS2024 detector configuration. Hence, around 6% of electrons are expected to be lost along the drift length, which must be accounted for within the simulation with the following model.

The electron lifetime is modelled in LZ with a decaying exponential probability. If an electron cloud is produced with a total charge, Q_0, it will decay exponentially with time as,

$$Q(t) = Q_0 e^{-\frac{t}{\tau_e}},$$

where τ_e is the electron lifetime. It is often easier in simulation to represent this in terms of physical distance travelled; this can be easily performed with a linear translation $d = v_d t$ and $D_A = v_d \tau_e$, where $v_d = 1.4\,\mathrm{mm\,\mu s^{-1}}$ is the mean drift velocity of the electrons in the liquid xenon and D_A is the attenuation distance of the electron path. Dividing this new translated expression by Q_0 gives the probability of an electron surviving the drift distance and hence can be expressed as

$$P_{surv} = e^{-\frac{d}{D_A}}. \tag{5.8}$$

This can therefore be implemented in a Monte Carlo algorithm to randomly determine whether a given electron survives the drift to the detectors' extraction region. The effects of electron lifetime can be seen in the data by examining the relationship between S2 pulse area and drift time. Longer drift times have a higher probability of attaching to the electronegative impurities and will therefore produce smaller S2s. Given that the electron lifetime is much longer than the maximum drift time of the detector, a small but non-negligible loss of pulse area is observed with drift time. In the simulation, the default value of the attenuation length is $D_A = 1\,\mathrm{m}$, which causes a 77% decrease in S2 area for an electron travelling the maximum length of the detector, demonstrated in Figure 5.8.

Correcting for this is simple, due to the constant monitoring of the electron lifetime in LZ, from

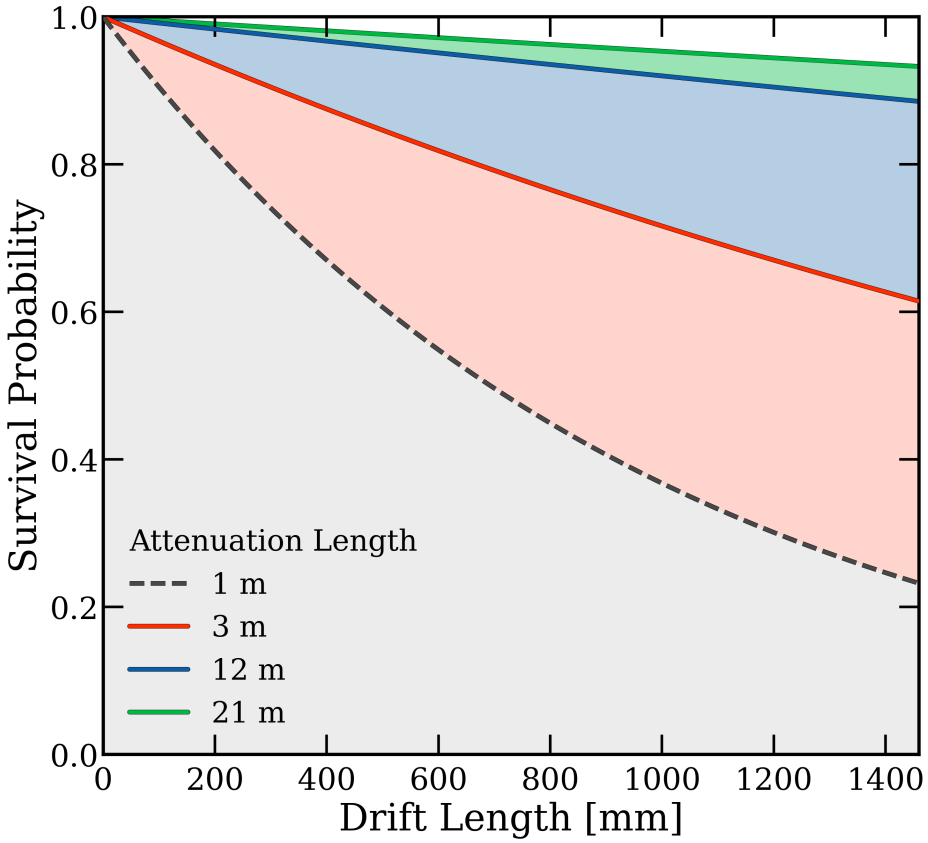


Figure 5.8: Survival probability for various values of the attenuation length used in Equation 5.8. The area of the shaded region under each curve is representative of the survival proportion of electrons. The initial default value in the simulation was $1\,\mathrm{m}$.

which we calculate $D_A = 1.4\,\mathrm{mm\,\mu s^{-1}} \times 15\,000\,\mu\mathrm{s} = 21\,\mathrm{m}$. Implementing this new value corrects the distribution shown in Figure 5.9, noting that the electron lifetime effects account only for the slope of the line, not the absolute magnitude of the S2.

Furthermore, there also exists a probability that the electrons close to the walls of the TPC can be attached to the dielectric PTFE surface [89]. LZ has observed a considerable amount of charge build-up on the detector walls, indicating this is a non-negligible effect to consider [63]. In addition to removing electrons from the S2 pulse, the build-up of charge on the PTFE surfaces impacts the shape of the electric field near the walls. This makes it necessary to include the wall-attachment probability in the field and drift maps from LZEF. With these probabilities predetermined, one must sample the values from the drift maps to determine if electrons near the edges of the detector become attached to the PTFE and are removed from the drifting electron cloud, which constitutes the S2 pulse.

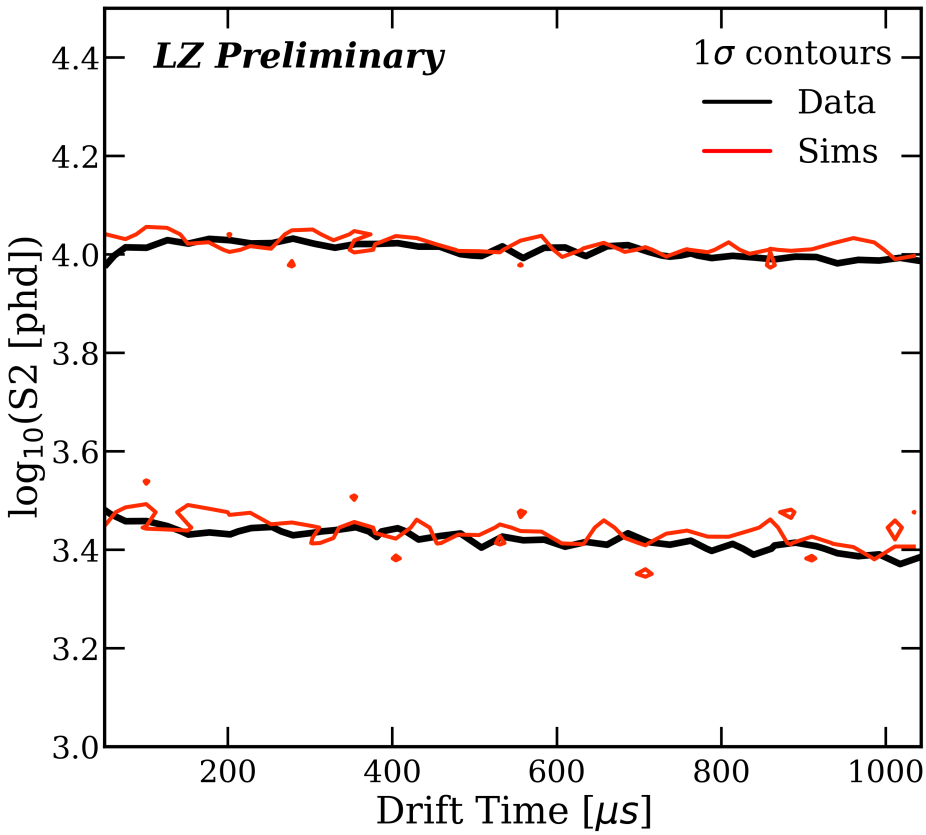


Figure 5.9: 1σ contour comparison of real and simulated CH_3T S2 size with drift time, demonstrating the agreement due to the electron lifetime modelling.

5.2.3 Extraction Efficiency

Electrons that reach the liquid surface must overcome the potential barrier at the liquid-gas boundary to be extracted into the gaseous xenon, where electroluminescence occurs. The extraction region of LZ is specifically designed to maximise the efficiency of electron extraction into the gas. This is required to preserve the maximum amount of the ionisation signal from the S2, similar to maximising electron lifetime.

In the absence of an electric field, the electron approaching the liquid surface induces a polarisation in the liquid xenon. Following Sorenson, the image potential can be described by the Schottky barrier

$$V_{image}(z) = \frac{e^2}{8\pi\epsilon_0 z}\frac{(\varepsilon_L - 1)}{(\varepsilon_L + 1)}, \tag{5.9}$$

where $\varepsilon_L = 1.96$ is the dielectric constant for LXe, e is the electron charge and z is the distance from the liquid surface [90, 91]. This represents the magnitude of the barrier in the absence of

an electric field. Physically, this is an attraction of the electron to the liquid, indicating that the electron is in a confining potential well. Introducing the electric field contributes an additional term to the total potential $V(z)$, written as

$$V(z) = V_{image}(z) + eE_{liq}z, \tag{5.10}$$

where E_{liq} is the electric field in the liquid defined in Equation 4.3. The inclusion of the electric field both decreases the potential barrier, since e is negative, and increases the electrons' energy. Overall, this increases the probability that electrons escape the liquid surface. The larger the field applied, the easier it is for electrons to overcome this barrier. This is depicted in Figure 5.10.

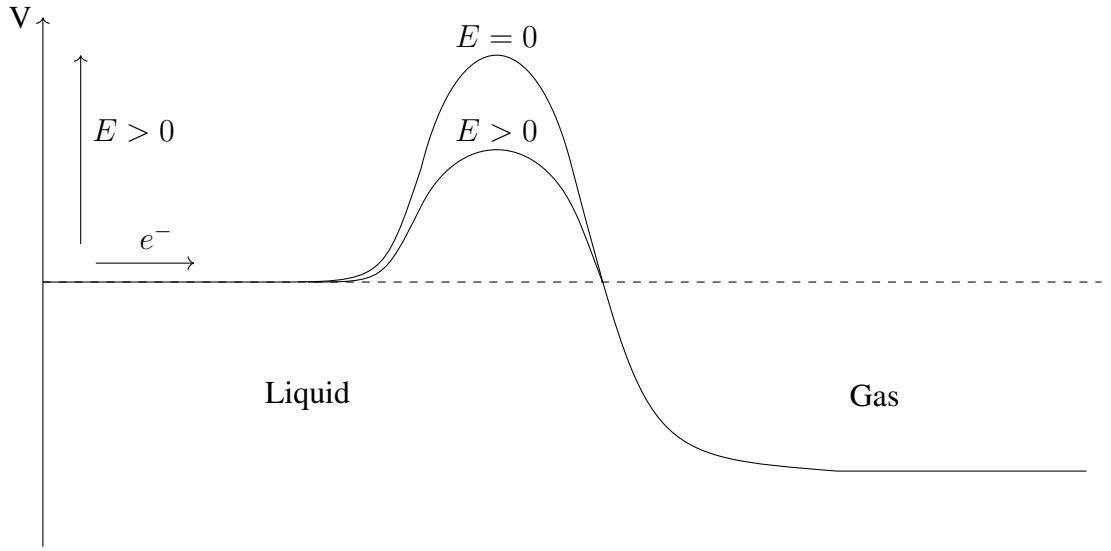


Figure 5.10: Simplified diagram showing the effect of increased electric field on the potential barrier from liquid to gaseous xenon. The stronger the field strength, the lower the barrier height and the higher the electron extraction efficiency.

The energy required to extract electrons at this boundary depends on the electric field strength in the liquid and its temperature. The temperature directly contributes to the thermal energy of the electrons and can increase the probability of extraction, similar to how the electric field increases the electron's kinetic energy. Many experiments have measured their extraction efficiency, though their methods of individual measurements may vary, which allows NEST to perform a global fit to these measurements and provide an empirical function for extraction efficiency:

$$\varepsilon_{ext}(E_{liq}, T) = 1 - A_1(T)\exp(-A_2(T) \cdot E_{liq}^{A_3(T)}) \tag{5.11}$$

where ε_{ext} is the probability of extracting an electron, E_{liq} is the electric field strength in the liquid xenon and T is the temperature. The coefficients, A_i for $i = 1, 2, 3$, are all temperature-dependent polynomial functions. For LZ's operating temperature, $T = 178.7\,\mathrm{K}$, these coefficients are $A_1(178.7) = 5.0 \times 10^3$, $A_2(178.7) = -6.9 \times 10^4$ and $A_3(178.7) = 1.5 \times 10^5$. The magnitude of A_3 further exemplifies the strength of the electric field dependence as well as the additional temperature dependence in Equation 5.11.

The NEST model was chosen over the default LZ model,

$$\varepsilon_{ext}(E_{liq}) = -0.038E_{liq}^2 + 0.53E_{liq} - 0.85, \tag{5.12}$$

which is dependent only on the electric field strength in the liquid. This model was tuned to the fixed operating temperature of the WS2022 detector configuration, which operates at $177.9\,\mathrm{K}$. Given the sensitivity to operating temperature observed in Equation 5.11, small changes here ($\sim$1 K) and other possible changes in the future cause this model to lose its predictive power and hence are abandoned for the NEST model. We therefore take the value of E_{liq} in Equation 5.11 to be $\sim 3.5\,\mathrm{kV\,cm^{-1}}$, which produces the NEST extraction efficiency curve shown in Figure 5.11.

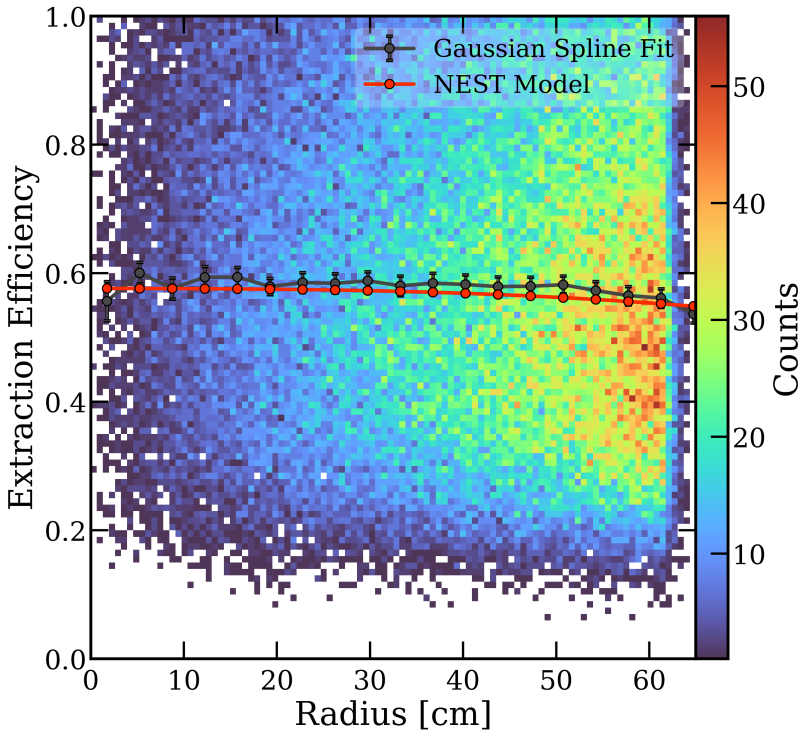


Figure 5.11: Measured extraction efficiency curve from CH_3T calibration data compared with the accepted NEST model.

Figure 5.11 also shows the comparison to the measurements in LZ from CH_3T calibration data. This measurement is carried out by first establishing that the size of an S2 can be written as,

$$S2 = N_e \times e^{-t/\tau_e} \times \varepsilon_{ext} \times \langle SE \rangle, \tag{5.13}$$

where N_e is the number of electrons ionised by the initial interaction and all other symbols have their previously defined meanings. Therefore, the extraction efficiency can be expressed as,

$$\varepsilon_{ext} = \frac{S2 \times e^{t/\tau_e}}{N_e \times \langle SE \rangle}. \tag{5.14}$$

The mean SE pulse area in each radial bin is found by fitting the SE pulse area distribution to a Gaussian. The S2 size and drift time, t, are known for each event by definition of the SS criteria. The true number of ionised electrons is unknown in the data, so here the value for each event is taken by sampling the distribution generated from the NEST beta yields model, which, for CH_3T, is $\mathcal{G}(156.5, 89.9)$. Since this is not known for each event, this increases the spread of the measured efficiencies in the heat map but will retain the same mean value, which is taken as the measurement for ε_{ext} in this case. Note that since this measurement requires the use of a NEST beta yields model, it cannot be classified as an independent measurement but does show valuable consistency with the expectation from these models. Figure 5.11 shows the agreement between the measured data and the NEST model, thereby affirming confidence in this improvement.

5.3 Afterpulsing

At this stage in the tuning process, the electron clouds generated for each event are now correctly transported to the liquid surface, accounting for drift times, diffusion and charge loss along the path. With the improvements to the extraction efficiency model, and the finalised tuning from Chapter 4, the S2 is constructed with significantly improved physical modelling. During the process of validating these newly constructed S2s, it was observed that the position reconstruction showed large overdensities in the (x, y) plane, shown in Figure 5.12. These reconstructed positions are calculated using a package called Mercury, which uses the PMT hit map for a given S2 pulse to reconstruct its position based on the distribution of light across the top-array PMTs [92]. The photons generated by the extracted electrons will predominantly be detected by the nearest

PMTs, allowing for accurate localisation of the events in the (x, y) plane.

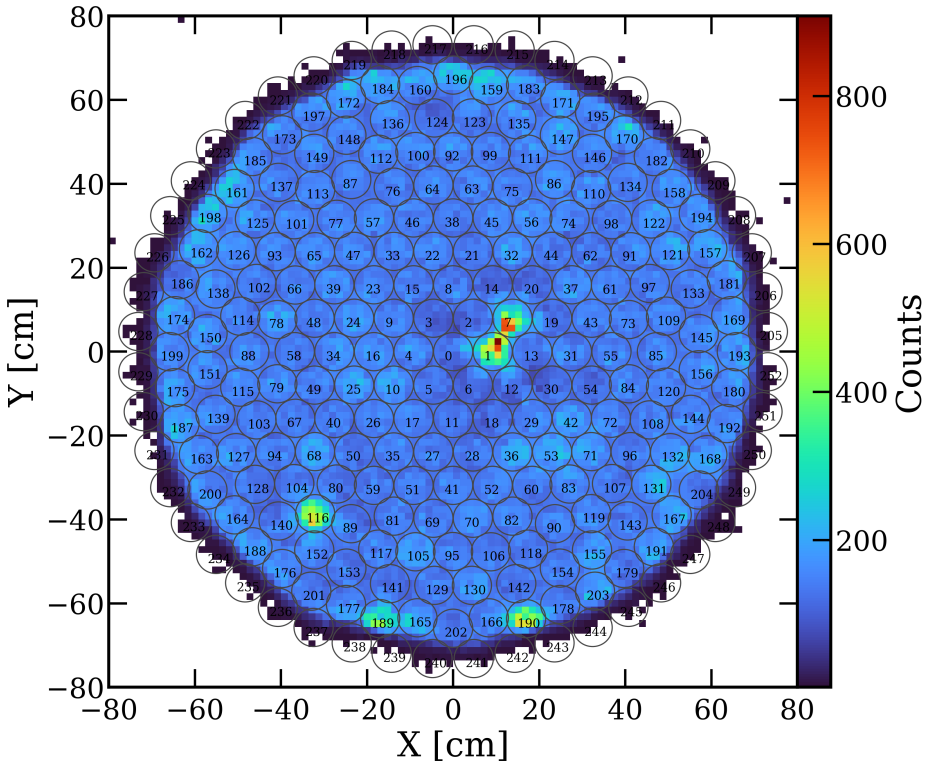


Figure 5.12: Heat map of the reconstructed positions of simulated S2 pulses. Despite the true position distribution being flat, notable areas of higher density arise localised to PMTs with high afterpulsing rates.

When simulating a CH_3T source, the starting positions are uniformly sampled to create a uniform spatial distribution of events throughout the liquid xenon volume. This should result in a flat distribution of the reconstructed positions of all S2s produced. Nevertheless, multiple dense areas arise from these simulations. The positions of these overdensities coincide with PMTs known to have high afterpulsing rates. The afterpulsing effect, discussed in Section 3.1.2.2, has minimal impact on the size and width of the S2 since the size and width of the afterpulses are negligible in comparison. However, the large quantity of S2 photons results in large quantities of afterpulses within these PMTs, which therefore appear to observe more light, relative to the adjacent PMTs. Hence, they bias the reconstruction algorithm towards them.

Despite the real PMTs having large afterpulsing rates, the biasing of their reconstructed positions is not observed in the calibration data, only in simulation. This indicates an inaccurate modelling of the PMT afterpulses in the DER. The afterpulse model in the DER consists of sampling the probability of an afterpulse occurring for each ion species, followed by sampling the size and time

delay of the afterpulse. The timing and average sizes of afterpulses for each species of impurity were derived from first principles and are shown in Table 5.1.

Species	Time [ns]	Number of Photoelectrons
H^+	280 ± 15	17.5 ± 5.3
H_2^+	390 ± 15	16.5 ± 5.0
He^+	520 ± 15	13.0 ± 4.3
CH_4^+	1010 ± 20	10.6 ± 3.9
Ne^+	1130 ± 30	8.0 ± 3.2
N_2^+	1330 ± 30	7.0 ± 3.0
Ar^+	1580 ± 30	4.3 ± 1.2
Xe^{++}	2020 ± 30	3.0 ± 1.9
Xe^+	2850 ± 50	2.0 ± 1.3

Table 5.1: Afterpulsing species in TPC PMTs and their initial expected timings and number of photoelectrons produced. The timings presented here should be understood as a time window around the mean as shown in Figure 5.13 [93]. The number of photoelectrons is presented as the mean and standard deviation, following a Gaussian distribution.

The mass of the species determines the time between the real SPE pulse and its associated afterpulse. LZ performs monthly LED calibrations within the TPC to calibrate the PMT responses and monitor the afterpulsing rates. It is important to monitor afterpulsing rates, as they indicate the health of the PMTs. If afterpulsing rates become too high, they can cause issues like those seen in the simulation and negatively impact the quality of LZ data; when this happens, these PMTs are turned off. During these LED calibrations, many individual pulses can be observed in each PMT, which are aligned in time so that the start of each pulse is taken as $t = 0$. The times derived for each species in Table 5.1 are then used to categorise the response region from each ion species in the residual gas, as shown in Figure 5.13, and measure the afterpulsing ratio (APR) for each impurity in the PMT. The APR for a given species is the ratio of the mean afterpulse area to the main pulse area. Additionally, from the compilation of multiple pulses, we can extract the afterpulsing probability (APP) that an initial pulse will produce an afterpulse of a given species. The

APP and APR values for each TPC PMT are documented and can be used directly by the DER, enabling a data-driven approach to this modelling. The probabilities measured from calibration are larger than the defaults in the early DER PMT parameter spreadsheets, indicating a decrease in the afterpulsing size is necessary to maintain a consistent afterpulsing ratio.

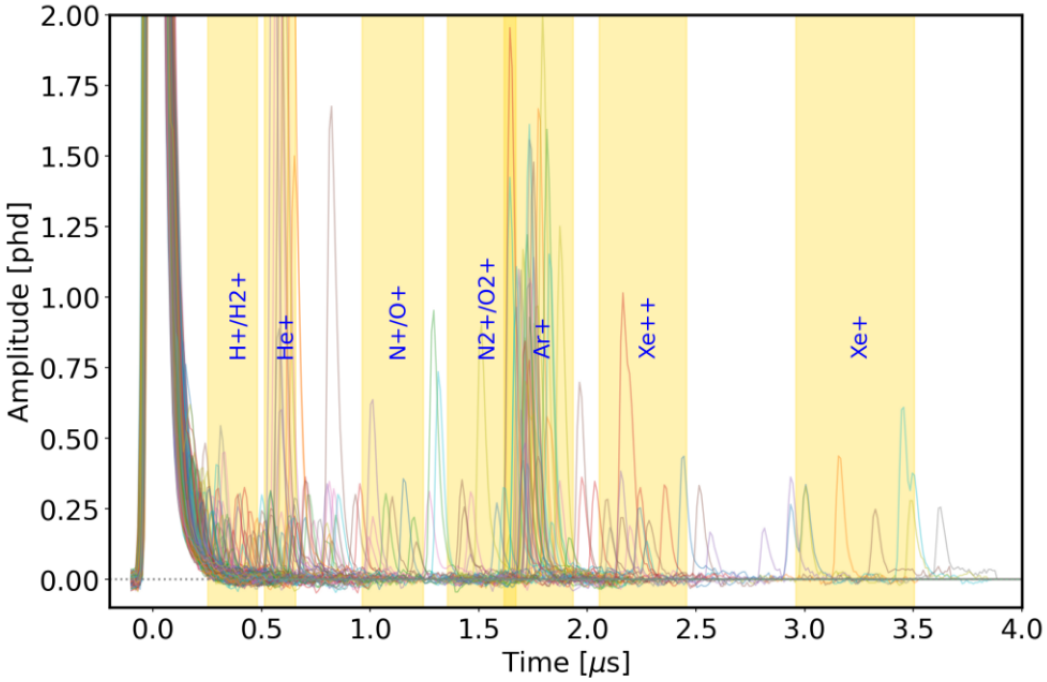


Figure 5.13: LED calibration data time-aligned to the beginning of the LED-induced S1 pulse, displaying the time windows for each afterpulsing species. All small pulses shown to follow the initial LED-induced S1 pulse are taken to be afterpulses. Figure from [93].

The DER requires the afterpulsing probability to determine whether an afterpulse is produced for a given SPE. It also requires the afterpulse size (APS) to know how large each afterpulse should be. The LED calibrations contain direct measurements of the APR and APP but not the average APS. This can be calculated with the useful relationship between APS, APP and APR,

$$APS = \frac{APR}{APP}. \tag{5.15}$$

By compiling several months of LED calibration data, the distribution of all afterpulse sizes for each ion species can be extracted. Truncated Gaussian functions are fit to these distributions, resulting in the expected distributions for the sizes of afterpulses to be sampled from in the DER. The means of the derived distributions are shown in Figure 5.14.

Comparing these values to those observed in Table 5.1, there is a large decrease in the size of the

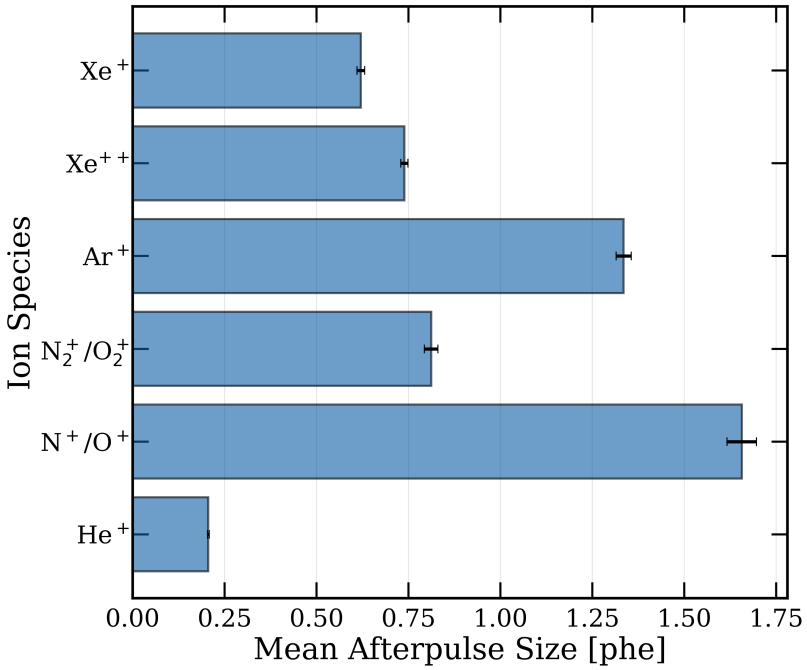


Figure 5.14: Mean afterpulsing size for each ion species simulated in the DER. The error bars show the error on the mean from the fitting procedure. Note that the afterpulsing size is measured in photoelectrons, phe.

expected afterpulses compared to those observed in PMT calibrations. This reduces the amount of light observed by the problematic PMTs and removes the overdensities at the (x, y) positions. With less light now observed in these PMTs, it serves as a good check to assess the impact this has on the S1 pulses since these are generally more sensitive to afterpulsing effects. In practice, afterpulses are expected to have negligible contributions to the S1 pulses and thus any noticeable changes to the S1s indicate further tuning is required. When investigating the new simulated data, a $\sim$20% decrease in the size of the S1 pulses is observed.

To identify the cause of this decrease, two simulated datasets were created to investigate the impact of afterpulsing on S1 size. The afterpulsing effect can be switched off in the DER in the run command; hence, for the two datasets, one included afterpulses and the other did not. The outcome of this comparison was that when afterpulsing was included, S1s were significantly smaller than when it was not. When quantifying the decrease, it was observed that it centred on the mean double photoelectron probability, 21%, as shown in Figure 5.15. This implies that the inclusion of the afterpulsing effect somehow removes the DPE effect, pointing to an error in the implementation of the DER logic.

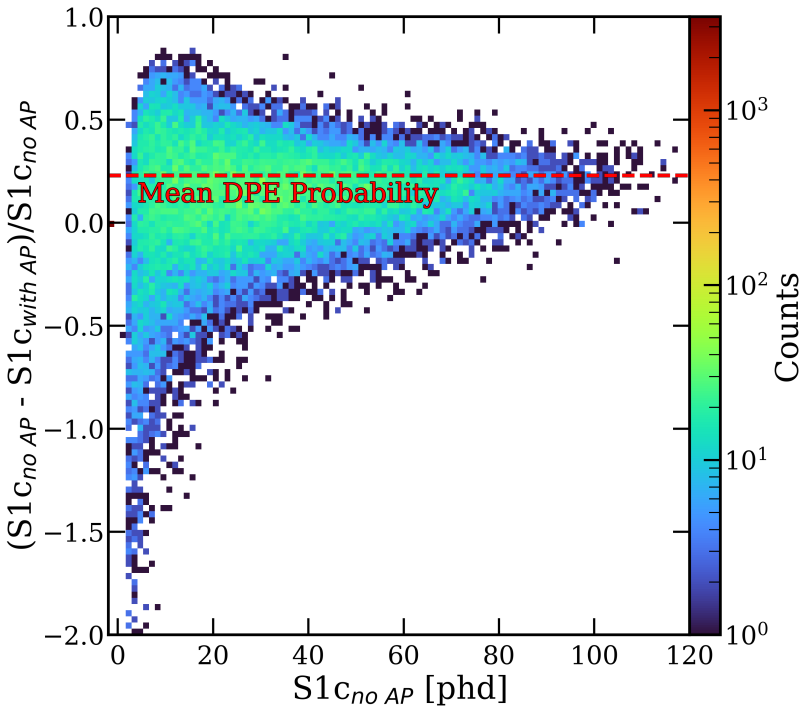


Figure 5.15: Investigation into the impact of applying afterpulsing effects on the S1 size for simulated CH_3T data. The red dashed line shows the mean double-photoelectron probability for PMTs in LZ, revealing a logical fault between the afterpulsing and double-photoelectron modelling.

Further investigation revealed that producing an afterpulse was the first check in a flow of conditional statements, each simulating a secondary PMT effect. The logic showed that whenever an SPE produced an afterpulse, it then had no possibility of exhibiting the DPE effect. This was only now observable due to the increased probability of afterpulsing, as directly measured in calibration. The logic in this implementation was corrected by separately determining if an afterpulse and double photoelectron are generated for each SPE.

5.4 ParticleBomb Generator

When simulating events with either the defined LZ calibration generators or the GEANT4 general particle source, much of what occurs in each event is subject to statistical fluctuations. While we can then select the subsample of events that meet our interests, in the case of rare effects this can lead to large amounts of wasted data and computing resources. To limit this, a general-purpose generator was devised which allows the user to run in one of four modes, depending on their requirements. In each of these modes, the number of quanta (photons or electrons) produced at

a given location is decided by the user in the macro, therefore completely avoiding the need for clustering and yields models, which have been shown to require further tuning at high energy.

S1 Mode

Randomly select a position within the LXe volume and emit a given number of photons from that position. These photons are propagated using GEANT4's stepping action as is done for all S1s in simulation. It should be noted here that S1s produced in this way are not subject to microphysical effects, such as singlet/triplet timing or electron-ion recombination, that occur in xenon interactions, and are approximations useful for light-collection efficiency studies. The output should include a single S1 waveform.

S2 Mode

Similarly, a position in the LXe volume is randomly selected, and a given number of electrons are released from the position. These follow the models outlined in this chapter. The output should include a single S2 waveform.

Double S2 Mode

This mode randomly generates an initial position and forms the S2 as in the S2 mode above. However, this mode also requires the user to define a radius from the initial interaction point. A second position is randomly sampled in the sphere defined by that radius, where a second S2 is produced. Currently, there is no implementation of a time delay between the two S2s and so they are produced simultaneously. This mode is primarily designed to test the time- and position-discrimination power of LZAP for nearby S2s, which is important for high-energy searches. The output should be two S2 waveforms.

Single Scatter Mode

The final mode produces both S1 light and S2 charge simultaneously from the randomly sampled position. The output should include a single S1 and a single S2 of defined sizes, such that these events will be classified by LZAP as single scatter events and the efficiency with which we classify these interactions can be measured.

To decrease computation times for S2s using this generator, all charge transport and electron life-

time effects are performed within the generator itself, decreasing the total amount of functions necessary. This allows electrons to be quickly generated and transported to the surface, where they are immediately killed by BACCARAT, and all extraction-region effects are simulated by the BaccMCTruth script. This allows large quantities of S2s to be generated with complete user control on a more reasonable timescale. For S2s less than $15\,\mathrm{electrons}$, 10^5 events can be completely simulated within $\sim 6\,\mathrm{h}$ on the NERSC Perlmutter supercomputer using just 2 CPU nodes [94], which would usually take $\sim 18\,\mathrm{h}$–$20\,\mathrm{h}$ for a low energy beta particle producing the same size S2s with S1s.

This generator was essential for the results presented in Chapters 6 and 7 due to the ease of producing large quantities of small S2s in regions where calibration proves difficult.

5.5 Calibration Validations

The final stage of this tuning, like the SE tuning, is to validate the final model outputs. This can be performed by comparison to various calibration sources within LZ. From several comparisons shown in this chapter, the models have been well-validated with a tritium source. Here, these comparisons are repeated for completeness to visualise the amalgamation of all tunings. This validation step serves to confirm which RQs from simulation can be trusted in analysis and which models require further work to improve.

5.5.1 Tritium

Since the Q-value for the tritium source is $18.6\,\mathrm{keV}$ [42], it has served throughout this chapter as a clean source for comparing models without introducing high-energy effects such as increased track lengths of high-energy gammas. As a result of this, the S2 pulse areas and widths are very well validated at this energy, as can be seen in Figure 5.16.

This final validation, however, does uncover small discrepancies that remain in RQs such as the TBA. Despite the TBA being well-tuned for SE response, the increased statistics from the larger S2 sizes introduce some level of disagreement, as seen in Figure 5.17. The figure also shows that the main source of the discrepancy comes from the bottom PMT array, while the top array remains in reasonable agreement. This is expected from Section 4.8, as it was noted that the LXe material

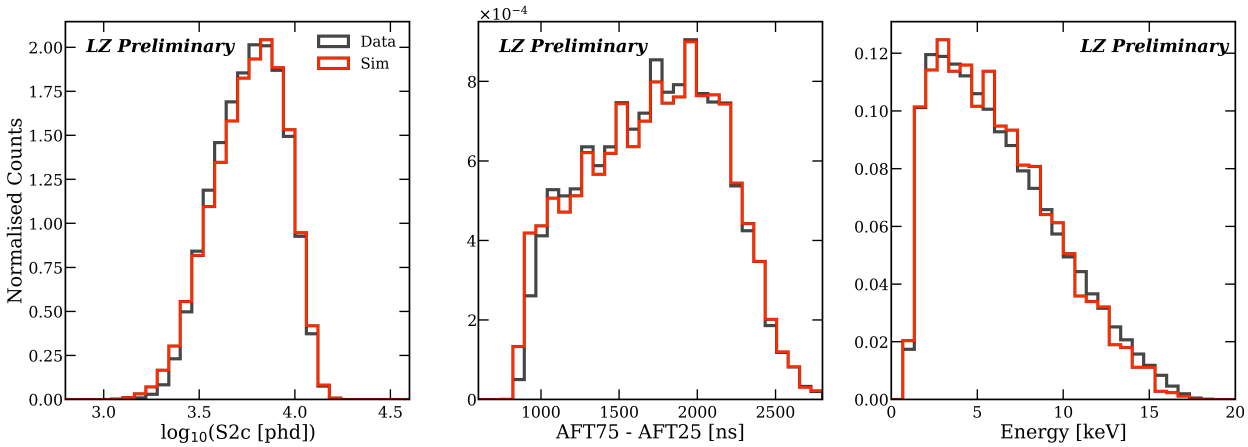


Figure 5.16: Direct comparison between simulation and calibration CH_3T data for the pulse area, AFT width and total reconstructed energy. These represent the core RQs used in the analysis and demonstrate the agreement between the simulation and the data.

parameters were not retuned to account for differences in the GXe parameters. In the future, a full optical model is required to resolve this discrepancy; until it is complete, the TBA parameter cannot be fully trusted in simulation. For low energies, the S2 TBA is mainly used to remove events that occur in the gaseous xenon with very high TBA values $\sim$1. The discrepancy observed is not significant enough to cause any of the simulated S2s to be removed by cuts in the TBA space. However, this discrepancy is detrimental to high-energy searches, where PMT saturation is frequently observed in the top array, leading to the frequent use of the bottom PMT arrays instead of combining both.

5.5.2 Neural Network

With LZAP producing over 100 RQs for each event, a direct quantitative comparison of all RQs is impractical during a tuning procedure. Instead, one can use a supervised classification neural network (NN). This approach is implemented by training on a subset of events labelled as either real or simulated data, then testing the ability of the NN to distinguish between the two populations. In principle, if a well-trained NN cannot distinguish between real and simulated data, this indicates a very strong simulation. Only RQs derived from S2s are tested in this implementation, since this is the purpose of this tuning. In future, this could be applied to RQs for all pulses. To supervise this learning a label of 0 is assigned to all real S2s and 1 to all simulated S2s from the CH_3T comparison shown previously. The test results are shown in Figure 5.18, with 64.6% accuracy. The optimal accuracy here is 50%, meaning the NN model cannot distinguish between simulation

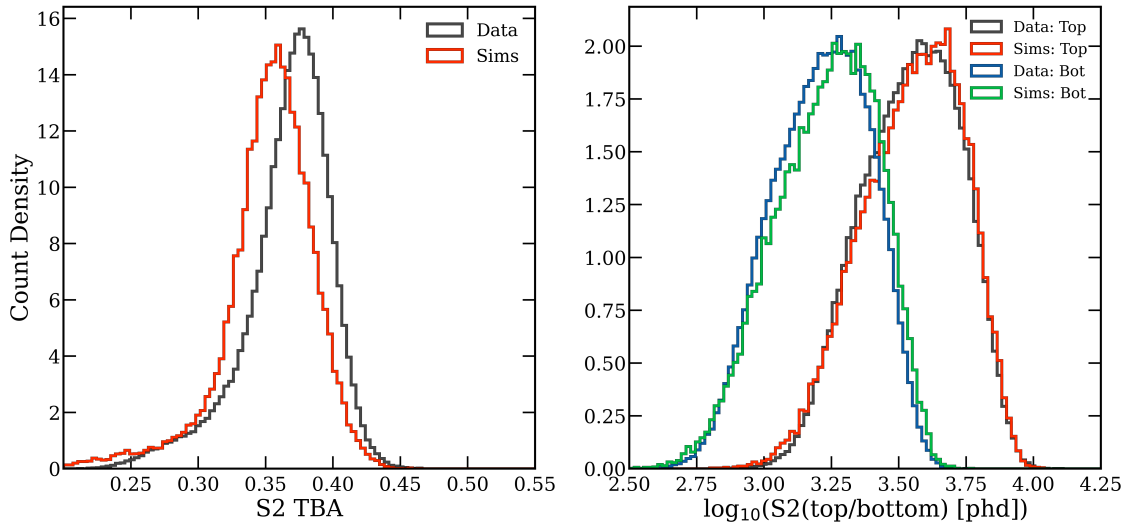


Figure 5.17: Top-bottom asymmetry comparison between simulation and calibration CH_3T data. The left-hand plot shows the total disagreement between the two populations, while the right-hand plot shows the pulse area observed in each PMT array. The main discrepancy arises from the bottom PMT array being overestimated by the simulation.

and data at all.

The additional benefit of using this method to validate the simulated S2s is that the NN can return the Permutation Feature Importance, which indicates the level of importance it assigns to each RQ for distinguishing between the two populations. In this instance, the larger the importance for an RQ, the weaker the RQ is in comparison to the data. Therefore, not only can this NN be used as a validation tool for the simulation but also a development tool with which weaknesses in the simulation can be quickly identified for future tunings, making this a powerful tool in the simulation tuning process. For this tuning, the Permutation Feature Importance is shown in Figure 5.19, showing that AFT99 has the largest importance of all features, followed by the TBA. The TBA discrepancy has already been discussed in Figure 5.17. The importance of AFT99 can be attributed to the lack of photoionisation and electron-train effects simulated in this production of the CH_3T data. These effects will only significantly impact the tails of the S2 shape, as is the case here compared with the AFT95 importance. These models are present in the simulation but still require tuning before they can be trusted.

Removing the AFT99 and TBA parameters from the training data reduces test accuracy from 64.6% to 50.9%, indicating that the NN has difficulty differentiating across most of the RQs used in the analysis. The distribution of the NN classification output is shown in Figure 5.20. This

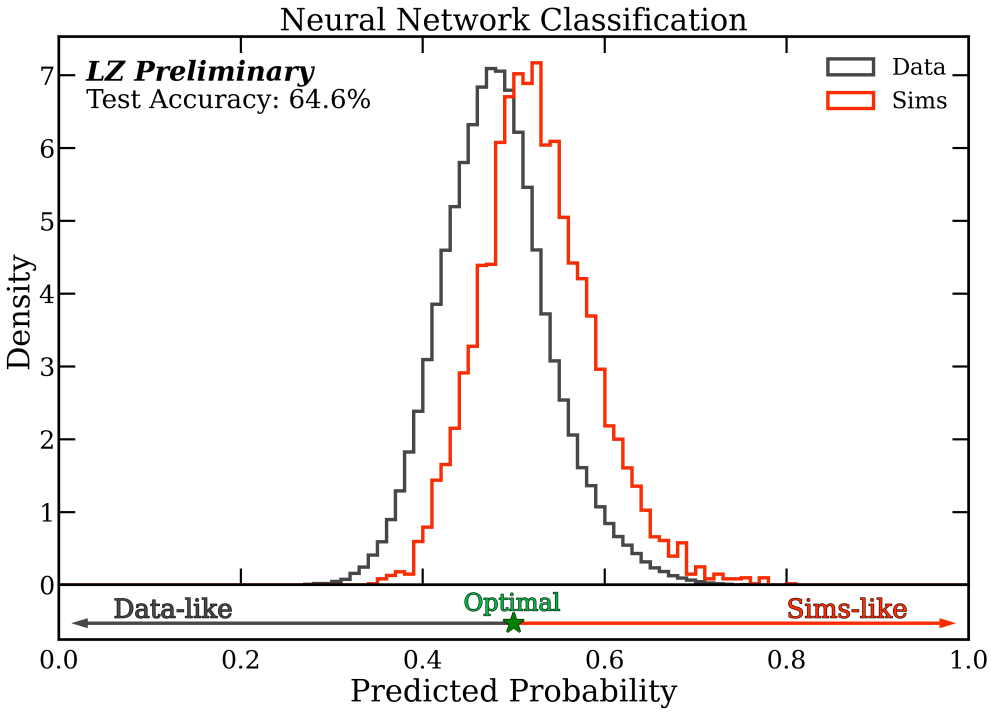


Figure 5.18: Neural network output for classifying a sample of testing data and simulation S2s. Test Accuracy here gives the percentage of correctly identified labels. For a perfect simulation, this should result in 50% accuracy. This plot demonstrates the model has difficulty distinguishing between real and simulated S2 pulses.

strongly indicates the strength of the foundation of the simulation for S2s with fine-tuning of the stated models left to future analysers.

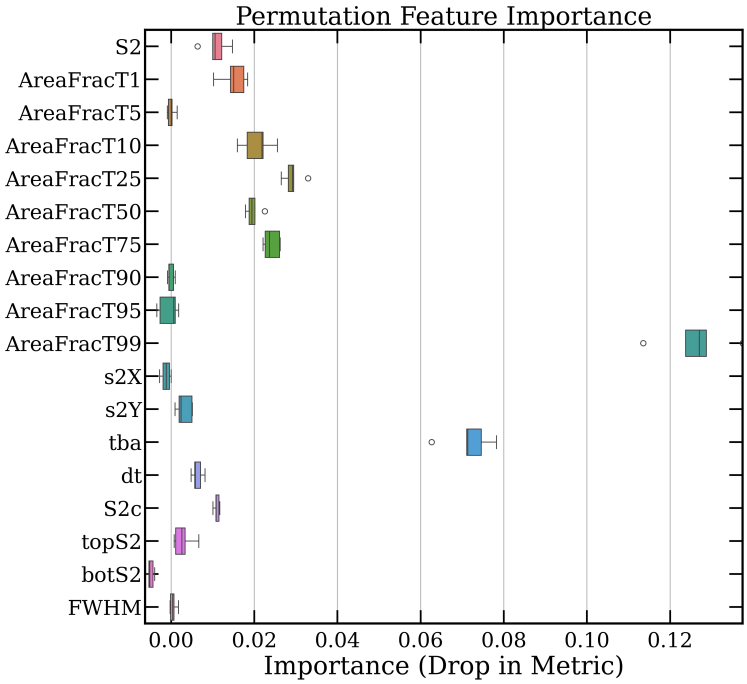


Figure 5.19: Permutation feature importance for all RQs used to train the classification NN. The most important RQs used to distinguish are AFT99 and TBA.

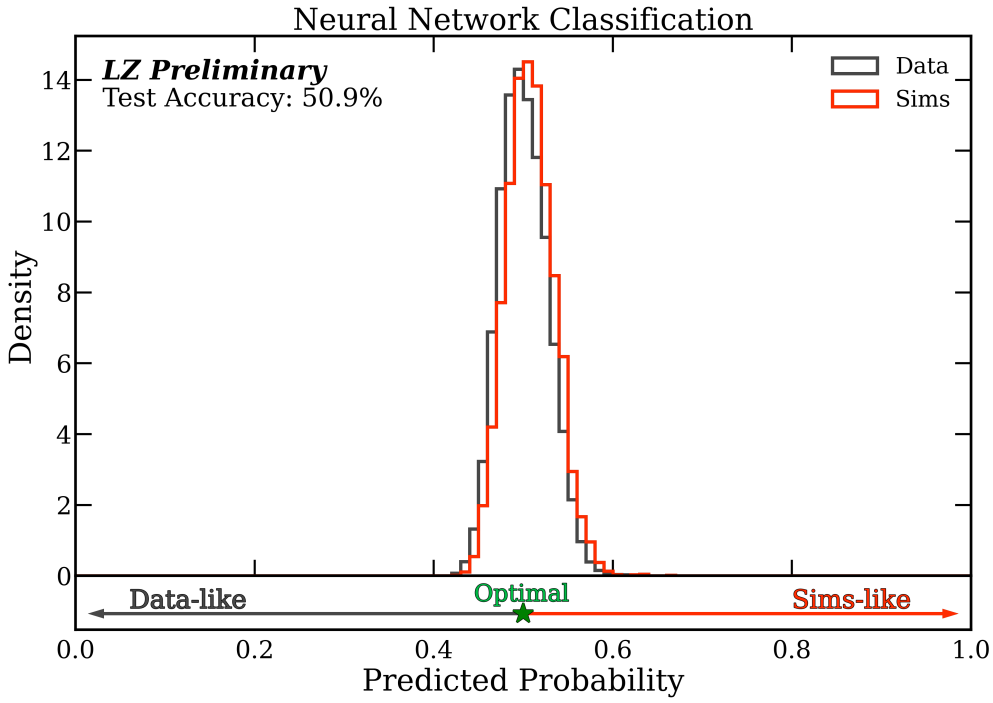


Figure 5.20: Impact of removing AFT99 and TBA parameters from the training data in Figure 5.18. These two RQs are known to be inaccurate or missing models in the simulation run, and further tuning work is left for future analysis.

5.5.3 High Energy

The primary purpose of tuning the simulations for this work is to produce accurate S2 pulses for the low-energy regime, demonstrated by the CH_3T comparisons discussed above. Most of the models tuned in this iteration of the simulation are independent of energy; hence, by tuning them to agree with the CH_3T calibration data, they should, in principle, also match at higher energies. However, as mentioned in Section 5.1, the lineage clustering and yields algorithm breaks down in the high-energy region. While this has little impact on the timing features of the S2, as can be seen from the agreement in measured diffusion coefficients in Figure 5.6, the yields modelling has a significant impact on the sizes of the S2 pulses. For the $0\nu\beta\beta$ decay searches, a dedicated tuning of the density-based clustering algorithm with a $0.1\,\mathrm{mm}$ neighbourhood radius and updated tuned β and γ yields models from ^{214}Bi spectra is completed. The deviation between data and simulation for the updated implementation of different gamma sources at various energies is observed in Figure 5.21.

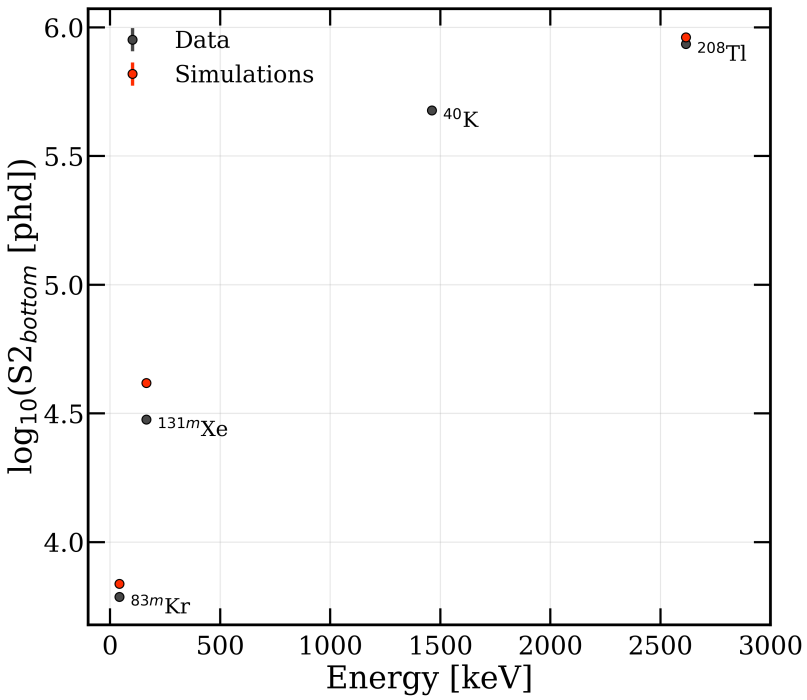


Figure 5.21: Mean S2 pulse area measured by the bottom PMT array for the labelled gamma calibration sources. Pulses in the bottom array are chosen to remove the PMT saturation effects observed in the top array PMTs for large S2s.

Figure 5.21 demonstrates much better agreement with this implementation at high energy, which then deviates at the lower energies. This discrepancy could be resolved by appropriately handling

the clustering and yield models. Work is ongoing within LZ and NEST to produce a universal algorithm which holds at all energies. Nevertheless, given that the results that follow concern S2s at threshold, interaction energies are on the order of $1\,\mathrm{keV}$, and the track lengths are limited to $\mathcal{O}(\mu\mathrm{m})$, they are thus reasonably point-like. The lineage clustering algorithm, in its current state, is sufficient to handle these interactions accurately, as evidenced by agreement with CH_3T calibrations.

At this stage, we have updated and validated the implementation of all physical models contributing towards the S2 response in the full chain simulation. The agreement observed throughout Section 5.5 is unprecedented within the full chain in LZ and is deemed acceptable to perform analyses with. Hence, having accumulated an in-depth understanding of the physical detector effects at play during the S2 formation, we are now in a strong position to apply this enhanced version of the simulation to study and tackle threshold detector effects.

6 | Fragmentation

The longitudinal diffusion of the ionisation electrons during their drift through the detector causes them to reach the liquid surface at different times. This increases the pulse width because the time at which each electron is extracted from the liquid surface is spread out. For large S2 pulses, there are sufficiently many electrons that the broadened pulse remains continuous in time. However, when approaching S2s of less than $15\,\mathrm{electrons}$ (the ROI for a low-mass WIMP, ^{8}B CEνNS and S2O searches), the electrons approach a more discrete distribution in time and hence, the diffusion can cause LZAP to fragment small S2s into multiple separate pulses, as shown in Figure 6.1. In this figure, we show a simulated S2 pulse from the ParticleBomb generator, in which all electrons are released from the same point in the liquid xenon. LZAP then classifies the solid line as a prominent S2, which will be used for analysis, and the dashed line as a separate pulse, which will be disregarded despite being roughly the same size and shape.

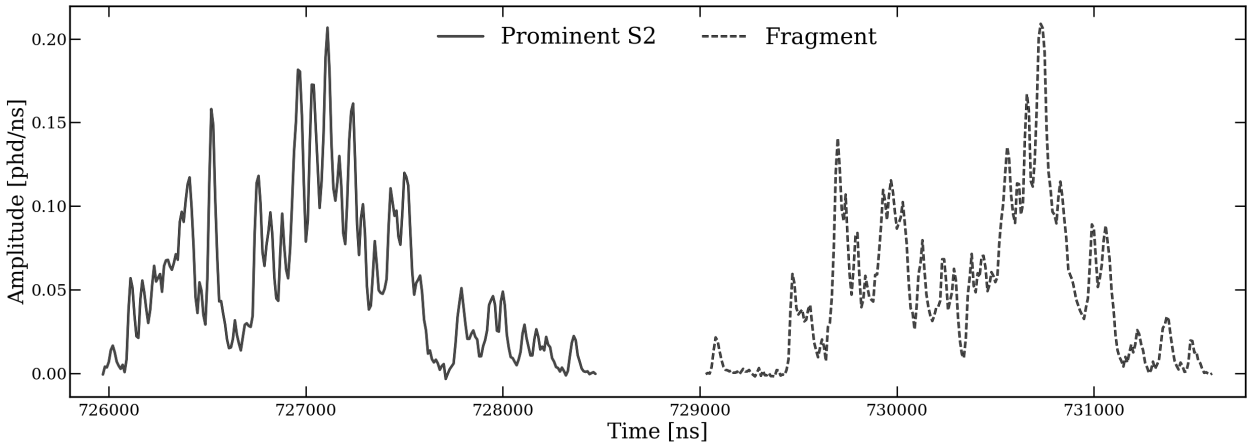


Figure 6.1: Simulated S2 waveform showing the fragmentation effect. Both pulses originate from the same position and time in the detector but are reconstructed as a single prominent S2 and a fragment. Fragments can be identified by LZAP as SE or S2 pulses.

In analysis, only the RQs taken from the prominent S2 are considered for each event. Therefore, any events that display this fragmentation effect are incorrectly reconstructed. This can lead to the loss of signal in two ways:

1. Mistaken multiple scatters,

2. Loss of pulse area and width.

To demonstrate these effects, we take the example of $6\,\mathrm{electrons}$ extracted into the gas. For Case

1, if the electrons are spread in time such that LZAP identifies two S2 pulses of $3\,\mathrm{electrons}$ each, LZAP would likely classify this event as a multiple scatter, so long as it has an S1 to accompany these S2s. As a result, this potential signal event is lost from our dataset (which should contain only single-scatter events) due to reconstruction inaccuracies in LZAP. For Case 2, if the $6\,\mathrm{electrons}$ are spread out such that only one of the electrons is removed from the prominent S2, LZAP would see a 5-electron S2 and an SE pulse. Only the S2 pulse is used for analysis, which is now assigned a smaller S2 size and width than in reality. Signal loss from S2s is not a foreign concept in LZ; signal is similarly lost from the electron lifetime and extraction efficiency effects. These losses are accepted in LZ because corrections are applied to account for them in analysis.

In this chapter, we introduce the concept of S2 fragmentation and use the full chain simulation to characterise its impact on the S2 pulses. We then offer three methods to handle its existence in the data. These approaches are: correct the signal modelling, cut fragmented events, or tag and reconstruct fragments. The following sections will discuss the impact of each of these approaches with examples from the ^{8}B CEνNS and, where appropriate, S2O searches.

6.1 Characterising Fragmentation

Characterising fragmentation effects in data is difficult because many physical processes can occur in the detector that mimic this event pathology. For example, an additional SE pulse could be produced within $2\,\mu\mathrm{s}$ of an S2 pulse by an S2 photon striking the gate grid wires and releasing an electron through the photoelectric effect. Alternatively, there is a constant SE background that persists in LZ, thought to be mostly from grid emission, which could coincide with an S2 pulse, thus producing an SE pulse within a similar time window. Finally, additional backgrounds from electron trains can often produce S2s followed by a flurry of S2 and SE pulses lasting a few seconds after the initial S2. It then seems obvious that characterising the fragmentation effects using the now well-validated waveform simulation is the optimal method for this, since it eliminates all possible background pathologies. It should also be noted that this analysis is not possible with the fast chain simulation due to its lack of waveforms; thus, this study is greatly simplified using the full chain simulation.

6.1.1 Simulation Measurements

The ParticleBomb generator is perfectly designed to measure such effects, since it produces a fixed number of electrons at a random position within the detector; hence, the exact number of electrons that should contribute to the S2 is known. Note that generating the initial number of electrons does not guarantee that all of those electrons are extracted into the gas and produce S2 signals. To account for this, the data is re-binned in terms of the electrons extracted rather than electrons generated. Henceforth, any numbers of electrons refer to extracted electrons unless explicitly stated otherwise.

By nature of the S2 mode of the ParticleBomb generator, the resulting events from LZAP should contain a single S2 pulse; thus, any event containing multiple S2s or additional SE pulses has demonstrated the fragmentation effect. With this knowledge, the simulation can be used to assess the impact of both misclassifying multiple scatter events and losing pulse area and width in the prominent S2s.

Misclassified Multiple Scatters

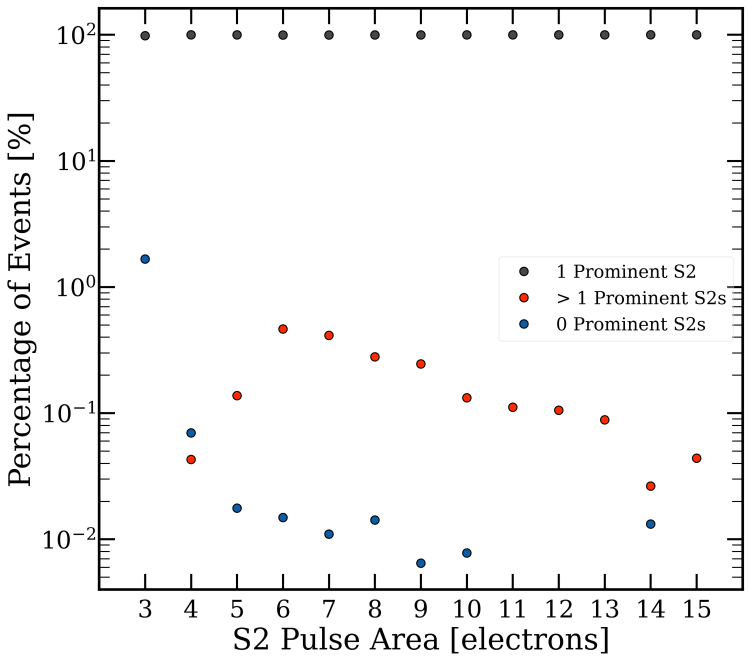


Figure 6.2: Percentage of events generated using the ParticleBomb S2 generator which result in either 1 prominent S2, more than one prominent S2 or no prominent S2s. A single prominent S2 is used as a proxy for a single scatter event, while more than one is used as a multiple scatter event. The percentage loss across multiple scatters is negligible.

To quantify the fraction of events lost to multiple scatters, as in Case 1, we need to determine the number of prominent S2s in each event. An event is classified as a Multiple Scatter by LZAP if it contains one S1 and more than one prominent S2; events with one S1 and only one prominent S2 are labelled Single Scatters. It is expected that the loss to Multiple Scatters will be dependent on the number of extracted electrons since more electrons are inherently necessary to produce multiple prominent S2s. The percentages of events classified as Single Scatter, Multiple Scatter, and Other are shown in Figure 6.2, which demonstrates that the loss to Multiple Scatters is negligible, with a maximum of $<0.5\%$ for $6\,\mathrm{electrons}$. This is deemed to be a negligible loss of signal events in this region.

Reduced Areas and Widths

Now inspecting Case 2, fragmentation is expected to be prevalent at the bottom of the detector due to the longer drift times and hence more prominent diffusion effects. The data is therefore binned in both drift time and number of electrons, and the probability of fragmentation is calculated in each bin by

$$P_f = \frac{\text{Number of Events with >1 S2/SE Pulses}}{\text{Total Number of Events}}. \tag{6.1}$$

Figure 6.3 confirms the expected behaviour of this effect. Firstly, the probability of fragmentation is much higher at the bottom of the detector, reaching up to 31% for $3\,\mathrm{electrons}$. Secondly, there is a general trend: the probability is higher for smaller electron sizes. This is due to the enhanced discretisation of smaller electron numbers, allowing for a discrete number of pulses.

The probability of fragmented S2s of these sizes is significant, and the extent of the S2 pulse area lost during fragmentation must also be characterised. Figure 6.4 shows the measurements of the fraction of pulse area lost as a function of extracted electrons. The violin plot indicates that, across all S2 sizes, the most likely outcome is the loss of a single electron. Then, for S2s of 4–$15\,\mathrm{electrons}$, there is also some possibility of losing 2 or even $3\,\mathrm{electrons}$ to this effect. This loss has a bigger impact on the smaller S2 sizes since they lose a larger fraction of the total S2 area. The probability distributions for each of the violins in the figure are fit to Gaussian and double Gaussian distributions where relevant. The triple Gaussian is neglected since the possibility of losing $3\,\mathrm{electrons}$ is deemed statistically negligible and an accurate fit was impossible to achieve. The loss of pulse width is not included here since the statistical analysis is performed in terms of

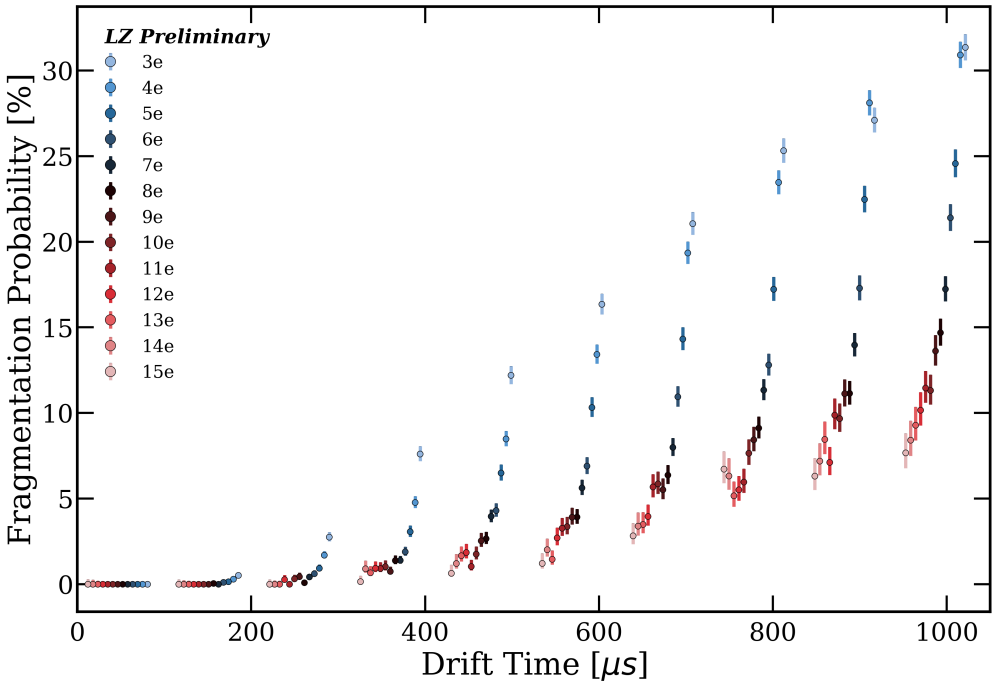


Figure 6.3: Probability of fragmentation of an S2 pulse consisting of 3–15 electrons in drift time bins spanning the total liquid xenon volume. Fragmentation remains a significant problem near the bottom of the detector, as expected from the diffusion relation.

S2 size. However, this will become relevant in Section 6.4 and discussed in detail there. Hence, the fragmentation effect has been characterised to the level necessary to apply corrections to the data.

6.1.2 AmBe Validations

While these simulations have been extensively validated to produce accurate S2 pulses, we must validate these measurements with data. In particular, we wish to replicate Figure 6.3 with a clean selection of calibration data. Figure 6.4 is more difficult to replicate in data since it requires the knowledge of the number of electrons extracted. We can estimate this value since the mean SE size is known, but there are sufficient fluctuations in this value that electron counting may lead to inaccurate measurements. In fact, even Figure 6.3 is difficult to replicate with the data. This measurement counts the probability of an S2 fragmenting, irrespective of the type of fragments produced. For example, most of the fragments observed in simulation are single electrons. In real data, there is a persistent background of single electrons, so simply selecting any event with an S2 and an SE pulse close by will overestimate the probabilities by including other background

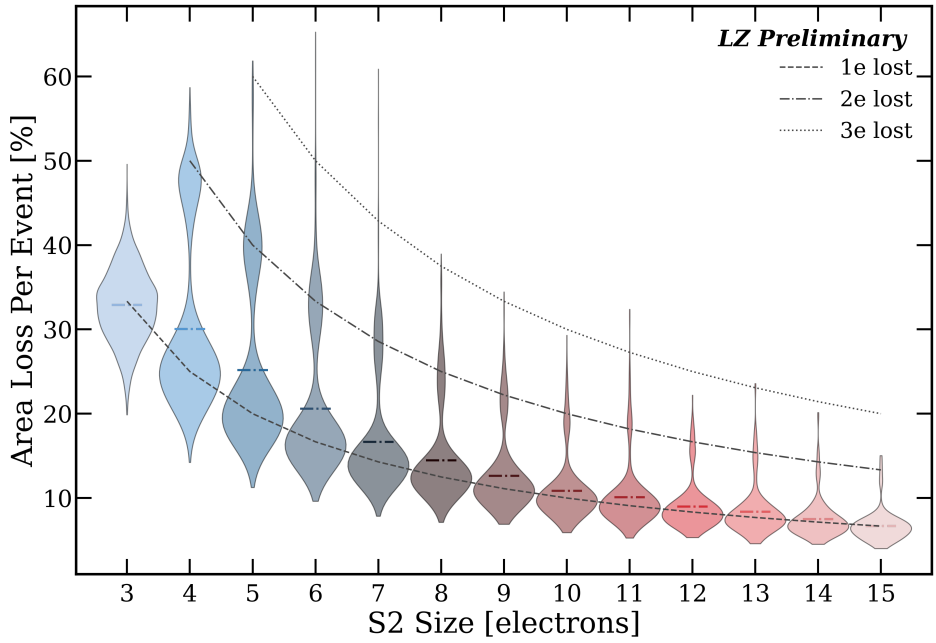


Figure 6.4: Distribution of pulse area lost as a percentage of the total area of the S2 pulses for each S2 size in the low-energy ROI. The marker shows the mean percentage loss on each distribution.

topologies. To accurately validate these measurements with data, one must first develop criteria to effectively identify these fragmented S2s.

Since all electrons in the S2 are released at the same position in the detector, one can measure physical bounds for the maximum time separation between electron pulses for them to be considered fragmented. This maximum time separation is governed by the longitudinal diffusion of the electrons, which is symmetric before and after the prominent S2, since the diffusion is normally distributed. However, due to the background pathologies mentioned previously, it is known that real data will contain events that resemble the fragmentation effect from different sources. These sources are not all symmetrical about the S2, so the measurements of the time window between the prominent S2 and its fragments are split into pre- and post-S2. Figure 6.5 demonstrates this separation of measuring times in both directions. The time between successive pulses is evaluated by the difference between AFT5 of one pulse and AFT95 of the other, since these are more stable than the start and end times of the pulses. Additionally, the electrons in all pulses should be reconstructed to the same (x, y) position. For SEs, position reconstruction is weaker because fewer photons are detected to measure their positions, which can cause the reconstructed positions to deviate from the true positions. Because of this, it is reasonable to expect a larger radial separation between the reconstructed positions and the true positions of the events given by the simulations.

These measurements are also shown in Figure 6.5.

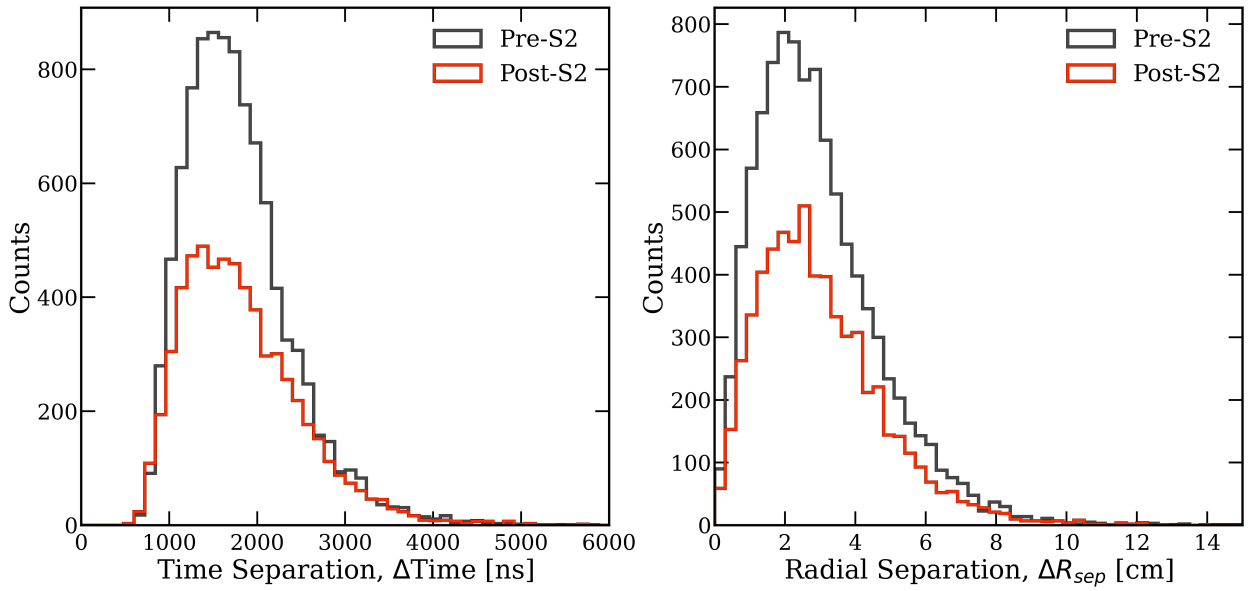


Figure 6.5: Time and radial separations between the prominent S2 in each simulated event and any adjacent SE or S2 fragments shown for both pre- and post-S2. Time separation is defined as AFT5 of one pulse to AFT95 of the other. Radial separation is determined from the reconstructed positions of each pulse.

The measurements confirm symmetric distributions in both time separation and radial separation for the fragments. From these distributions a simple selection criteria of $<5000\,\mathrm{ns}$ in time and $<10\,\mathrm{cm}$ in radius is derived. In Section 6.4 this selection criteria is improved with the incorporation of data quality cuts. To produce viable comparisons between simulation and data, we neglect fragmentation to SEs entirely, focusing only on the probability that one S2 fragments into multiple S2s, following the criteria mentioned above. This is due to the SE background, which would bias the selection with false positive events. The data used for this comparison are taken from AmBe neutron calibration data since it provides the largest number of S2s within the range of 3–$15\,\mathrm{electrons}$. These neutrons can produce S2s in the range of 3–$15\,\mathrm{electrons}$ but with limited statistics, $\mathcal{O}(10^3)$ events, making this validation difficult. To address the limited statistics, the comparison is based only on drift time and is not broken down by individual S2 sizes. Nevertheless, the comparison shown in Figure 6.6 indicates consistency between simulation and data. The consistency observed between simulation and data therefore allows us to measure this effect in the simulation and then apply it to the data.

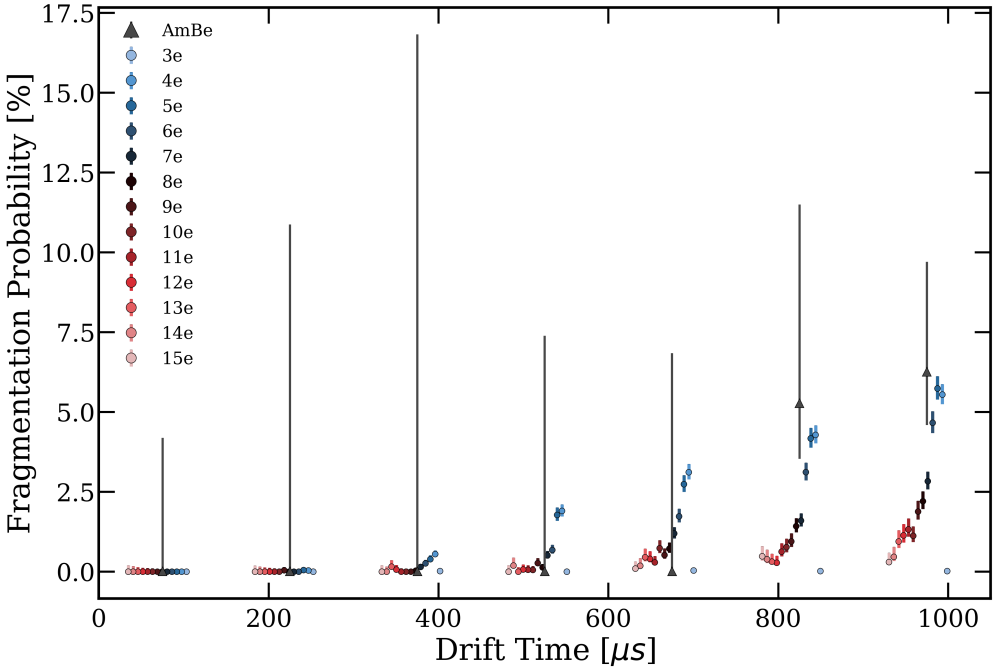


Figure 6.6: Probability of fragmentation to multiple S2 pulses. AmBe calibration S2s are displayed as triangles, compared with the simulation measurements shown as circles. The error bars are large for the AmBe calibration data due to limited statistics. Nevertheless, the simulation is consistent with the calibration data.

6.2 Signal Modelling

With the measurements of the fragmentation probability from Figure 6.3 and lost area from Figure 6.4, it is possible to apply the fragmentation effect as an efficiency loss to the signal modelling. In principle, this can be applied to any signal model, but in this case it was applied directly to the ^8B CEνNS signal model. For statistical inference in LZ, the fast chain simulation is used to produce large-scale, $\mathcal{O}(10^7)$, signal PDFs in S1 vs S2 space. This allows us to compare real data events with the signal model PDF and extract the likelihood that each event belongs to the signal model. As previously stated, the fast chain simulation cannot produce waveforms and therefore cannot reproduce the fragmentation effect. To apply the fragmentation effect to the signal PDF, the following process is performed:

1. For all events in the signal model simulation, identify the drift time and number of electrons extracted

2. For each event determine if the S2 will be fragmented by sampling the probability from

Figure 6.3

3. If the S2 is fragmented, proceed to sample from the distributions in Figure 6.4 to determine the pulse area lost from fragmentation

4. Subtract the lost area from the S2 pulse and reassign its new position in S1 vs S2 space

Figure 6.7 illustrates the total effect of this implementation, showing minimal impact on the total S2 spectrum but a loss of signal below the $3.5\,\mathrm{electron}$ threshold. This is expected, since this effect will cause the S2 size to decrease by one electron for most pulses; hence, the pulse area loss will cascade to fill the space of smaller S2s, but the smallest S2s will fragment to be below threshold. From these studies, it is noted that 11.5% of signal events are selected as fragmented, and the average lost area for fragmented events is 20.1%.

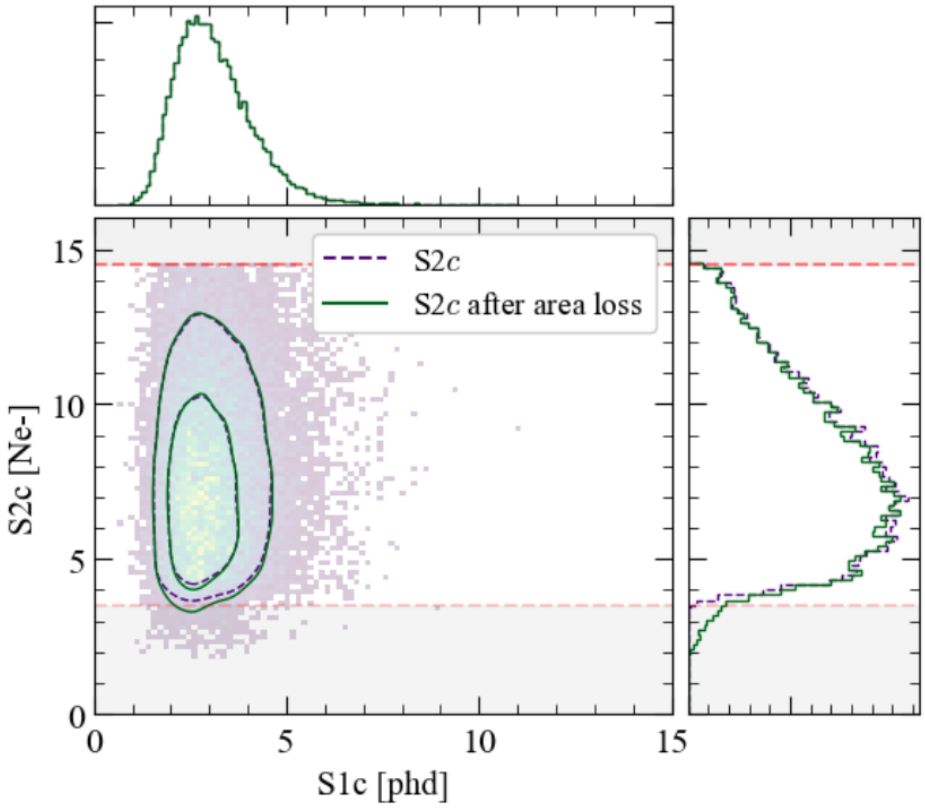


Figure 6.7: Impact of applying the fragmentation probabilities and area loss distributions to the ^8B PDF generated using LZLAMA. This area loss causes some signal loss by pushing signal events out of the ^8B ROI (below $3.5\,\mathrm{electrons}$ in $S2_c$). Figure from Alberto Usón.

This is the final version of the ^{8}B signal model used for LZ's published CEνNS result, which reports the first evidence of ^{8}B CEνNS interactions on xenon with 4.5σ significance [1]. This was

devised as the simplest approach to correcting for the fragmentation effect, given the known deficit of losing signal below the threshold.

6.3 Fragmentation Cut

While altering the signal modelling works for signal detection, it does so without accounting for data quality. The RQs determined by these fragmented S2s do not represent the true information about the event but rather a skewed reconstruction artefact. Many low-threshold searches are sensitive to small changes in the S2 size by even a single electron that may be fragmented. To preserve the quality of the data, a cut is designed to remove any events that exhibit fragmentation using the time and space criteria described previously, while maintaining the expected signal modelling prior to the corrections shown in Figure 6.7.

6.3.1 Cut Logic

Primarily, fragmentation results from longitudinal diffusion, which spreads electrons over time. Hence, for each prominent S2 pulse in an event, a physically motivated time window around this pulse can be enforced, which is required to be entirely free of any other S2 or SE pulses that may have been fragmented. From Figure 6.5, we define this time boundary to be $5000\,\mathrm{ns}$ between subsequent pulses around the prominent S2. Additionally, it is known that a single electron from the detector's constant SE background may fall within this time window and cause us to remove several well-reconstructed S2s incorrectly. To reduce this number, we enforce a secondary requirement that the fragmented pulse must be reconstructed to within $10\,\mathrm{cm}$ of the prominent S2 pulse, as measured in Figure 6.5. This ensures that all electrons are close enough in space and time for the pulse to most likely originate from fragmentation. A depiction of the cut logic is shown in Figure 6.8 and is demonstrated on a combination of fragmented simulation events and data in Figure 6.9.

With this definition of the cut, it is possible that it does not solely remove fragmented events but also other background topologies present in the data, such as electron trains. This can therefore be thought of as an S2-environment cleanliness cut and is a powerful tool for all searches including S2 pulses.

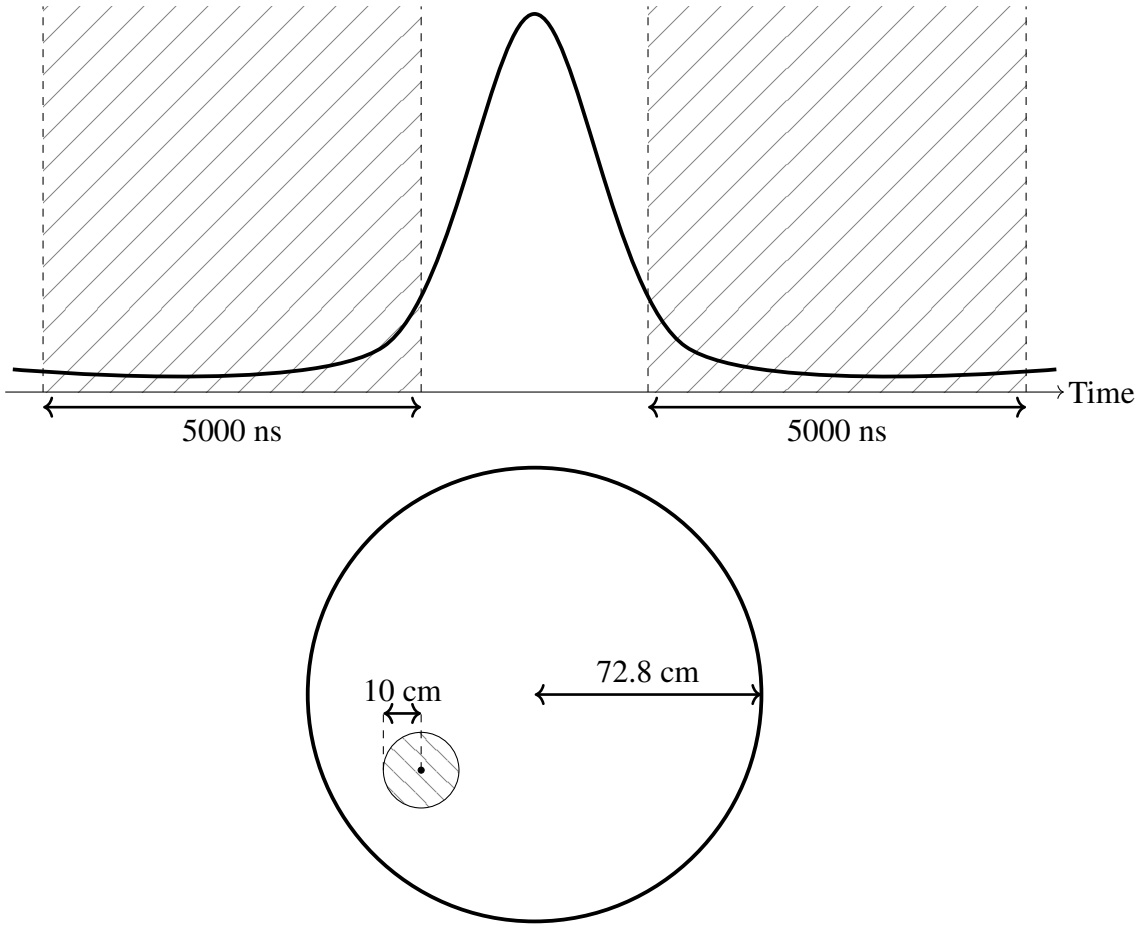


Figure 6.8: (top) Example S2 pulse with time boundaries explicitly shaded out. Any other S2 or SE pulses that exist in this time-space will cause this event to fail this cut so long as they also satisfy the spatial cut in (bottom). The spatial cut shows the $10\,\mathrm{cm}$ radius drawn around the prominent S2 position in the TPC. Any other S2 or SE pulses that are in this space, in addition to the time space, will be cut. The radius of the entire detector is given here for reference.

6.3.2 Signal Acceptance

For a cut to be applied in the analysis, we must be confident in its signal acceptance. Cuts are designed to maximise signal acceptance, preserving as much signal as possible. Here we define signal acceptance as the probability that a genuine signal interaction passes the cut. Even events where fragmentation occurs can be physically sound since fragmentation is a physical outcome of the longitudinal diffusion of the detector. By applying a cut to eliminate these events, we are inherently eliminating some physically sound signal events with poor reconstruction to retain only high-quality data. From this standpoint, we can assess how many of the signal-like events are retained by this cut. The signal-like events are constructed by selecting S1 environments from AmLi neutron calibrations and pairing them with S2 environments from CH_3T calibrations to

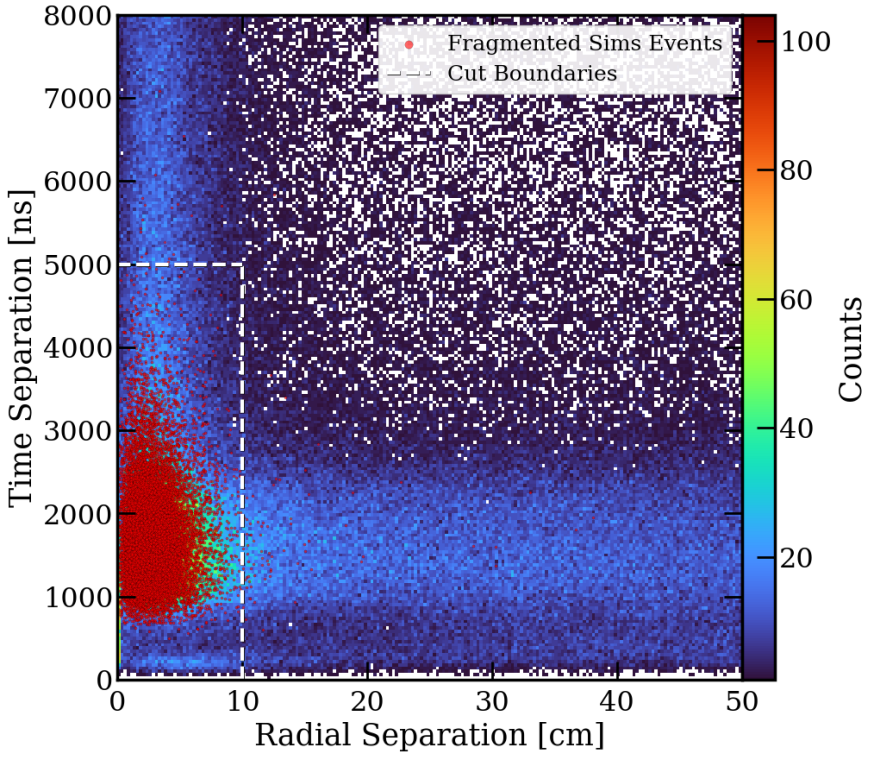


Figure 6.9: Fragmentation parameter space with the spatial and time separation between pulses illustrated. The red simulated data points indicate where the fragmented events lie, and the white dashed lines indicate the region of this parameter space removed by this cut.

resemble NR events. This process is called ChopStitch. The signal acceptance measured from this is shown in Figure 6.10.

This result is consistent with expectations. There are more fragmented events being cut at the lower S2 areas, while the larger areas tend towards 100% since fragmentation is less likely to occur for larger S2s. Note here the range of S2s tested in this cut extends to $50\,\mathrm{electrons}$ to cover the extended range of the S2-only search discussed in Chapter 7. Also note that some bins show 100% signal acceptance, meaning no events in those bins were removed by the fragmentation cut. The rest of the plot indicates a continuous increase with prominent S2 area due to the decreasing probability of fragmentation. Since this dataset is already produced after applying all other cuts to maintain a clean signal-like dataset, we can conclude that this is a reasonable estimate of the signal acceptance under this paradigm.

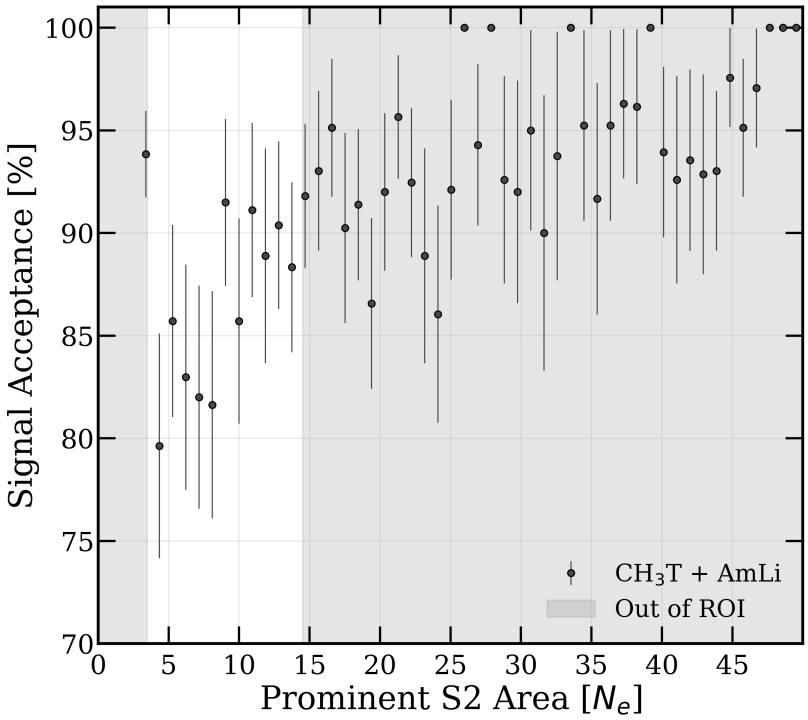


Figure 6.10: Signal acceptance of applying the fragmentation cut to a sample of CH_3T + AmLi ChopStitched data. This is constructed by sampling the S1 from the AmLi neutron calibrations and the S2 from the CH_3T calibrations to form NR events.

6.3.3 Background Rejection

The second parameter to optimise when designing cuts is the background rejection power. That is, how strongly does this cut reject background events? Again, to do this we must first determine what is meant by "background" in this context. In these low-energy regimes, the dominant background is from accidental coincidence events. Given this, it is prudent to evaluate the background rejection of some sample of generated accidental events. Similar to the signal model, these events are generated by randomly selecting S1 and S2 pulses from separate calibration events and pairing them together to artificially construct new events. However, these events need not resemble signal events. The details of this are neglected here but are well validated within LZ for background modelling [48]. Applying this cut to this background set, we observe the background rejection shown in Figure 6.11.

The rejection power shown by Figure 6.11 is consistent with the expectation from Figure 6.10. The datasets used to evaluate these efficiencies should contain a similar number of fragmented events, since they both have a uniform distribution of S2s with drift time. Therefore, it is expected that

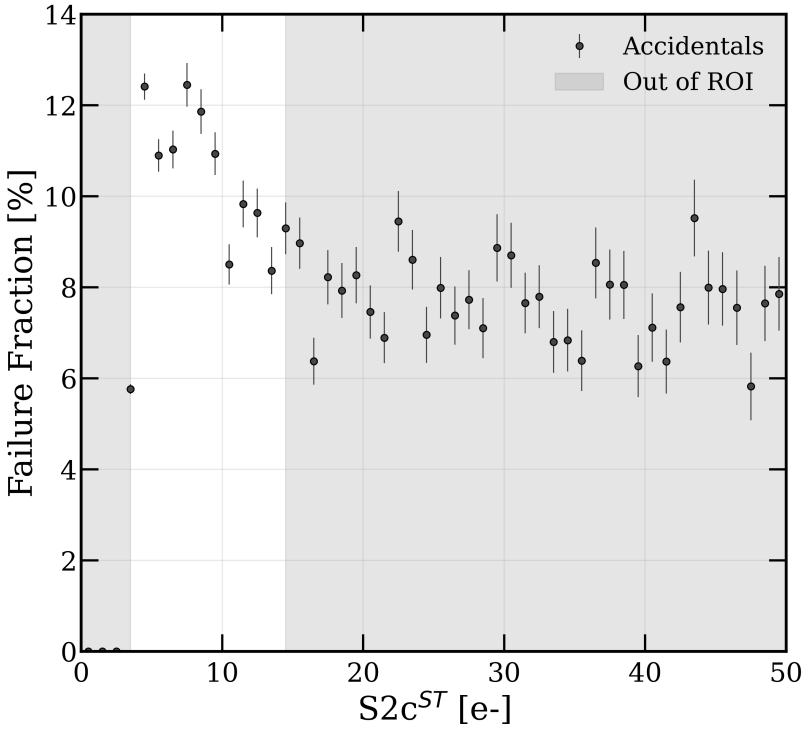


Figure 6.11: Percentage of accidental ChopStitch events that fail the fragmentation cut as a function of the spatio-temporally corrected S2 size in terms of electrons.

the signal acceptance and background rejection from these datasets are inverses of each other, as is observed in the two figures.

Despite Figure 6.11 seemingly having a small background rejection power, accidental events are not the topology that this cut is designed to remove. In fact, this cut is designed to preserve data quality and hence turns any poorly reconstructed S2s into background events. Selecting fragmented events from the data requires the logic of the cut and produces a circular bias that will result in 100% background rejection. From Figure 6.5, it is shown that even in simulation, we do not reject 100% of the fragmented S2s since we apply a cut of $5000\,\mathrm{ns}$ which is before the end of the tail of the time distribution. Additionally, from Figure 6.5, position reconstruction for SE pulses, with fewer photons than S2s, can be easily skewed by effects such as afterpulsing in certain PMTs, pushing their distance from the prominent S2 above $10\,\mathrm{cm}$. We therefore evaluate the background rejection by selecting all events from the simulated dataset with more than 1 S2/SE pulse, applying the cut logic in time and space, and evaluating the probability that we specifically target fragmented events with this cut. The outcome of this is shown in Figure 6.12. In this figure it is shown that this cut topology has a very high rejection power for fragmented events, cutting

an average of 99.3% of all fragmented S2s in simulation. The events that are not cut by this logic mostly include SE fragment pulses with poorly reconstructed positions due to the low number of detected photons, which therefore fall out of the $10\,\mathrm{cm}$ radius. Increasing this radius to catch the additional events will allow for a larger number of background SE pulses to be included in this cut, decreasing its overall accuracy.

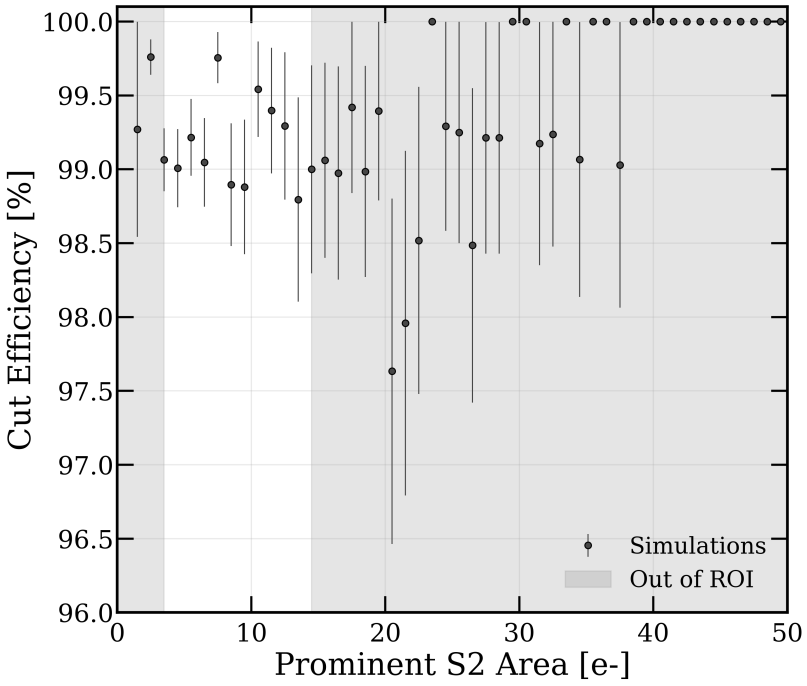


Figure 6.12: Cut efficiency for successfully removing simulated fragmented S2s binned by prominent S2 size in terms of electrons extracted. This results in a flat efficiency of 99.3% for targeting fragmented S2s. This also demonstrates that S2s up to $40\,\mathrm{electrons}$ can still exhibit fragmentation though its impact is less important since these lose a smaller fraction ($\sim$3%) of the total pulse area.

6.3.4 S2-Only Impact

For S2-only searches, where the S2 pulse is the only signal pulse in the event, maintaining high-quality S2 pulses is extremely important; hence, the fragmentation effect should be accounted for in this search. The S2O regime also has a different set of background sources, with the largest contribution coming from the grid wires, particularly the cathode. Cathode events, having the longest drift times in the detector, are most prone to fragmentation. This implies that the fragmentation cut should cut a higher proportion of cathode S2s than signal from the majority of the LXe bulk. Although the transition between cathode and LXe bulk S2s is continuous, there is no guarantee in this cut logic to target cathode backgrounds other than their higher probability of being frag-

mented. To test the drift time distribution of events removed by this cut, we must use single scatter events, since S2O events do not have drift time information since they do not have the associated S1 pulse present.

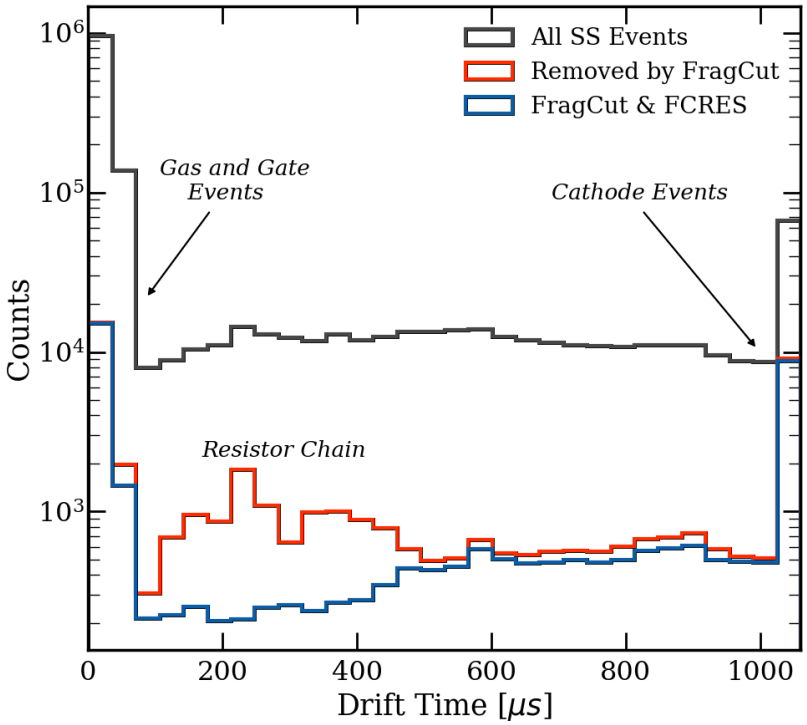


Figure 6.13: Showing the proportion of SS events removed by the fragmentation cut. A secondary peak is shown here, due to events from the resistor chain, as demonstrated by applying the resistor chain cut (FCRES), which removes any events with S2s near the (x, y) positions of the resistors. The final distribution shows a gradual increase in the number of events cut as drift time increases.

Figure 6.13 demonstrates the impact of applying only the fragmentation and resistor cuts to a random sample of single scatter data. For the main bulk of the liquid xenon (drift times from $\sim$100–1000 µs), the blue histogram confirms that more events are cut at the larger drift times due to increased diffusion. There is also a large peak at the maximum drift time, the cathode drift time, which shows that the number of events cut here increases by an order of magnitude. An unexpected but pleasant result is the removal of an extra order of magnitude of gas and gate events. Gate events are also background events originating from the gate grid wires near the top of the detector. However, without the diffusion that would cause them to be fragmented, one can conclude that these are removed here more likely due to rapid electron flurries occurring after grid emissions. This flurry of electrons has a different time distribution from fragmented electrons and is visible in Figure 6.9 as the band beneath all fragmented simulated events, closer to 0 ns. In practice, one

could adjust the boundaries of the fragmentation cut to raise the lower bound so as to remove only fragmented S2s. However, since this also serves as an S2 cleanliness cut, we also wish to remove the gas and gate events in this region, as they do not constitute signal events. In SS searches, LZ selects a fiducial volume with a lower drift time limit of 71 µs to remove gas and gate events. For S2O searches, this is not possible, and so allowing these to be removed by this cut inherently increases the quality of the S2O data.

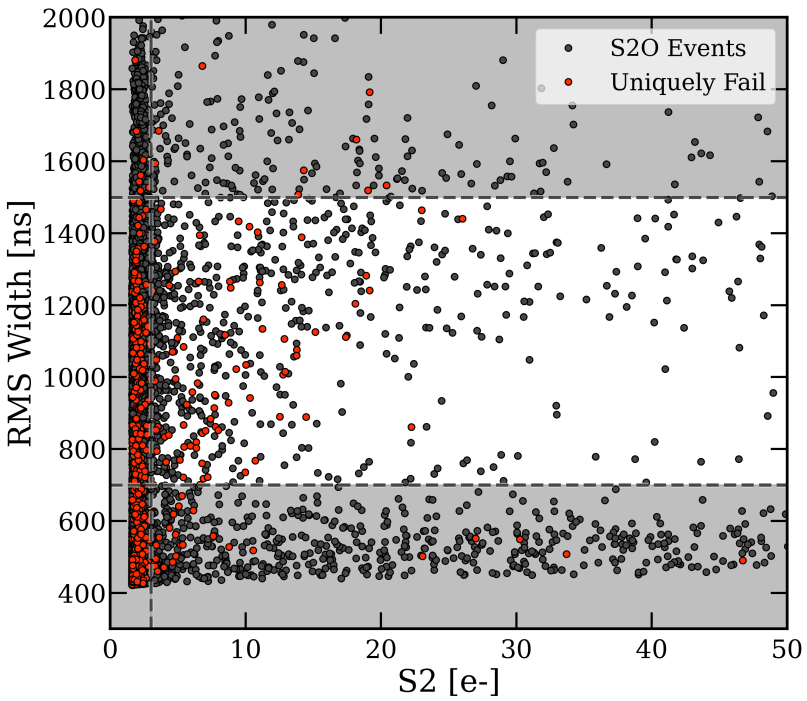


Figure 6.14: RMS width against S2 size phase space for all S2-only events in the 10% unblinded dataset. Areas outside of the region of interest are shaded in grey. The black data points here show the events in the search that pass all cuts except the ROI cuts in RMS width and S2 pulse area. The red data points show the events that uniquely fail the fragmentation cut and pass all other cuts applied here. In the highlighted region of interest, 67 events are uniquely removed by the fragmentation cut, 12% of the total number of events.

Since an S2 cleanliness cut will have some overlap with other cut topologies, some already noted here, it is important to understand how many events are being uniquely removed by this cut and what kind of events they are. To this end, we introduce the parameter space for the S2O search. With the only signal pulse being the S2 itself, the best dimensional space in which to visualise the data becomes S2 pulse area and RMS width. Since the widths indicate depth, we can use this space as a proxy to fiducialise in depth and remove most gate and cathode backgrounds, though this is not absolute and still results in the inclusion of some gate and cathode events in the final dataset.

For the latest cut tunings available at the time of writing, the fragmentation cut uniquely removes 67 events in the region of 3.5–$15.5\,\mathrm{electrons}$ and RMS width between 700 to $1500\,\mathrm{ns}$ for a small unblinded sample dataset ($\sim$10% of the full S2O dataset). The events uniquely removed by this cut are highlighted in the entire parameter space in Figure 6.14.

The cuts in RMS width are present to fiducialise in depth. Below $700\,\mathrm{ns}$ RMS width we expect a population of gate events, with very short drift time they experience less diffusion and hence the width of such S2s are narrow. Likewise the cathode S2s are expected to lie mostly above the $1500\,\mathrm{ns}$ RMS line. Additionally, we apply a lower bound on the S2 area to the S2O search, since most of the cuts applied were originally tuned above this threshold for the ^{8}B CEνNS and low-mass WIMP analysis. Figure 6.14 shows that the fragmentation cut removes a large population of S2s below the pulse area threshold. This is likely due to pileup of spontaneous emission of electrons from the grids, which form accidental S2s. In the accepted region of this parameter space, this cut removes 67 events that would previously have been considered candidate events. These events are incorrectly reconstructed in both pulse area and width – the only parameter space taken into account for the S2O search – so it is crucial to remove them.

6.3.5 ^{8}B CEνNS Impact

The analysis for the ^{8}B CEνNS and low-mass WIMP (LMW) search was finalised prior to the creation of this cut. However, this cut can be applied to the final dataset to investigate whether any further events would have been removed and to assess their nature: signal or background. This cut removes a further 2 events from the 19 events in the official dataset, shown in Figure 6.15, bringing the new dataset down to a total of 17 observed events. The effect of this cut on the statistical analysis and final ^{8}B discovery significance is dependent on the nature of the two events.

For this discussion, we will label these events $E2$ and $E4$ related to their respective $S1_c$ values in Figure 6.15. Two simple checks can be made to firstly gather an understanding of whether these are signal or background events:

1. Check drift time - it is more likely to have fragmented events from the bottom of the detector
2. Inspect event window by eye - what do the S1 and S2s look like?

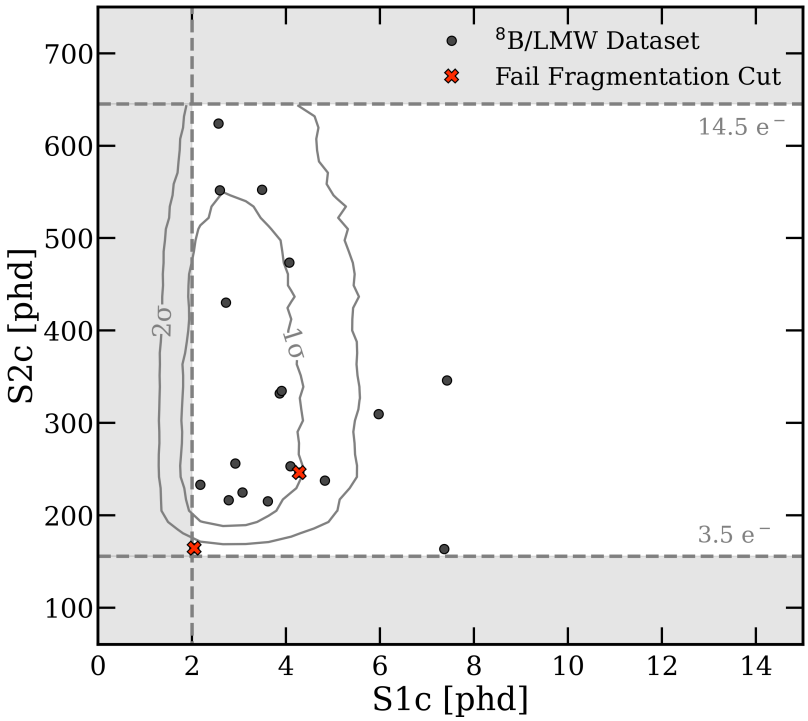


Figure 6.15: Final ^{8}B/LMW dataset with salt removed. The fragmentation cut would have removed two events. The contours show the ^{8}B signal model contours from simulation.

Check 1 results in Figure 6.16, which shows the radius and drift time of the entire dataset, indicating the positions of $E2$ and $E4$. From this figure, it is likely that $E4$ is from a real signal event that has been fragmented, since its drift time places it near the bottom of the detector. $E2$, however, is much closer to the liquid surface and thus has a higher chance of being an accidental event. Check 2 concurs with this initial hypothesis: $E2$ occurs in an event window with a high SPE rate, and the S1 pulse is likely due to an accidental pileup of 3 SPEs, hence the mismatch in the drift time. Therefore, we make the assumption here that $E2$ is an accidental event. The event window for $E4$, however, is reasonably clean, and both S1 and S2 appear reasonable relative to each other and to the event environment. We can therefore assume this cut removes 1 signal event and 1 background event.

Before performing a full PLR statistical analysis with the new dataset, we must first investigate the impact of this cut on both the signal and background models. The signal model is relatively simple to implement by reverting the changes applied in Figure 6.7 and returning to the original signal model. Then the signal acceptance from Figure 6.10 is applied to the signal model PDF to account for the signal loss caused by this cut. Accidentals dominate the background model for

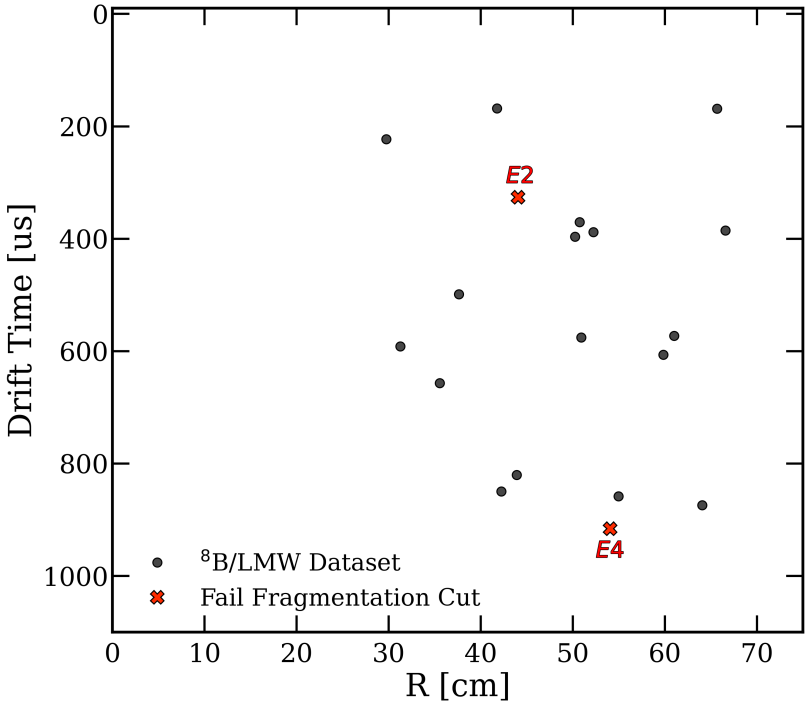
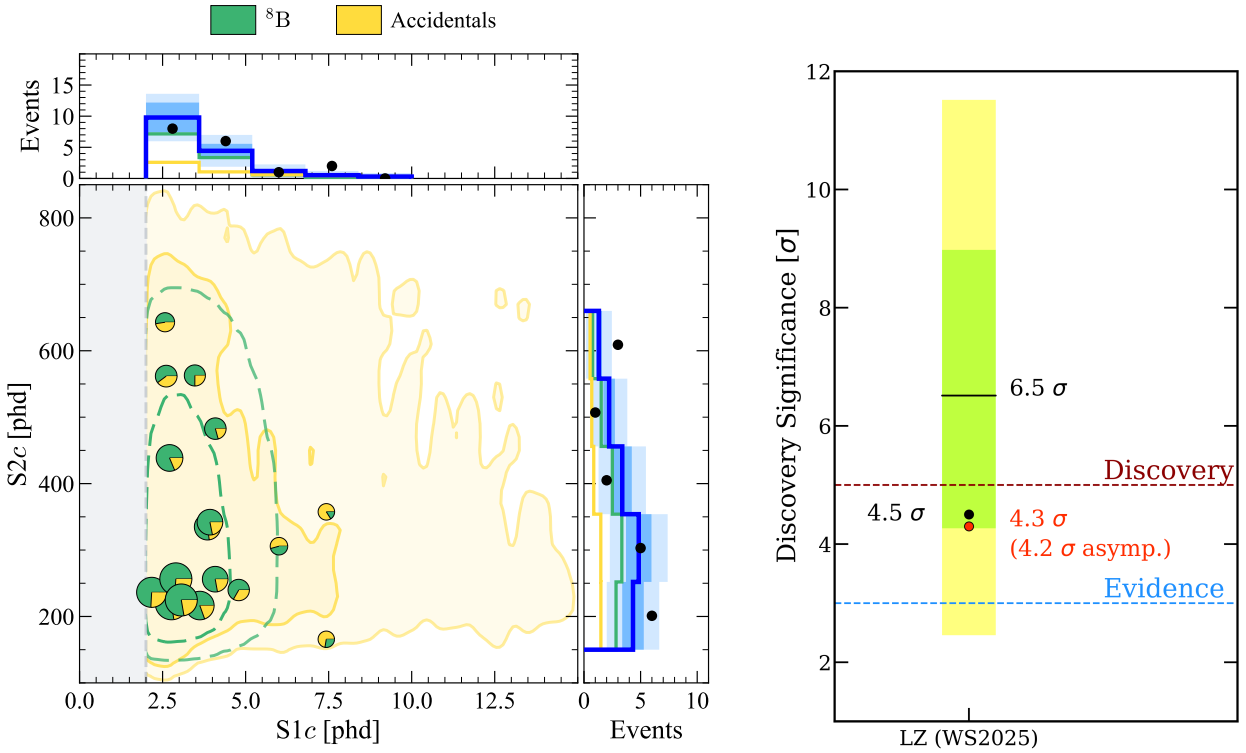


Figure 6.16: 2-dimensional projection of the reconstructed position of all events in the final dataset. Highlighted are the two events which fail the fragmentation cut.

the ^{8}B CEνNS search; hence, the cut must be applied to the accidental population to produce a new background model PDF, which is then used to estimate the number of accidentals expected in the final dataset. Prior to applying this cut, 6.55 ± 0.44 accidental events are expected to be present in the final dataset. Applying this cut reduces this value to an expected 5.53 ± 0.24 events, consistent with the assumption that this cut has removed 1 accidental event. The PLR analysis is then performed on the new signal and background models using the reduced dataset, resulting in an overall reduction in observed significance from 4.5σ to 4.3σ in Figure 6.17b.

The reason for this decrease is due to the position of $E2$ (assumed accidental) and $E4$ (expected signal) events in S1-S2 space relative to the ^{8}B signal contours. $E2$ is sufficiently far from the signal model contours that removing this event will not substantially affect the observed significance. On the other hand, removing $E4$ from inside the 1σ contour has a substantial impact on the significance since an event is lost which had a high likelihood of being a signal event. From this study, we conclude that applying a flat fragmentation cut to the data removes approximately equal amounts of signal and background, which can lead to detrimental effects. This outcome motivates the need for a post-LZAP reconstruction algorithm that can correct for fragmentation rather than remove it, preserving the signal and allowing the other accidentals-targeted cuts to handle those



(a) Distribution of the events in the final dataset. Each data point is represented by a pie chart showing the likelihood of each event belonging to each model. The size of each data point represents how ^{8}B-like the event is.

(b) New ^{8}B signal significance when applying the fragmentation cut into the analysis, resulting in a reduction of observed significance.

Figure 6.17: Final impact of the fragmentation cut on the ^{8}B CEνNS observed significance.

backgrounds appropriately. This is further exemplified by examining the correlation matrix of the cuts applied in this analysis.

6.3.5.1 Cuts Correlation

Despite this having a negative impact on the ^{8}B significance, the finalised cuts list can be used to ensure that the fragmentation cut primarily targets an independent topology in the data not accounted for by the pre-existing cuts. To evaluate this, we produce the correlation matrix of all cuts used in this analysis, shown in Figure 6.18, which highlights the fragmentation cut. Figure 6.18 demonstrates that there are a few correlations between the fragmentation cut and a handful of other cuts, but none are unexpected. The correlated cuts primarily concern the shape of the S2 waveform. The narrow S2, S2 amplitude, and S2 rise time cuts are designed to target S2s with anomalous shapes relative to their pulse area or drift time. Fragmented S2s are more likely to fail the narrow S2 cut due to the inherent loss of pulse width; the same reason they are more likely to

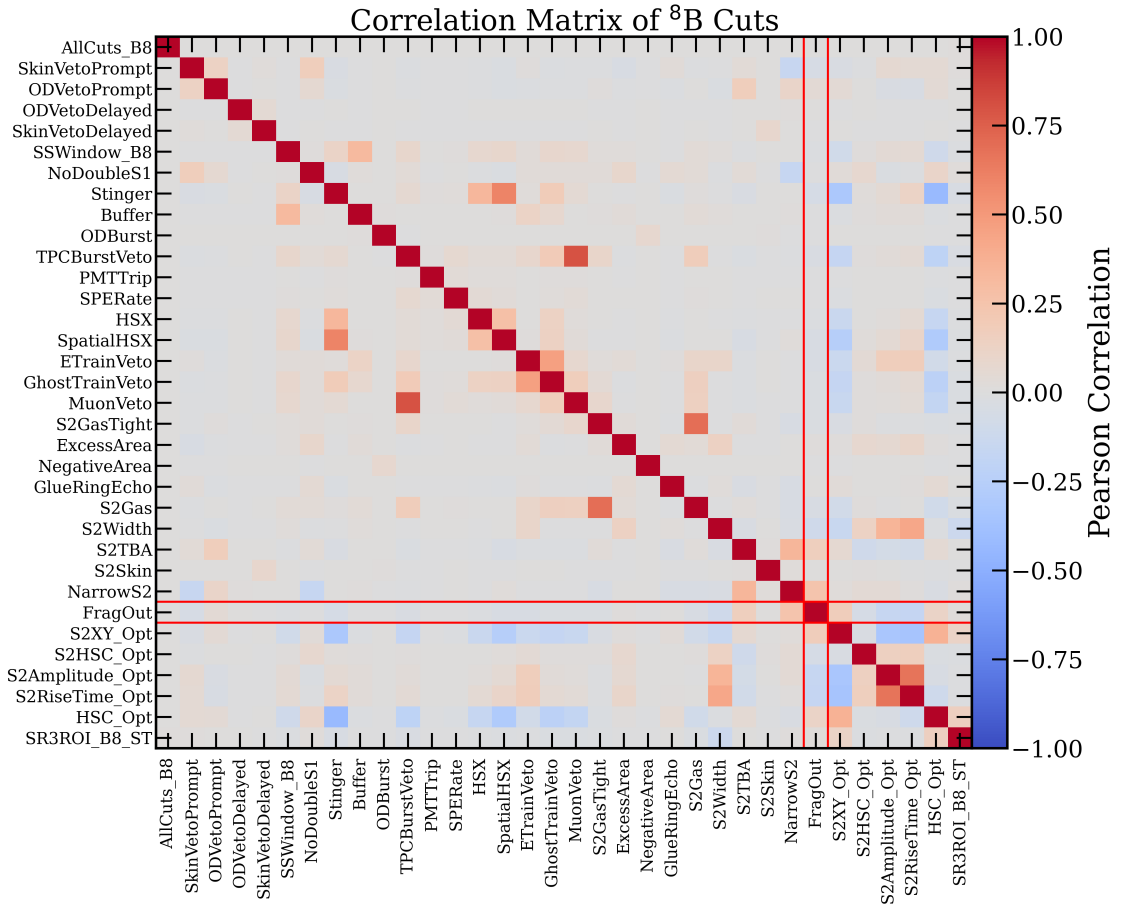


Figure 6.18: Correlation matrix between the cuts applied to the ^{8}B analysis. The fragmentation cut is highlighted with red borders, showing minimal correlation with most cuts. All correlated cuts are S2-shape cuts.

pass the rise time cut. The 'S2XY_Opt' cut refers to the quality of the reconstructed position. The correlation with the fragmentation cut here is more likely due to the tendency of smaller S2s to have poor position reconstruction and also to fragment. This correlation matrix demonstrates two main points. Firstly, the fragmentation cut can be reasonably applied to analyses as an additional cut, with little correlation with the other cuts applied, and thus targets a specific event topology. Secondly, from the correlated S2 shape cuts, fragmented events are incorrectly passing or failing certain cuts, which could change the dataset used for the final analyses when corrected.

6.4 Fragmentation Tag

The final approach to handle this effect would be to correct the mis-reconstruction. This involves firstly developing an accurate tag for this pulse topology, then developing a reconstruction algo-

rithm to piece the fragmented pulses back together. In the future, improvements will be made to the LZAP pulse finding algorithm, which will reduce the frequency with which S2s are fragmented. However, since LZAP is used to analyse all LZ data, any changes to this framework must undergo rigorous validation that takes months to push to the collaboration. Following these updates, a full re-processing of all data is necessary, which in turn can also take a few months to complete. For this reason, we choose to take a post-processing reconstruction approach, meaning we use the information available to us from the current fragmented pulses to reproduce what we would expect to see with the LZAP improvements, without the need for re-processing. This reconstruction algorithm will be discussed in Section 6.4.2.

6.4.1 Tag Logic

Firstly, the fragmentation tag must be developed. The logic contributing to this tag is similar to the fragmentation cut. That is, in the data, the simplest method to determine whether a pulse is fragmented is to draw windows in time and space around the prominent S2, as determined from simulation. However, the existing background topologies that also arise within this parameter space introduce additional cut requirements to increase the accuracy of the tag specifically to fragmented events. The first additional criterion to implement is to add a lower bound on the time window between subsequent pulses. The simulation measurements in Figure 6.5 demonstrate that fragments only occur when the time separation is above $600\,\mathrm{ns}$. This is now enforced as a lower limit during tagging to prevent inaccurately tagging events from the gate. The second criterion to consider is the possibility of a photoionisation peak following $\sim$2500 ns after the S2 due to photoionisation of the gate grid wires producing an additional SE. The probability of gate photoionisation by an S2 photon is $\sim 0.005\%$, leading to a total of $0.43\,\mathrm{photons}$ producing photoionisation for a $15\,\mathrm{electron}$ S2 and hence is negligible in this region of interest. However, when approaching $40\,\mathrm{electrons}$, which from Figure 6.10 is still possible to be affected by this logic, the probability suggests that $1.15\,\mathrm{photons}$ will produce a photoionisation electron. For large pulses like this, we impose a limit that the fragment pulse must be $> 5\%$ of the prominent S2 pulse's area to be included, as measured in high-energy searches where photoionisation peaks are prominent.

By applying all other cut definitions to a random sample of S2s from the WS2024 data, we can assess the impact of these cuts in this parameter space and specifically target event topologies that

mimic fragmented pulses. This is shown in Figure 6.19.

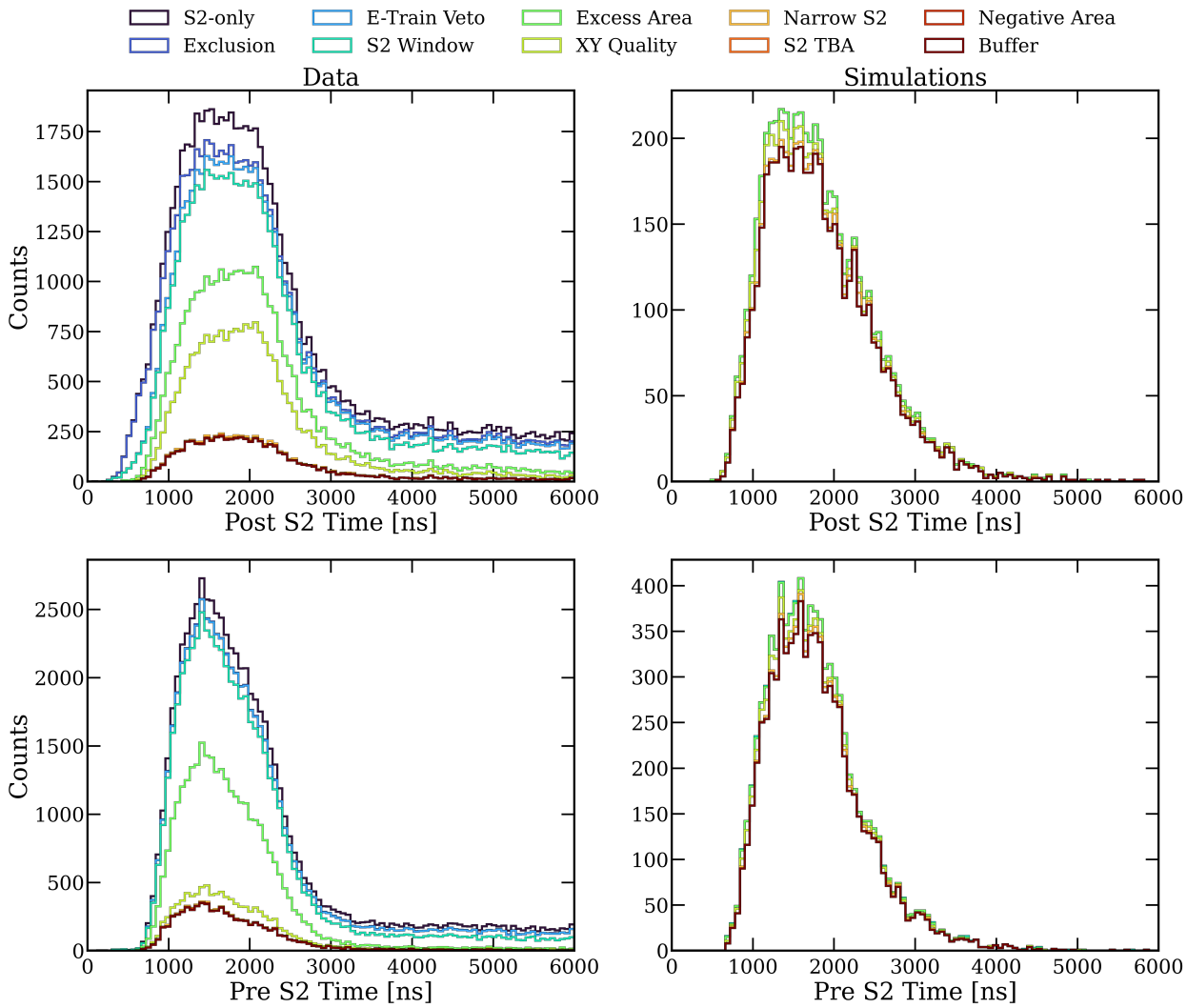


Figure 6.19: Cuts applied sequentially to a selection of S2-only data and simulation, showing the impact of each cut on the time between consecutive pulses pre- and post-S2. A full description of the cuts is omitted here, and only specific cuts will be discussed in the text where relevant.

This tag can be applied to both the ^{8}B CEνNS search (with both S1 and S2) and the S2O search, since it is defined without reference to the S1 pulse. Given the differences in information attainable by the two event types, they require different definitions of certain cuts. To remain equally applicable to both analyses, we choose the necessary cuts to be effectively equivalent. Initial cuts are applied to remove periods of poor data quality from the dataset. This may include electron trains (e-trains), events with PMT trips, hotspots or high SPE rates. Hotspots are periods of high SE and S2 activity caused by continuous electron emission from a high-voltage grid wire. They are localised in (x, y) position for some amount of time, meaning that a cut around the (x, y) po-

sition for that period of time removes these events. While only e-trains and hotspots can produce a false positive for a fragmented pulse, applying these additional cuts reduces processing time by immediately rejecting poor-quality data. Next, we apply a fiducial volume cut. This is different for S1-S2 and S2O analyses since S2O cannot perform a cut in z, only (x, y). The only additional cut found necessary to apply is the Excess Area cut. This cut removes events where the total pulse area of all pulses (excluding SPEs) prior to the S2 exceeds the pulse area of the S2. In practice, this should only occur for events with S1 pulses larger than their associated S2 pulses (such as alpha decays) or for events with high SE rates. For S2 sizes of less than $50\,\mathrm{electrons}$, it is more likely that the Excess Area cut is removing high-SE-rate events rather than alpha decays. The effect of these cuts is shown in Figure 6.19 for the S2O dataset, demonstrating the necessity of applying cuts to the data that will not affect the simulation.

With these pre-tag cuts applied and the time and space windows modified as necessary, we must now validate the tagging capability. The simplest approach is to apply the tag to both the S2O simulated dataset and a real sample of WS2024 data which should contain the same proportion of fragmented S2s. Applying this to the simulated S2O dataset returns a total of $3.02 \pm 0.03\%$ fragmented S2s. Applying this to the real data yields a total of $4.25 \pm 0.03\%$, which does not agree with the expectation from the simulation. The errors in these measurements are calculated with pure binomial statistics, assuming no systematic uncertainties.

To understand this discrepancy, a selection of events was examined manually. The result of this investigation revealed a significant proportion of the events from the data contained multiple SE pulses within the event window that were originating from the same (x, y) position as the prominent S2 but were sufficiently far away in time that they could not have been separated by diffusion; sometimes further in time than the full drift time of the detector. Most events observed contained multiple SE pulses of this description. An example of one such event is shown in Figure 6.20. Since these S2 pulses were small ($\leq 5\,\mathrm{electrons}$), the current hypothesis is that these events are caused by a coincident emission of multiple electrons from grid wires close enough in time that they are merged to form an S2 in LZAP. This corroborates the observation that there are multiple other SE pulses from the same position before and after this S2, as this could be thought of as a low-activity hotspot in which a single point on a grid wire sporadically emits electrons. We hereby refer to the S2s produced in this manner as Multiple Single Electron (MSE) pulses. That

is, they are more likely a coincident overlap of multiple SE pulses than a single site S2 pulse. A full discussion of MSE pulses as a background for the S2O search is in Section 7.2. For now, they can be removed from the fragmentation tag by requiring that there exist no other SE pulses from the same (x, y) position (within $10\,\mathrm{cm}$ radius) and outside of the time window ($5000\,\mathrm{ns}$) of the fragmentation tag.

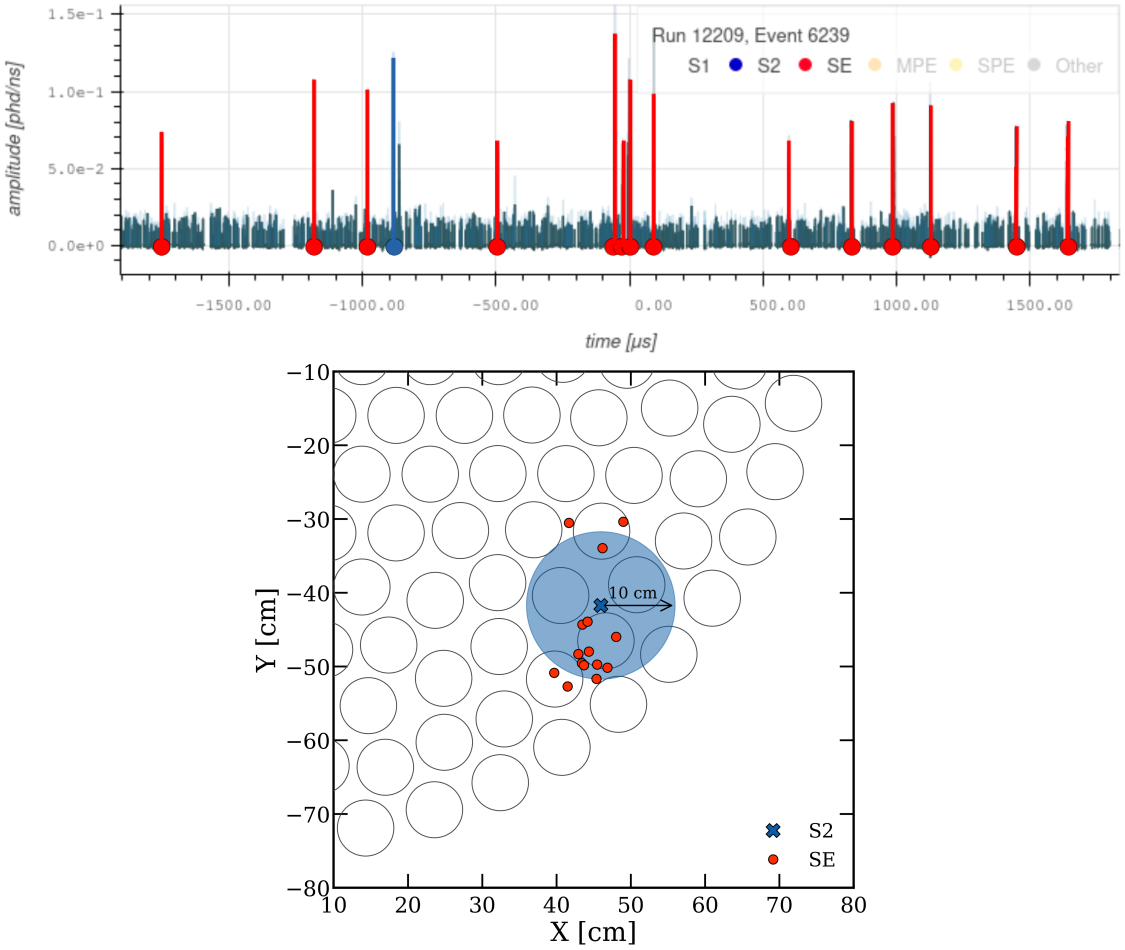


Figure 6.20: (top) Example MSE event window. The red lines indicate SE pulses within the event window, and the blue line shows the S2. (bottom) The same event shown in the reconstructed position space where the colours remain the same. Since all SE pulses are so close to the S2 pulse position, this is most likely an S2 caused by coincident overlap of multiple SE pulses. A 10 cm circle is drawn around the prominent S2 position to illustrate the radius within which electrons are tagged as MSE, motivated by the simulation.

Introducing this additional criterion reduces the proportion of fragmented events in the data to $3.25 \pm 0.03\%$. This is still in disagreement with the $3.02 \pm 0.03\%$ from simulation. Further investigation into the remaining tagged events reveals no further obvious topologies responsible for this discrepancy. Instead, the systematic uncertainty associated with the tagging ability must

be taken into account. The tagging efficiency is determined with two methods. The first method involves using the simulated S2s which are fragmented to calculate the true-positive fraction since this dataset does not contain any of the other topologies that may appear to look like fragmentation. The results from the simulation indicate a 96.4% tagging efficiency. However, this method fails to account for false positives where other event topologies may be tagged as fragmented. This can currently only be achieved by using the data that has been tagged as fragmented and manually assessing each event window to avoid circular logic. After manual inspection, due to the observation of false positive tags, the purity of correctly tagged fragmented events is 94%. Applying the 94% purity to the 3.25% of fragmented events observed in data brings this value down to $3.05 \pm 0.03\%$, which is in direct agreement with simulated expectations. This provides sufficient evidence that this tag is efficient enough to be used in data.

6.4.2 Post-Processing Reconstruction

All tagged events should then be reconstructed to reform the true S2 RQs. In practice, this is implemented as a single algorithm that can be attached to any analysis module within LZ's framework and is generally applicable to all analyses that observe the fragmentation effect (have S2s less than $50\,\mathrm{electrons}$). All analysis performed by LZ at this point is carried out in RQ space; that is, the total waveforms are not needed, just the RQ variables we would extract from the total S2.

Some RQs are simple to account for, most notably, the pulse area. This has a few denominations in the RQ list since it can be split into top area (the amount of light collected in the top array) and bottom area (the amount of light collected in the bottom array), as well as positive and negative areas (above or below the baseline). These are simple to reconstruct because they can be easily summed to find the total area. The new areas can then be used to evaluate things like top-bottom asymmetry defined in Equation 4.1.

Reconstructing the total width of the S2 requires a more complex algorithm since the waveforms themselves are not saved. The approach taken is to approximate the pulses as statistical distributions such that one can combine them and re-evaluate the quantiles of the total distribution, which are precisely the AFT times. We must first identify the distribution parameters for each pulse. For large S2 pulses, it is known that the pulse shape fits well to a Gaussian distribution $\mathcal{G}(\mu, \sigma)$; however, small S2 or SE pulses often resemble a boxcar distribution, depicted in Figure 6.21, given

by

$$B(x, a, b, A) = A[H(x - a) - H(x - b)], \tag{6.2}$$

where A is the amplitude of the distribution, H is the Heaviside step function, and a and b are the x coordinates that define the start and end points of the box, respectively. This distribution works since we can safely approximate the production of EL photons to be constant for the time in which an electron is traversing the gaseous xenon, at least up to 96% of the distance as indicated in Section 4.4.

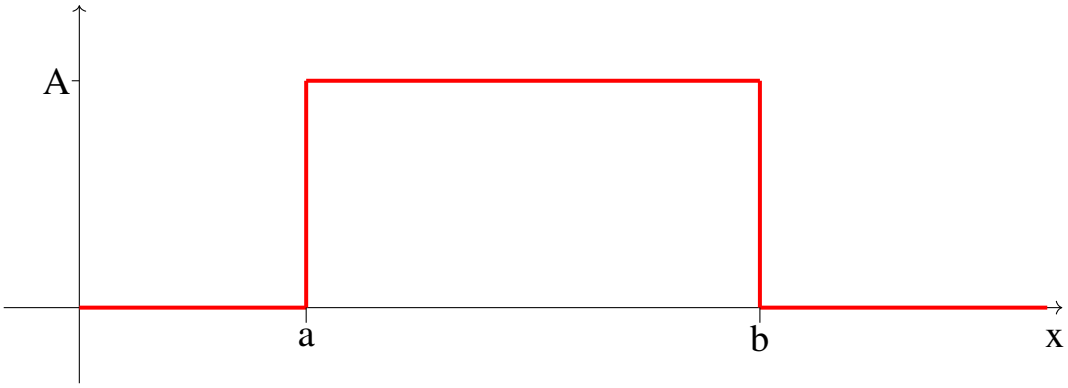


Figure 6.21: Simple diagram of a boxcar distribution as defined in Equation 6.2.

In practice, it is best to fit each fragment pulse to both distributions and evaluate the goodness of fit using a simple mean squared error for each, to decide which distribution to use. To find the distribution parameters without the waveform, (μ, σ, a, b), we use the AFT quantiles of each pulse to estimate the parameters for both distributions. We first examine how this is accomplished for Gaussian pulses.

For percentiles of Gaussian pulses, we can employ the relation that

$$x_p = \mu + \sigma z_p, \tag{6.3}$$

where x_p is the x value at the p^{th} percentile, μ and σ are the mean and standard deviation of the Gaussian pulse, and z_p is the p^{th} standard normal quantile taken from $\mathcal{G}$(0,1). The values for z_p are pre-determined for all percentiles. Translating this into the context of a Gaussian S2 pulse, the x_p values are the AFT values of the pulse, which are known for $p \in \{1, 5, 10, 25, 50, 75, 90, 95, 99\}$, providing us with 9 data points to perform a least-squares regression fit to Equation 6.3, returning the parameters μ and σ.

Similarly, for boxcar pulses we can employ the relation that

$$x_p = a + (b - a)p, \tag{6.4}$$

where all variables are as previously defined. This provides a simple linear fit once again returning the parameters a and b for the boxcar distribution in Equation 6.2. Note that the distributions for both the Gaussian and boxcar functions can be normalised here to remove the amplitude. In reality, the amplitude of either pulse is irrelevant at this stage since we only require the width of the pulse to be reconstructed.

Following this fitting, we can construct a combined cumulative distribution function (CDF), $\mathcal{F}$, weighted by the pulse areas of all fragment pulses, including the prominent S2. For simplicity here, we will take the case where the prominent S2 is best described with $\mathcal{G}(\mu, \sigma)$ and associated CDF $\Phi((x - \mu)/\sigma)$ and the fragment pulse is a single electron described by the boxcar, $B(a, b)$, with associated uniform CDF, $U(a, b)$. This can be written as,

$$\mathcal{F}(x; \mu, \sigma, a, b, w) = w\Phi\left(\frac{x - \mu}{\sigma}\right) + (1 - w)U(x; a, b), \tag{6.5}$$

where $w = A_{pS2}/A_{tot}$ is the weight of the prominent S2 pulse contribution relative to the total weight of all fragment pulses. Once combined in this manner, the new AFT parameters are the solutions to this combined CDF where $\mathcal{F}(x_p) = p$ for each value of p.

Since these are approximations of the pulse shapes, we must validate the algorithm's accuracy. In this instance, the simulated S2 dataset was produced using waveforms so that we can test this algorithm on the waveforms for the subset of fragmented S2s. One example is shown in Figure 6.22, where 2 SE pulses were fragmented from the prominent S2. The original AFT95 - AFT5 width is shown on the prominent S2 pulse. The reconstructed AFTs from this algorithm are given by the shapes at the base of the pulses. Since we have access to the waveform in this case, we can also integrate over it to identify where the real AFT95 - AFT5 boundaries lie, as shown in the green-shaded region of the pulse in the figure. This example demonstrates that the true integrated width visually agrees with the expected widths from this algorithm. This fragmented S2 further exemplifies the impact that this mis-reconstruction can have on the widths of the pulses. The original width is less than 30% of the true width of this pulse.

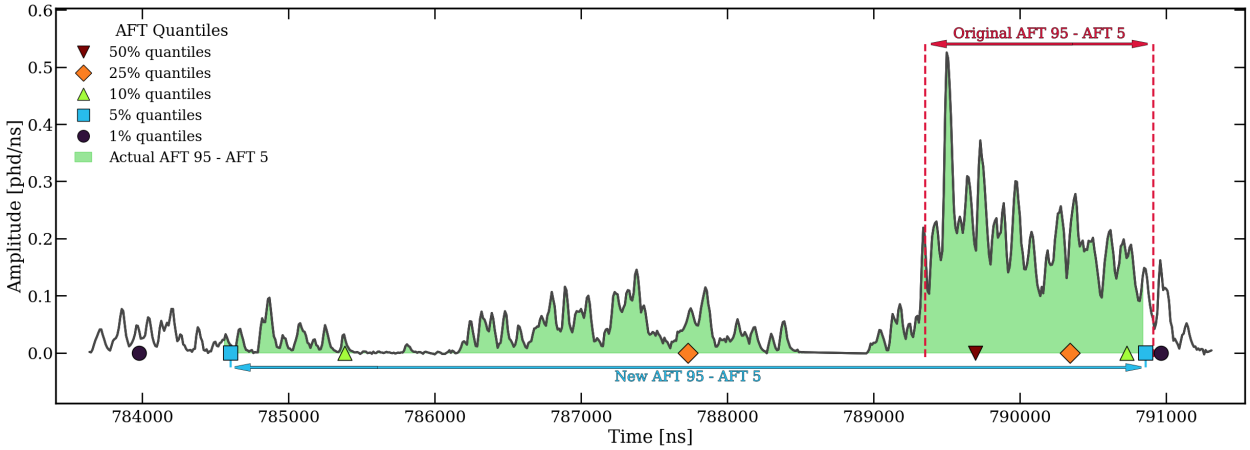


Figure 6.22: Extreme example of the fragmentation effect on the width of a simulated S2 pulse. The 'Original AFT95 - AFT5' width indicated here shows the width of this pulse taken directly from LZAP classification of the prominent S2. The new quantiles derived from the post-reconstruction processing algorithm are indicated along the bottom of the pulse. The 'New AFT95 - AFT5' width according to this reconstruction is given by the blue arrows. The shaded green area of the pulse is a result of the integrated width of the waveform.

By considering the difference between the integrated width and the reconstructed width of the total pulse for the entire population of these simulated fragmented S2s we can evaluate the accuracy of this reconstruction. The results of this are shown in Figure 6.23 and demonstrate an average deviation from the integrated width of 0.7% of the integrated width, with a maximum deviation of 10% above or below. In future iterations, we may continue to add other distributions, such as exponentially modified Gaussians, which may provide better fits to the pulse shapes and thus yield more accurate width estimates, further reducing this error margin.

Furthermore, applying this width reconstruction to the population of 3–4 electron S2s we can observe how this impacts the distribution of widths with drift time. It is known in TPC technologies that small S2s have a larger width variance at the bottom of the detector due to diffusion. In general, one would expect S2s originating from longer drift times to produce wider pulses, however, for small S2s they can remain consistent with the diffusion from the short drift times. Figure 6.24 indicates that this concept could be partially resolved by accurate reconstruction. Since these small S2s from the longer drift times are more likely to be fragmented, it is more likely that their apparent widths have been mis-reconstructed to be smaller than expected. Correcting for these shows a relationship between widths and drift time more reminiscent of the larger S2 pulses. This becomes increasingly important for S2O searches where the width of the pulse is used as a proxy

for depth.

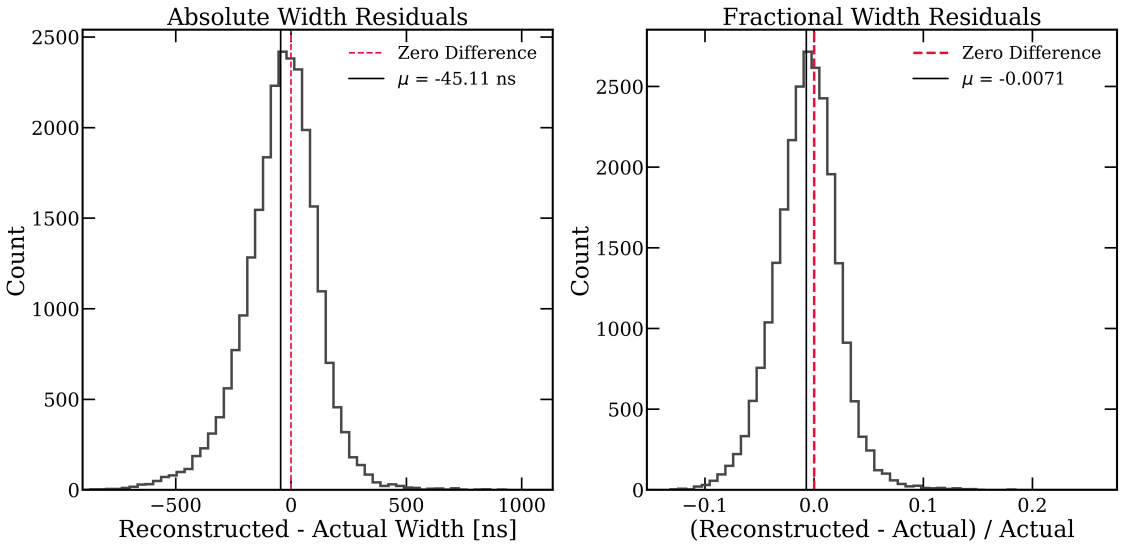


Figure 6.23: Absolute and fractional residuals from the reconstructed AFT95 - AFT5 widths and the actual integrated widths from the simulated waveforms.

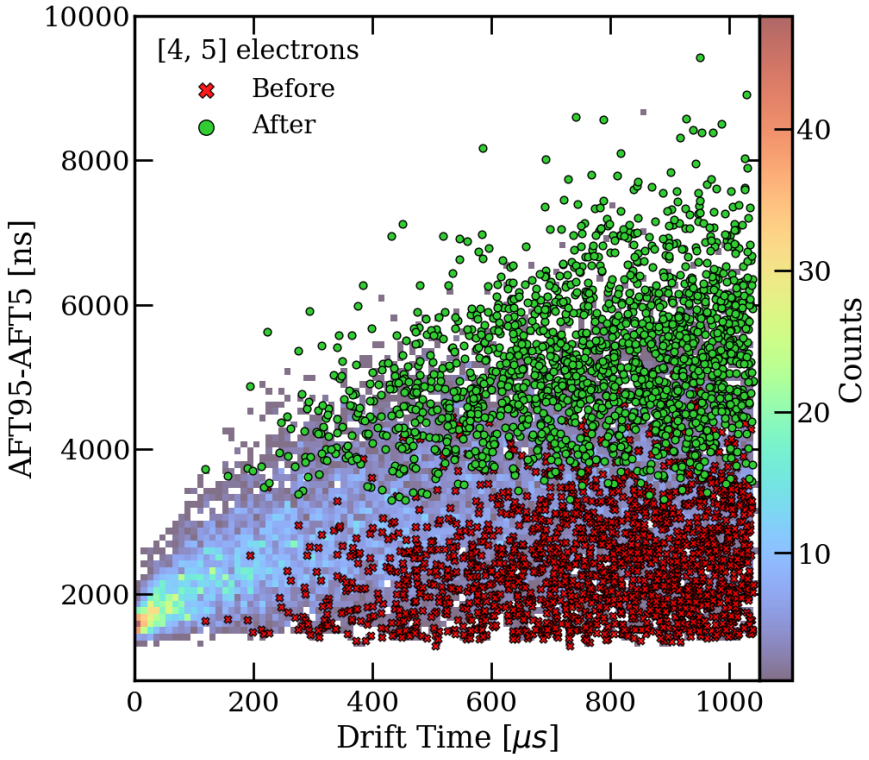


Figure 6.24: Impact of the width reconstruction on a distribution of simulated S2s of 4–5 electrons. The heat map shows the total distribution, and the red crosses indicate the previously fragmented widths that then become the green dots as the corrected widths, now conforming to the expected drift time-width relationship.

In principle, we could go further and completely reconstruct all RQs for these pulses. In practice, it is only required that we reconstruct RQs that are either used in the statistical analysis or in cuts to define the final dataset. All event-level RQs will remain unchanged by this reconstruction; hence, we only need to check which cuts are defined using the S2 pulse RQs. For these, we require the pulse area (in its various forms), TBA, AFTs and (x, y) position. The (x, y) position will not change by anything more than $10\,\mathrm{cm}$ when recombining the fragmented pulses since they are chosen to be from similar (x, y) positions by definition. Because of this, it is acceptable to retain the prominent S2 (x, y) position. Therefore, we have accurately reconstructed all RQs relevant to analysis and can apply this algorithm to the S2O and ^{8}B CEνNS analyses to assess the impact that this has.

6.4.3 S2-Only Impact

This algorithm is applied to the S2O search to retain as much signal as possible while also gaining discrimination power against the gate and cathode backgrounds, due to the more accurate S2 widths produced by the reconstruction. The first check is to observe how the dataset changes when applying all preliminary cuts before and after the reconstruction, particularly how many events are lost or gained and where they lie in the S2 area vs the RMS width parameter space. The final effect of this is shown in Figure 6.25, which depicts that 128 events are gained and no events are lost when applying this tag and reconstruction algorithm. The majority of the events gained by the reconstruction have large RMS values and small pulse areas. This is due to two effects: firstly, since most of the fragmentation occurs at the bottom of the detector, the pulses will naturally be broader and have larger RMS widths. Secondly, the inclusion of the additional fragmented pulse will increase the pulse area above the threshold, allowing these events to become part of this region of interest.

A small number of these gained events are now included following the re-evaluation of the data-quality cuts using the new RQs. Figure 6.26 shows how the quantity of events passing each cut changes once the events are reconstructed. The figure shows large increases in the number of events passing the lower bound (S2 $> 3.5\,\mathrm{electrons}$) and fragmentation cuts. Additionally, based on the correlation between the fragmentation cut and the remaining cuts applied here, there are significant changes in the S2 shape cuts. It is expected that this change will produce a more

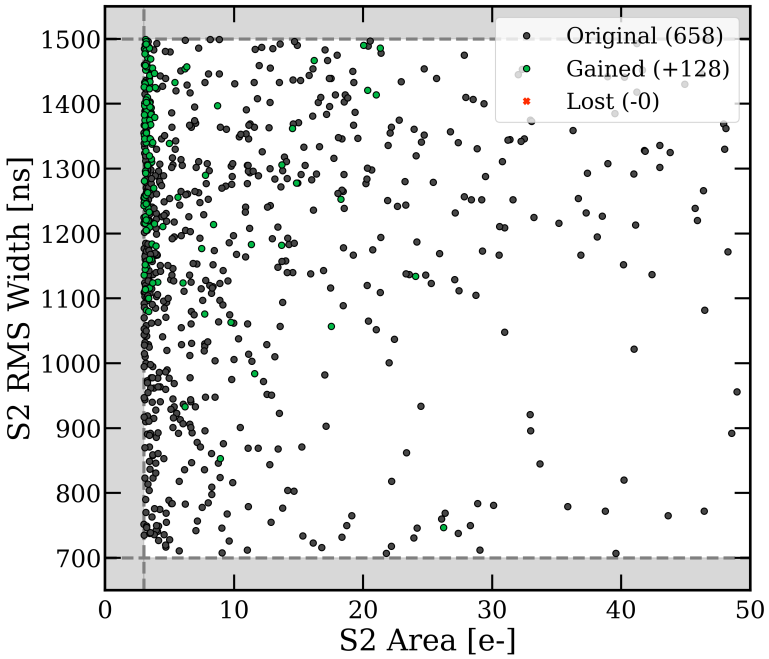


Figure 6.25: Effect of reconstruction on the final S2-only dataset demonstrating the gain and loss of events in this parameter space.

accurate data distribution, although these cuts may require re-tuning in this new parameter space.

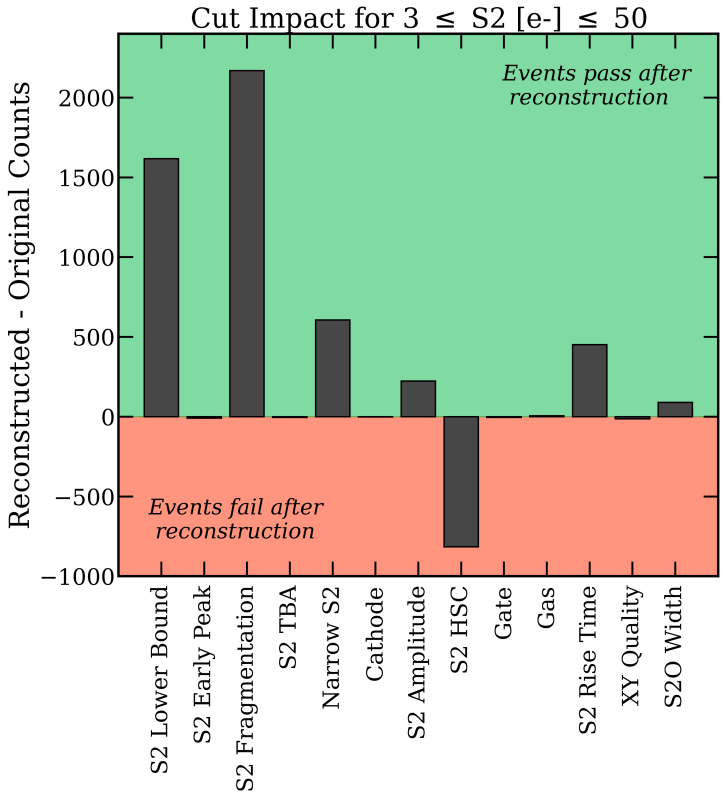


Figure 6.26: Impact of the reconstruction algorithm on the evaluation of all data quality cuts for the S2-only analysis.

At the time of writing, the statistical analysis framework is under construction; therefore, the impact of this on the sensitivity of the S2-only search cannot be evaluated. However, since this parameter space now retains more events and should inherently have greater discrimination power against the main cathode background, re-tuning the cuts in this space will likely improve sensitivity to low-mass dark matter as well as ^{8}B CEνNS neutrinos.

6.4.4 ^{8}B CEνNS Impact

This reconstruction algorithm is then applied to the ^{8}B CEνNS analysis. The main difference in applying this reconstruction algorithm as opposed to the cut is that this has the ability not just to remove further events from the final dataset but to transform which events remain in the final dataset. Correcting RQs can cause some events to pass cuts they previously failed, and vice versa. The first stage in this analysis is to investigate how the final dataset is transformed. This is shown in Figure 6.27, where only the two events that were previously removed by the fragmentation cut are now reconstructed to their correct positions in S2 space. No additional events were gained or removed, which is due to a strong tuning of all cuts applied in this analysis.

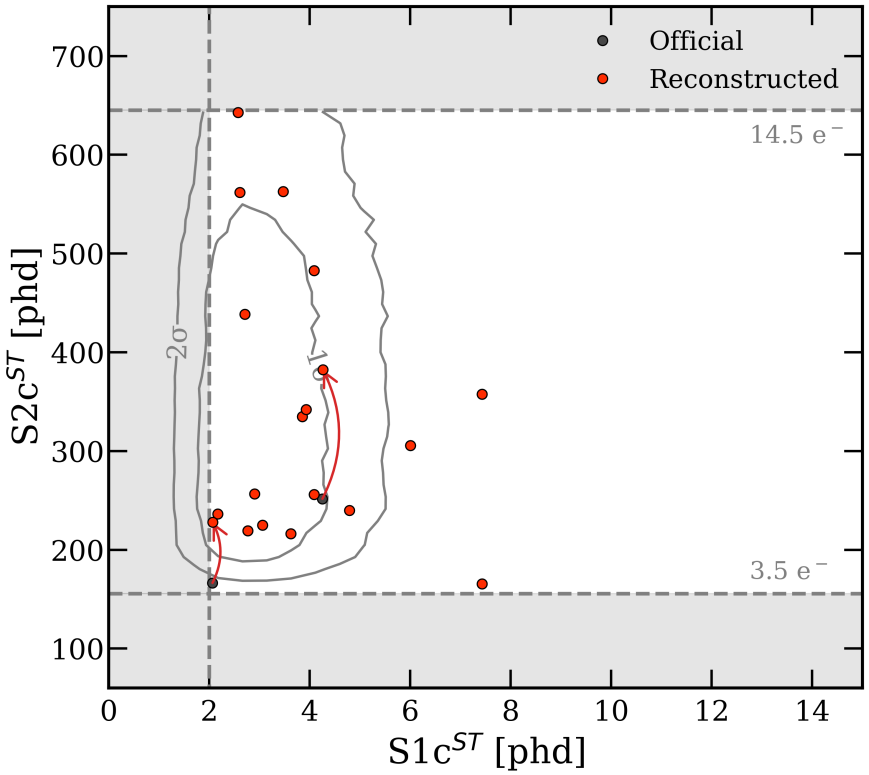
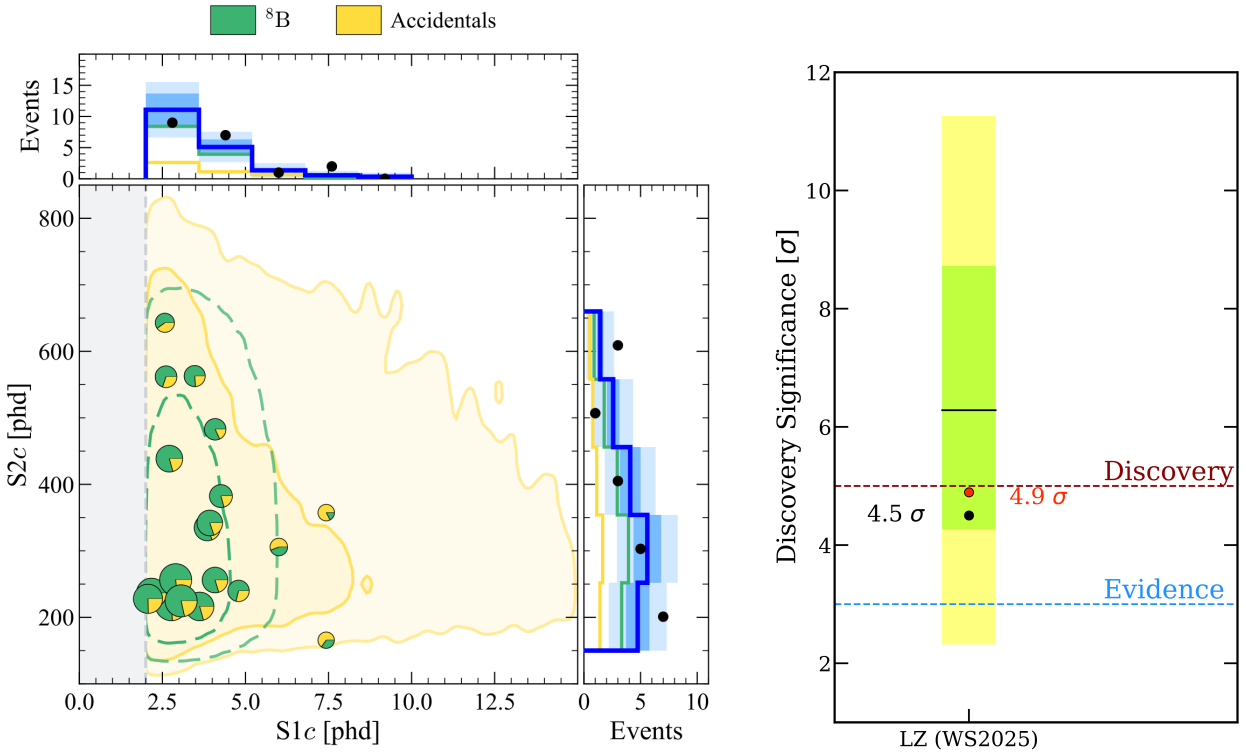


Figure 6.27: Two events transformed into new corrected S2 space; all other events remain in their previous positions. No new events were added to the dataset, and no events were taken away.

Additionally, this reconstruction algorithm is applied to the accidentals production to investigate the impact on the predicted number of accidentals and the overall distribution of these events. Recall that the published analysis predicts 6.55 accidental events should make it into the final dataset. Applying the reconstruction algorithm decreases this value to 5.86 events. This is due to the S2 shape cuts now being more effective, as many rely on the correct S2 shape relative to the drift time because of the diffusion effect. Finally, the 96.4% tagging efficiency is applied to the signal model PDF. The statistical inference is repeated with the new inputs, resulting in an increase from the official ^{8}B CEνNS significance of 4.5σ to the improved 4.9σ, shown in Figure 6.28b, with a best fit value of $13.14^{+6.92}_{-5.13}$ ^{8}B CEνNS interactions observed.



(a) Distribution of the reconstructed events in the final dataset. Each data point is represented by a pie chart showing the likelihood of each event belonging to each model. The size of each data point represents how ^{8}B-like the event is.

(b) Increased observation significance for ^{8}B CEνNS neutrinos interacting with liquid xenon caused by reconstructing fragmented S2s.

Figure 6.28: Final impact of the reconstruction algorithm on the ^{8}B observed significance.

A few noteworthy clarifications must be considered here. This result does not constitute a new result for LZ. However, we should state that although this result follows after the un-salting procedure and hence is applied to an unbiased dataset, the tagging and reconstruction algorithms were

frozen before application and were tuned entirely on simulation data. Additionally, due to the reconstruction changes, the single salt event present in the original ^{8}B dataset was removed during this analysis, leaving no salt events in the final dataset. A more careful statistical inference study is necessary to fully validate this result.

This study supports the notion that there is a substantial benefit to reconstructing S2s at threshold to correct for the fragmentation. In future, improvements should be made to the LZAP software to catch and reconstruct these effects during waveform handling such that approximations like that of the AFTs in this method do not have to be made and instead can be directly measured from the waveforms. However, as previously stated, improvements to LZAP must be thoroughly validated prior to use, and further re-processing of the data can take months to complete. This algorithm serves as a simple approximation of what can be achieved by reconstructing, without the need to re-process all the data. This can be applied to any upcoming analyses at threshold with no extra computing cost.

To recapitulate, we have identified, characterised and corrected for the effects of fragmentation in low-energy S2s caused by limitations in LZAP reconstruction. Through extensive use of the waveform simulations from the full chain, these effects are now well-understood, and we offer three direct approaches to handling them when producing results. By demonstrating their impact on the ^{8}B CEνNS discovery significance, we firmly conclude that the reconstruction method provides the greatest improvement on the current analysis. Under the nominal statistical model, the illustrative reanalysis pushes the published 4.5σ up to 4.9σ. For future analyses, the S1 threshold criterion will be relaxed to include 2-fold S1 pulses, which, in addition to the improvements from the reconstruction algorithm demonstrated in this chapter, may enhance the discovery sensitivity of the ^{8}B CEνNS interactions with xenon.

7 | Low Threshold S2-Only

Using the advanced modelling of S2 pulses we have established, we now wish to apply our understanding to advance the S2-only search towards the few-electron threshold and below. In this chapter, we introduce the S2-only search, its dominant backgrounds and salting strategy, and then present three approaches that aim to improve background rejection in this regime using not only the RQs from the S2 but also the waveforms themselves. These approaches predominantly target grid events, but the techniques discussed are applicable to other backgrounds and often other searches. The three methods proposed are as follows: categorising Multiple Single Electron S2s with a convolutional neural network (CNN), using a neural network to predict the drift time of an S2, and implementing a UMAP dimensionality reduction algorithm to identify anomalies and increase background rejection. Each of these methods has a series of benefits and drawbacks to their implementations, which will be discussed in the sections to follow.

7.1 S2-Only Searches

The S2-only regime in dual-phase TPC dark matter detectors is used to probe the parameter space for light dark matter masses [95, 96]. This can be achieved because the detection efficiency for S1 scintillation photons is $\sim 10\%$, and the resulting S1 must meet the 3-fold coincidence requirement. By contrast, each electron extracted into the gas produces a large proportional scintillation signal allowing interactions with undetectable S1s to still produce observable S2s [97]. In addition to the challenges that arise in the low-energy regime, as evidenced by the ^{8}B CEνNS analysis, the S2O regime introduces new background topologies in the absence of the S1. The region of interest for the S2O analysis lies between 3.5–$15.5\,\mathrm{electrons}$, almost directly aligned with the ^{8}B analysis since this is still where the signal region lies, as can be seen in Figure 7.1.

7.1.1 Bias Mitigation

To analyse data in the S2O regime, we must ensure a bias mitigation strategy is in place. Since this is the first S2O analysis that LZ will publish, over 300 live-days of data have already been accumulated. Therefore, instead of salting the data with S2O events, the S2O analysis uses a blinding strategy in which we unblind 10% of the WS2024 dataset. All cuts and analysis frameworks are

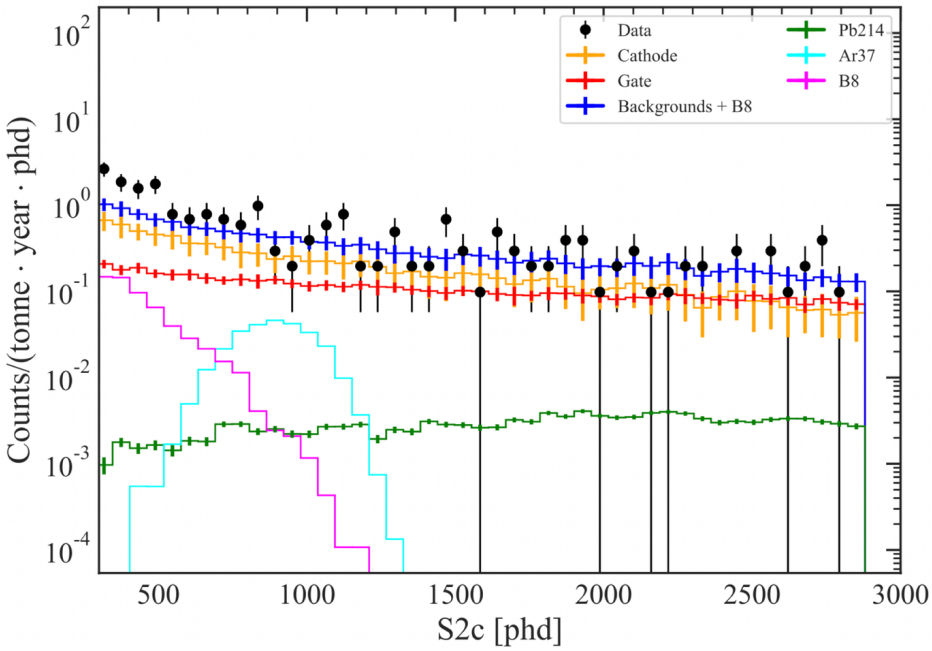


Figure 7.1: Background spectrum for the S2O regime from WS2022 data with 60 live-days of exposure. The largest contributions are from the gate and cathode.

tuned within this unblinded dataset and the AmBe calibration data. In future stages, it is planned to unblind further sidebands of data that are not in the ROI for this search, but at the time of writing this has not occurred. All results presented in this chapter utilise this 10% unblinded dataset and are therefore exploratory until applied to the full dataset.

7.1.2 Backgrounds

The full background spectrum for low-energy S2O searches is shown in Figure 7.1. The largest contribution to the background model is the high-voltage gate and cathode grids. These S2s are formed by both radioactive decays within the grid wires and by spontaneous emission of electrons due to the high voltage applied across these wires [98]. Analyses that include the S1 can remove these backgrounds, since they are localised to drift times of 2.5 µs for the gate and 1050 µs for the cathode. It is likely, however, that these S2s contribute to the accidental backgrounds in those analyses by producing the S2 by spontaneous emission of electrons which are then paired with an S1 from a different source. Tackling these grid backgrounds in the S2O regime will also benefit the S1-S2 analyses by removing accidental backgrounds. The spectrum shown in Figure 7.1 is taken only for the WS2022 science run with 60 days of live time since the full background model for the WS2025 science run is not yet finalised at the time of writing [46]. This spectrum shows a clear

excess in the lowest 4 bins. This is thought to be due to an inaccuracy in modelling the cathode background.

The uncertainty in modelling the cathode background strongly limits the sensitivity projection in the S2O regime. The extent of this on the ^{8}B CEνNS discovery significance is shown in Figure 7.2. The cathode model used to produce Figure 7.1 carries a 12% uncertainty. Improvements to the cathode modelling are in progress, parallel to this work, decreasing this uncertainty to 7%, limiting the ^{8}B CEνNS discovery significance to $\sim 1.5\sigma$. The sensitivity limitation can be removed with two approaches: improving the cathode modelling or improving the cathode background rejection.

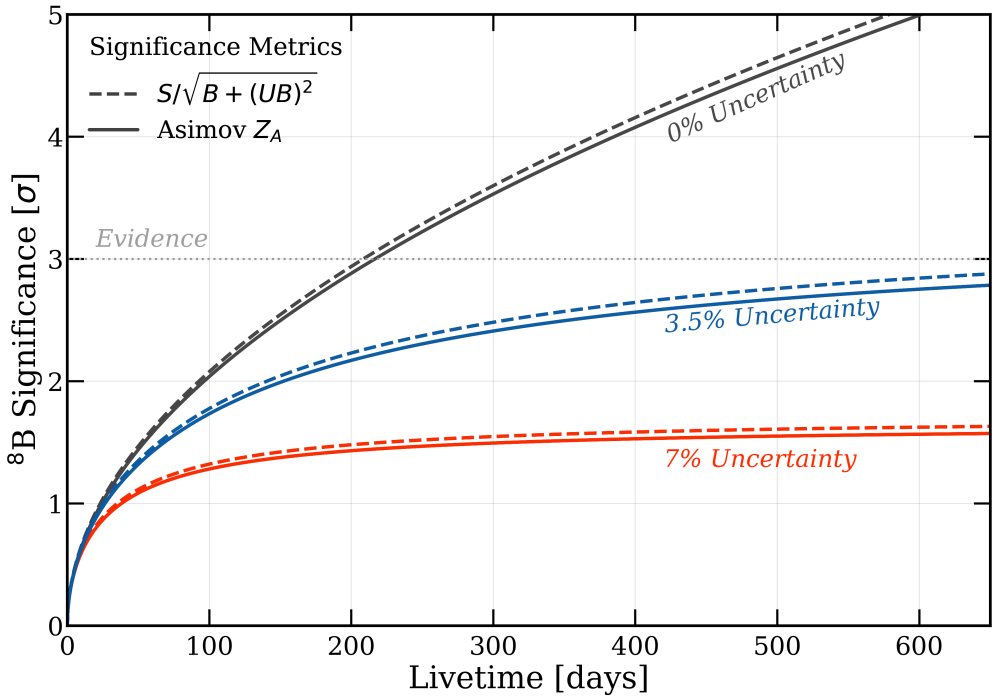


Figure 7.2: ^{8}B CEνNS interaction discovery significance as a function of time for 3 different levels of uncertainty on the background model from the cathode. The significance is evaluated with both a signal-to-background ratio (dashed) and the Asimov test statistic (solid) [99]. The current best estimate of the uncertainty for the S2O regime is 7%.

7.2 Multiple Single Electron S2s

The first approach is to attempt to classify common topologies that are indicative of grid emission. We introduced the concept of the Multiple Single Electron (MSE) S2s in Chapter 6 as an event topology where multiple SE pulses exist throughout the event window that originate from the same position in the detector. This topology is consistent with the spurious emission of elec-

trons from a single position on either grid. Dual-phase TPCs of this nature have been known to observe 'hotspots' in the detector where a single position produces large numbers of SE and S2 pulses for a period of time. However, the physical mechanism by which these electrons are released is not known [98]. Data quality cuts are applied to remove the areas and times in which hotspots occur above an SE rate threshold, but these are extreme cases of detector instability. The spontaneous emission of electrons can occur sporadically, resulting in a single S2 or SE. MSE events are consistent with the idea of a low-activity hotspot that can produce events that reside in the S2O regime, allowing for identification of grid events without the need for an S1 pulse.

7.2.1 Cut Development

MSE events are tagged by identifying events like that of Figure 6.20 which contain an S2 with several SE pulses that are far enough in time to not be fragments of the S2 ($> 5000\,\mathrm{ns}$) but are reconstructed to be from the same position ($< 10\,\mathrm{cm}$). In principle, the more SE pulses that are present in the event window, the more likely the S2 is an MSE pulse rather than a physical S2. Firstly, the relationship between pulse area and number of SEs present in the time window must be characterised. This is shown in Figure 7.3 to have a large excess of events below the analysis threshold of $3.5\,\mathrm{electrons}$, with up to 180 SE pulses in the same event with the same (x, y) position in the TPC. Above the analysis threshold, there is an exponential decay down to a constant background of MSE events with fewer than 10 SE pulses in the event window. Note that the electron train veto was applied to this dataset, since electron trains can directly produce multiple electrons within the event window and must be removed to reduce contamination from grid events. No further cuts were applied to retain statistics for the grid events dataset.

For future analyses, a cut on the number of SE pulses in the event window would be able to remove a large number of background events from the grid wires. Figure 7.3 contains 5.4×10^5 events below the analysis threshold for this dataset, which is just 10% of the total dataset which has been unblinded thus far. Above the analysis threshold, there exist ~24,000 events that can be removed with a simple cut on the number of SE pulses allowed in the event window. The proportion of the MSE events removed by varying the threshold for an acceptable number of SE pulses is shown in Figure 7.4. Allowing only events with a single SE pulse within the event window is the strongest cut. Yet it removes only 34.4% of events above the threshold, since most of these events have only

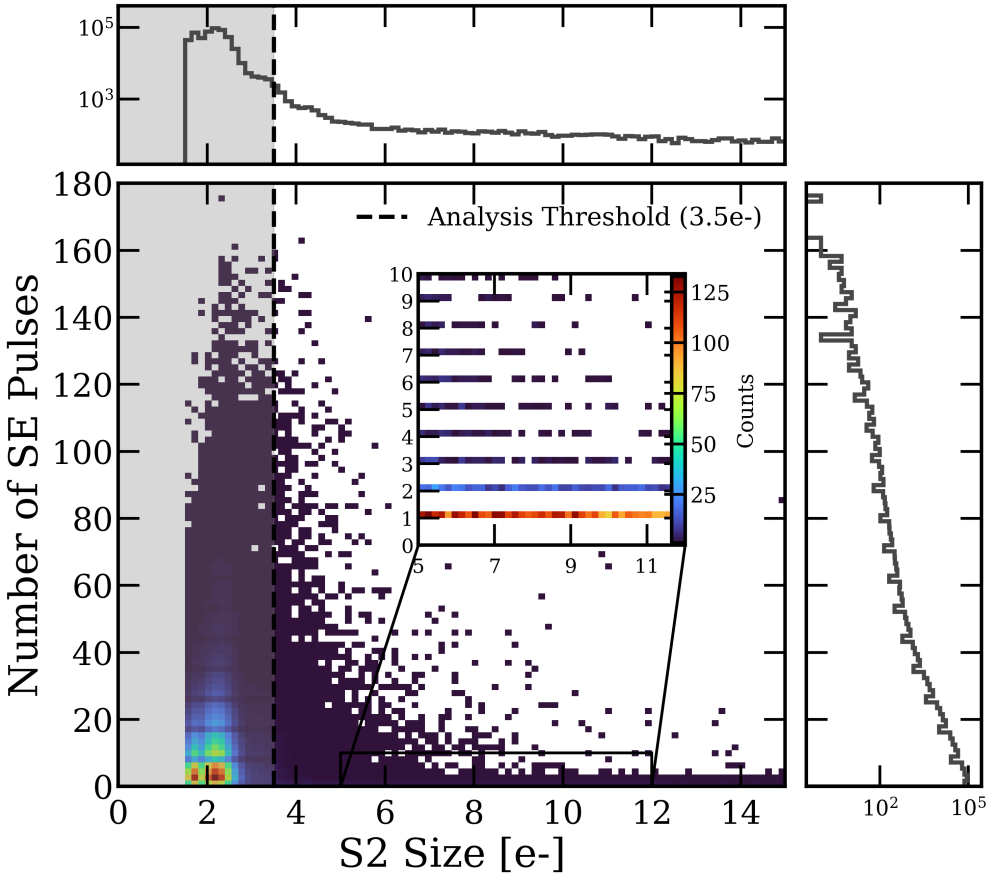


Figure 7.3: Number of SE pulses identified within $10\,\mathrm{cm}$ of the prominent S2 in the event window, further than $5000\,\mathrm{ns}$ away. Most of the MSE pulses are below the analysis threshold of $3.5\,\mathrm{electrons}$. Above this, the most likely outcome is one SE pulse within the event window.

one SE pulse within the event window. In principle, one could extend this to remove 100% of the MSE events by requiring that no other SE pulses are present from the same position. However, by taking a sample of S1-S2 events with the drift times of the cathode and the gate, it is found that only 8.9% of events in either dataset include at least one other SE pulse within the event window. Therefore, even removing 100% of the MSE events only removes 8.9% of the grid events. Furthermore, the majority of MSE events in this selection contain just one SE in the event window. Providing a harsh cut on this space has the potential to also remove large quantities of signal-like S2s.

We propose a pivot on this approach by using this topology to accumulate a high-statistics dataset for studying grid events. This method provides a much larger number of events, a factor of 110 increase, than selecting events with the grid drift times since most of the grid events lie in the S2O regime from spontaneous electron emission rather than radioactive decay.

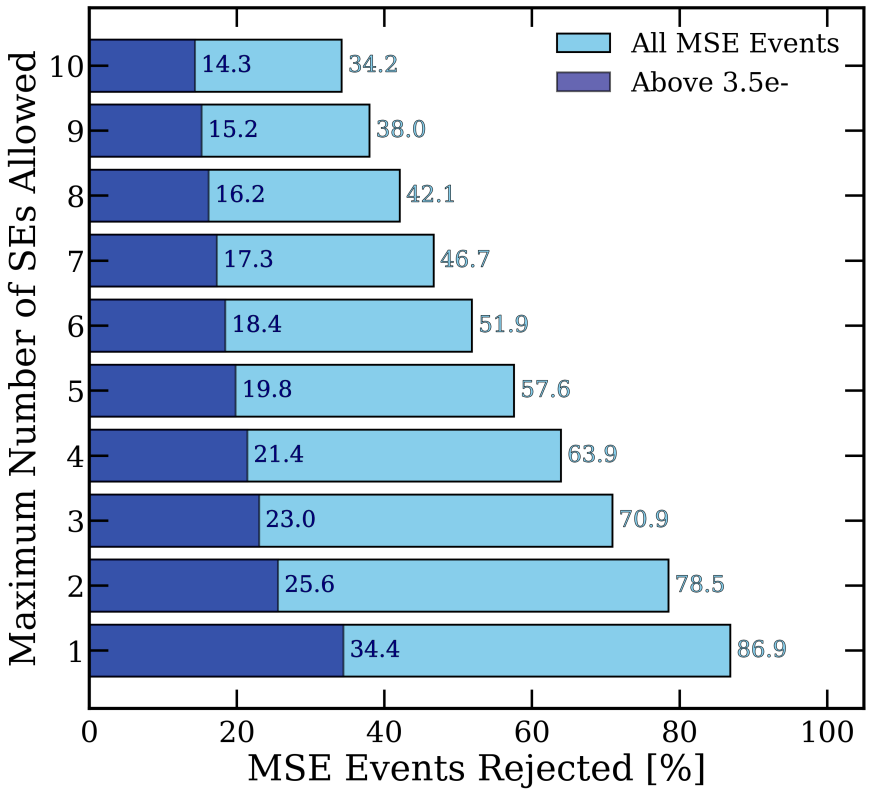


Figure 7.4: Percentage of MSE events removed by varying the threshold for the number of SE pulses in the event window for all MSE events up to $15.5\,\mathrm{electrons}$ (light blue) and just the fraction above the $3.5\,\mathrm{electron}$ threshold (dark blue).

Since we have established that 91.1% of the grid events will not have any other SE pulses in the event window from the S1-tagged events, the direct cut on the number of SE pulses is not as strong as a cut as is necessary for the S2O backgrounds. Instead, this topology is utilised to produce a large dataset of grid events, $\mathcal{O}(10^5)$ events. In the instance where the S2 has no accompanying SE pulses, the only indication of this S2 originating from grid emission or a real scattering interaction is the shape of the waveform. It would be expected that spurious emission from the grid wires will space the electrons out in a non-physical manner. This becomes a difficult background to remove since events from the gate may be spread sufficiently by the time delay to appear like an S2 from the centre or even bottom of the detector. Similarly, electrons emitted from the cathode will diffuse enough that the time delay between subsequent electron emissions is not discernable in the final S2 pulse. For this reason, we propose that the most appropriate method of attempting to distinguish between physical S2s and MSE pulses would be using a CNN trained on the S2 waveforms themselves.

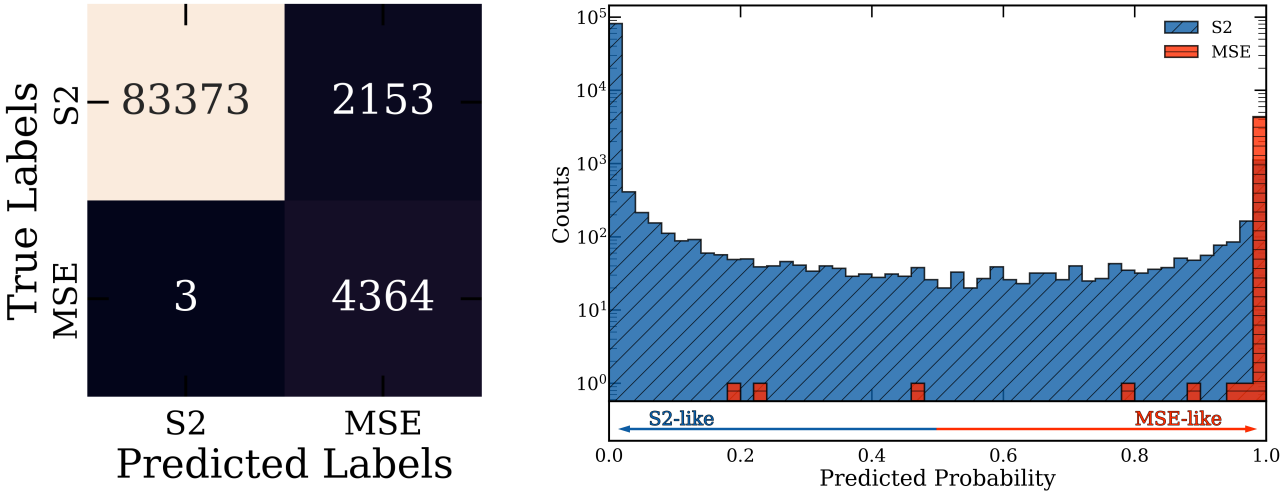
7.2.2 Neural Network Architecture

Prior to producing the architecture for this NN, the training datasets are determined. The MSE waveforms are selected from the unblinded S2O dataset to include any event with an S2 and at least one SE pulse which is within $10\,\mathrm{cm}$ of the S2 but is greater than $5000\,\mathrm{ns}$ away from the S2 to not be confused with fragmented pulses. The only additional criterion is the application of the electron train veto. The signal-like waveforms are taken directly from simulated S2s from the ParticleBomb generator used for the results in Chapter 6. This decision is made to ensure that no MSE or alternative background topologies are present in the training dataset, thereby maximising the potential separation between the signal and background populations. It is noted that it would be preferable to train on a high-statistics set of S2s from AmBe calibrations. At this stage of the analysis, the calibration data is limited to $\sim$1000 events that have waveforms. Future analysers should use a larger sample of AmBe S2s taking the waveforms from both single and multiple scatter events. This is preferable to simulated S2s since it is possible that the NN may learn to differentiate between data and simulation rather than physical vs unphysical S2.

Training is carried out using only the 1-dimensional S2 waveforms to distinguish between the two populations. In future iterations, it may be fruitful to include some RQs in the training data that the model cannot extract from the waveform itself. For example, the (x, y) positions of the S2s. This is not included at this stage due to the risk of overfitting to the hotspot S2s that are concentrated at specific (x, y) positions in the detector.

The optimal network structure for this task is a deep 1D residual CNN. This takes a 1D 1000-sample S2 waveform as input and produces a probability between 0 (signal-like) and 1 (MSE-like). The classification boundary applied is that any probability at or above 0.5 is classified as an MSE-like pulse, and anything below 0.5 is classified as a signal-like S2. The classification matrix is shown in Figure 7.5a with an accuracy of 97.6% and all the predicted distributions are shown in Figure 7.5b to have a clear, distinct separation of MSE-like events with a continuum of S2-like events.

As mentioned previously, the main concern of this approach is that by training with simulated waveforms, it is possible that the CNN can separate these populations by identifying the simulated waveforms from the real waveforms, as opposed to MSE from signal S2s. This concern can be



(a) Classification matrix

(b) Test histogram on simulated S2 (blue) and MSE (red) waveforms.

Figure 7.5: Final results from training the NN with a selection of MSE events and simulated S2s. The NN achieves a classification accuracy of 97.6% and separates the MSE population from the simulated S2s.

evaluated by taking the model, trained on simulated S2s, and testing with the small sample of AmBe waveforms available. The output of this results in 1368 AmBe S2s classified as MSE-like with a mean predicted score of 0.999. Only 2 events in the calibration dataset are classified as signal-like S2s. From this result it must be concluded that this method cannot be applied to analysis without being trained on a larger sample of calibration S2s such that one can be sure that the separation power is a result of the grid events, not the identification of simulated waveforms. Along with adding the RQs to the training data, future iterations should extend from a single-node output to a 3-node output. This would allow the CNN not only to classify the grid from the signal but also to further classify the gate, cathode, and signal. This would resolve the main drawback of using MSE events, which is that the grid from which they originate is unknown since the drift time is unknown.

Although we are unable to apply this approach to data, we can use several advancements from this work to reduce sensitivity limitations in the S2O regime. Namely, we have devised a dataset of grid S2s with a factor of 110 more statistics than is currently used to determine the cathode modelling from S1-tagged cathode S2s. If we can determine the drift time from the S2 waveforms, we may be able to reduce the uncertainty in the cathode modelling.

7.3 Drift Time NN

By moving to the S2O regime and losing the S1 pulse in the event, the position of the initial interactions can now be reconstructed only in (x, y) and is completely lost in z due to the lack of drift time. This is important, especially for the S2O regime, where the grid wire backgrounds could easily be removed with a drift time cut, since they are localised in z. The diffusion of the electrons along the length of the detector has a measurable impact on the width of the S2 pulses. Wider pulses generally originate from deeper in the detector following the relationship derived in Chapter 5, given by Equation 5.3. In principle, with a known diffusion coefficient, it is possible to invert this equation to predict the drift time given the width of the pulse. However, when approaching S2s of less than $20\,\mathrm{electrons}$ the correlation between depth and width is not as strong as for larger S2 pulses. This effect is demonstrated in Figure 7.6 where the variance of the S2 width increases as the S2 size decreases.

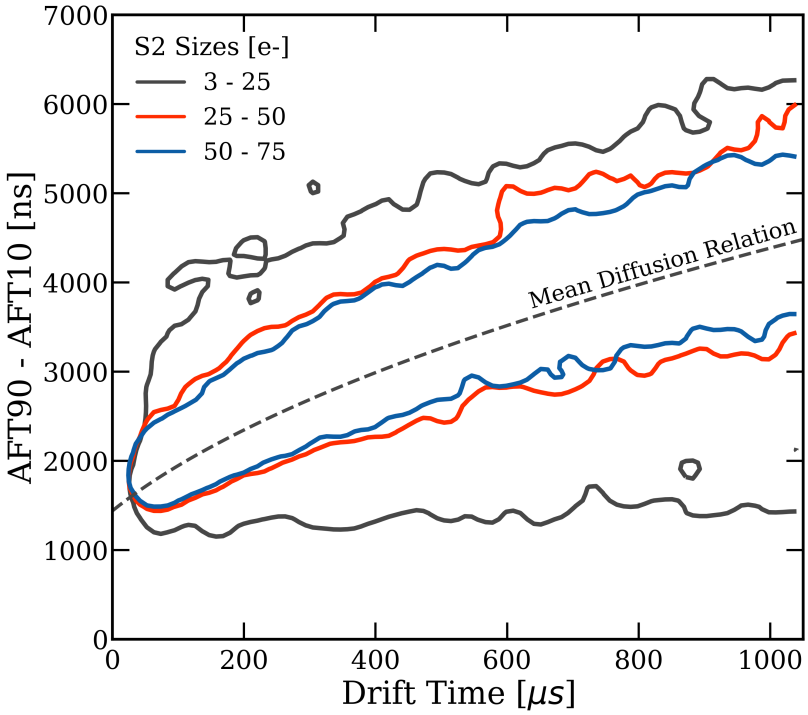


Figure 7.6: Diffusion relation contours shown for increasing S2 sizes. The smallest S2 sizes (3–$25\,\mathrm{electrons}$) show a much larger variance. Above $75\,\mathrm{electrons}$, the correlation becomes slightly stronger, approaching the Mean Diffusion Relation line. A drift time cut is applied to the top $60\,\mu\mathrm{s}$ to remove a large number of gas and gate events which dominate the contours. The general trends shown extend down to $0\,\mu\mathrm{s}$ in drift time.

From this result, it is concluded that, while an approximate drift time may be predicted for larger

S2 sizes, an accurate drift time measurement from the width of the small S2 pulses is essentially impossible. The small widths at long drift times are present for the smaller S2s since the sampling from the diffusion Gaussian is more likely to produce a negligible spreading in time and thus produce a width closer to the SE width despite coming from the cathode. Since this is a physical result, one can envision a scenario in which the S2 waveform can provide a stronger indication of the drift time than the AFT widths alone. This could be due to the substructure of each S2 that percentile time stamps cannot reasonably capture. For this reason, it is worth establishing a neural network that uses the waveform to predict the drift time of the S2. It is also noted here that, since the shorter pulses from the cathode for smaller S2s is a physical effect, it may be impossible to predict cathode positions accurately. However, the second-largest background for the S2 is the gate, which is at the top of the detector and should, in principle, be more accurately predicted by an NN, since it involves less diffusion smearing.

7.3.1 Neural Network Architectures

To train the NN to predict the drift time of events, there are two possible datasets. One can make use of the single scatters data, which contains the drift time from the S1 pulse, or one may use the simulated S2s with a known true z position. The benefit of using the SS data is that it will inherently account for any localised field anisotropies in the detector. The drawback is that as S2 sizes decrease, the SS dataset will be contaminated by a fraction of accidental events in which S1 and S2 are incorrectly paired, leading to an inaccurately estimated drift time. During training, we decided to first test the NN's strength with the simulated S2s. These S2s are generated under perfect-run conditions, meaning that if the NN cannot train to identify the simulation's drift time accurately, then it will be impossible to train on real S2s, which contain more perturbations and variation due to detector conditions.

Multiple architectures were attempted for this result, namely, S2 RQs through a multi-layer perceptron (MLP), a 1D CNN trained on just the waveform, a combined CNN and RQ dual pipeline, and finally, the most powerful, is a Feature-wise Linear Modulation (FiLM) residual CNN, which will be referred to as the FiLM ResNet. A direct comparison of the training loss for a small subsample of the training data for all 4 approaches is shown in Figure 7.7, demonstrating the immediate gain of including waveforms in training as opposed to just RQs, as well as the increased speed and

accuracy of training with the FiLM ResNet.

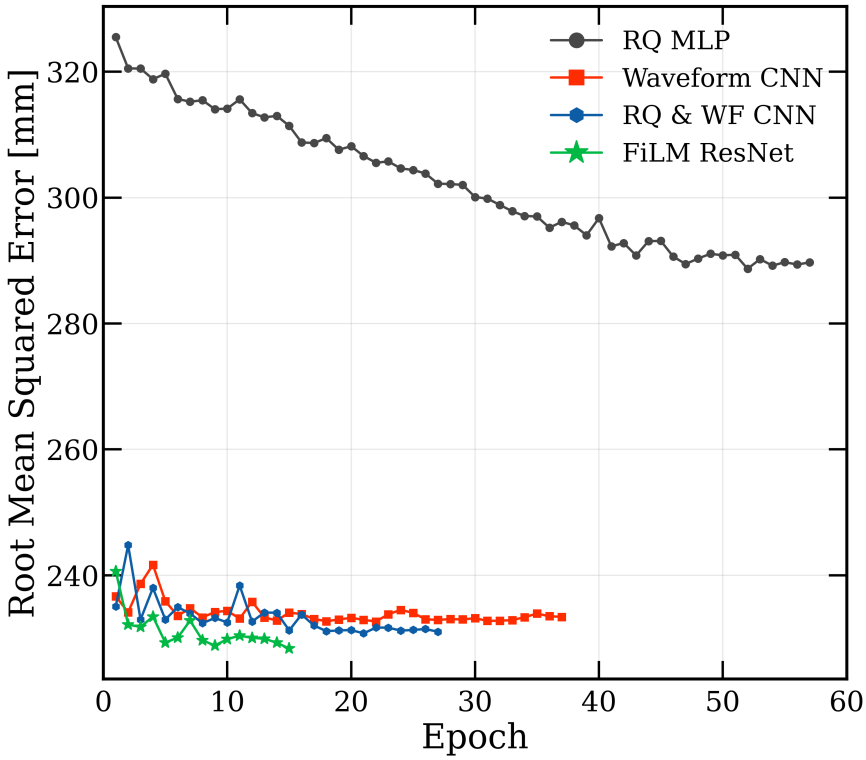


Figure 7.7: Comparison of the root mean squared error of each training epoch for each architecture approach. A significant improvement is achieved by training on the waveforms using a CNN. The FiLM ResNet achieves the best performance, with the smallest error and the shortest training time to reach stability.

This is most powerful because the FiLM approach allows the NN to distinguish between different trends in the pulse area and radius. The pulse area, as demonstrated in Figure 7.6, impacts the strength of the relationship between the waveform shape and the drift time. Similarly, the radius for a given number of electrons will cause the S2 pulse area to change slightly due to the radial field introduced by the deflection of the gate and anode grids. The FiLM ResNet forces the NN to learn these separate distributions by modulating the waveform using the available RQs. The complementary approach is the dual RQ and waveform CNN pipeline, which proves less powerful because it trains separately on RQs and waveforms and combines them only for prediction. In contrast, the FiLM ResNet enables synergised training that leverages all available information.

7.3.2 Results

The results of training the FiLM ResNet with simulated S2s uniformly generated between the cathode ($z = 0\,\mathrm{mm}$) and the gate ($z = 1456\,\mathrm{mm}$) are shown in Figure 7.8.

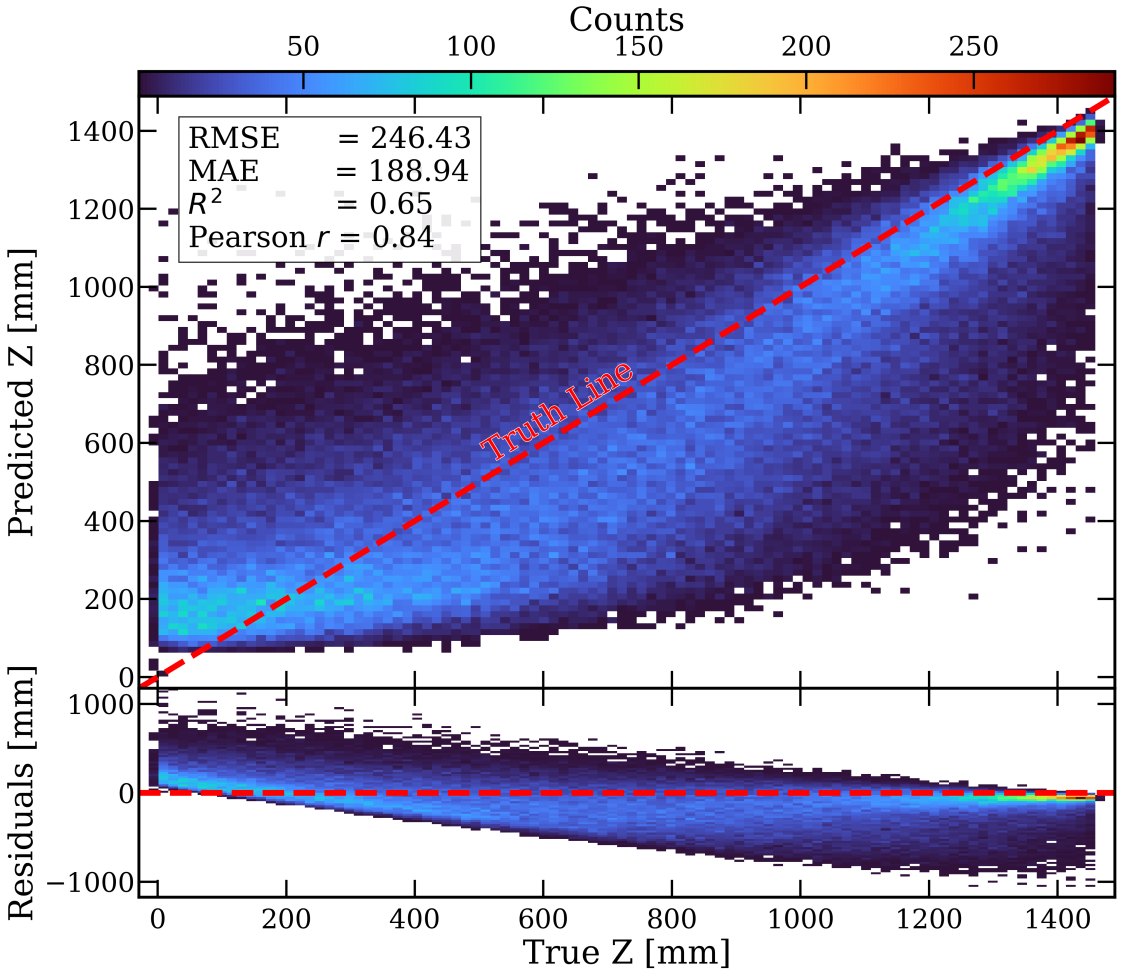


Figure 7.8: Predicted z height of an S2 against its true z height for the FiLM residual CNN trained on simulated S2 waveforms. Four regression loss metrics are evaluated in the top left corner, where RMSE denotes the Root Mean-Square Error; MAE denotes the Mean Absolute Error; R^2 denotes the coefficient of determination, and the Pearson r is the Pearson correlation coefficient. The residuals of Predicted z - True z are shown in the bottom panel.

This output indicates that the relationship between the predicted and true z position is not strong enough to make any reasonable predictions of the drift time of the S2 from the waveforms, with a Mean Absolute Error (MAE) of $188.94\,\mathrm{mm}$. To exemplify this, if an S2 were to be predicted at $z = 600\,\mathrm{mm}$, the true z position could be anywhere in the detector volume. Three distinct features of this figure are of interest. Firstly, there is a saturation of the predicted z at around $100\,\mathrm{mm}$ with no S2s predicted to be generated beneath $100\,\mathrm{mm}$. This is thought to be the result

of range compression, in which the model tends to regress toward the mean at a boundary, pulling the z prediction upward. Secondly, there is a strong correlation near the top of the detector. This is expected because there is less broadening due to diffusion, even for small S2 sizes. However, the correlation is biased underneath the truth line, meaning the network is consistently under-predicting the drift time. This is also likely due to the range compression issue. Lastly, the residuals display a near-linear dependence on the true z value. This inspires the opportunity for a post-training calibration that may be able to correct for the range compression issues at the top and bottom of the detector and improve the prediction power of the network.

Four calibration methods were applied to this result. The strength of the improvement was determined by comparing the RMSE values from before and after the calibration. The four approaches are detailed in Table 7.1.

Method	Function	Description
Linear	$z_c = Az_p + B$	Linearly apply corrections to each z_p value. The same correction is applied to all values.
Bin	Piecewise $z_c = f(z_p)$	Splits the data binwise and applies a different correction factor to each bin.
Isotonic	$z_c = f(z_p)$	Minimises the total squared error between true and predicted z by applying correction factors in a strictly increasing monotonic function.
GBDT	$z_c = f(z_p, RQ)$	Gradient-Boosted Decision Tree (GBDT) which utilises both the z_p values and the RQ features to learn a non-linear calibration.

Table 7.1: Post-training calibration methods applied to Figure 7.8 where z_p is the predicted z position and z_c is the calibrated z position.

The strongest approach is the Gradient-Boosted Decision Tree (GBDT), which decreases the RMSE value from $245.64\,\mathrm{mm}$ to $221.68\,\mathrm{mm}$ as shown in Figure 7.9. This is not a significant improvement on the nominal result. However, the figure shows that the calibration resolves the misalignment at large z values, with the strong correlation now falling directly on the truth line. Therefore, it should be expected that if the correlation present between predicted and true z were stronger, even if incorrect, it could be corrected with post-hoc calibration techniques like this. The

calibration cannot overcome regression to the mean at low z values. Fixing this would require sensitive handling of low z S2s during the training of the FiLM ResNet.

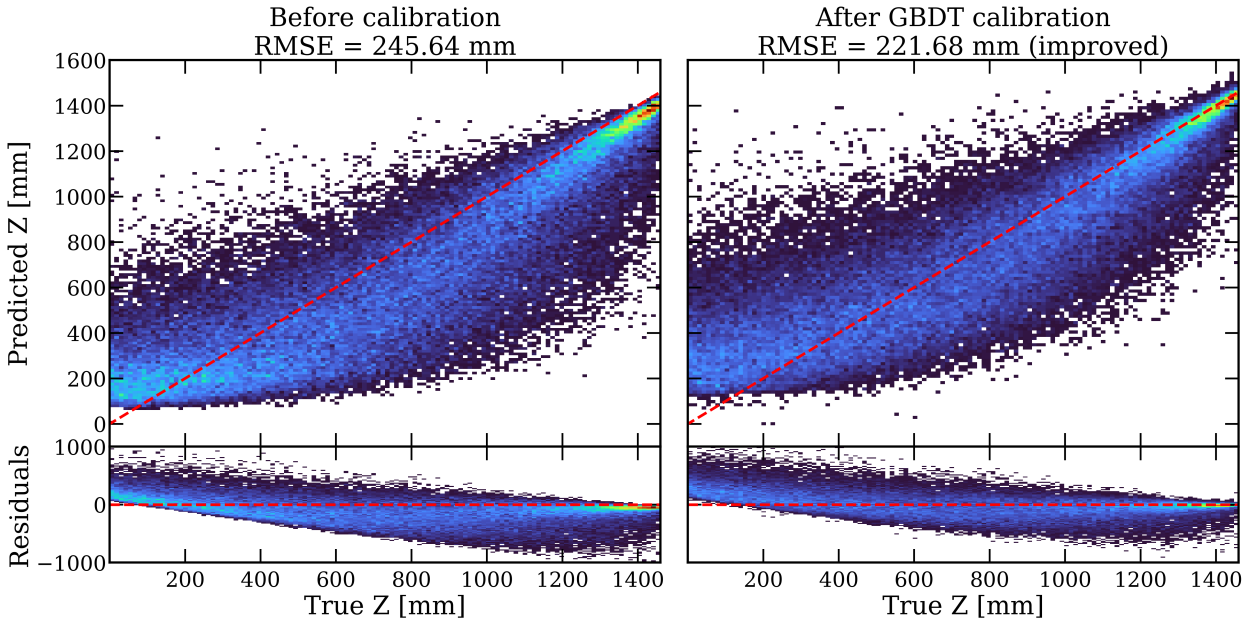


Figure 7.9: Comparison of the original and post-calibration outputs from the FiLM ResNet. The largest improvement came from the GBDT approach, reducing the RMSE by $\sim$24 mm. The calibrated output does not show a significant improvement over the standard output, except for a better-calibrated upper edge.

The results of this study demonstrate the level of sophistication that would be required to accurately predict the drift time of the S2 without the presence of the S1 pulse. Despite training a model on over 6×10^5 simulated waveforms with a sophisticated network structure, a predicted z value of 400 mm could still translate anywhere from 0–1300 mm even after post-hoc calibrations with a GBDT. One could continue to iterate on the network architecture, but it would require a significant increase in computing power to train such a model, which is already reaching the memory ceiling available with the NERSC Perlmutter supercomputer on a dedicated GPU node [94]. Although the possibility of a drift-time-predicting NN is enticing, this approach is not pursued further because it is too expensive in both time and computational power for a potential null result. Analysers should take alternative approaches in the future.

7.4 Anomaly Detection

The final method of improving the sensitivity of the S2O search is to utilise the Uniform Manifold Approximation and Projection (UMAP) dimensionality reduction technique to isolate anomalous data points or increase cut efficiencies against targeted backgrounds [100]. One difficulty with performing analysis in the RQ space is that each S2 pulse has over 100 RQs derived from it. To identify background topologies or anomalous events, analysers are required to investigate all relationships among these RQs across the entire population of events. UMAP reduces the dimensionality of this data by using similarities between events to construct a lower-dimensional projection that is visually more interpretable. A study of this nature has been performed in LZ previously, using the single scatter S2 waveforms to attempt to identify anomalous data points [101]. Here, we expand on this work by moving from SS to S2O events and further by including the RQ and waveforms of the S2s as the training dimensions.

7.4.1 UMAP Theory

UMAP works by preserving the local-topological structure of a high-dimensional dataset to construct an embedding in a low-dimensional space [102]. This is done by constructing a weighted k-nearest-neighbour graph, where the edge weights represent the strength of the local relationship between the points, and then identifying a low-dimensional coordinate system that best matches this graph by minimising a cross-entropy loss between the high- and low-dimensional neighbourhood representations.

The main parameters of this algorithm are the number of neighbours in the graph (`n_neighbours`), which controls the scale of the preserved structure, and the minimum distance (`min_dist`) between points, which controls how tightly the points cluster in the low-dimensional representation. Low values (1-10) of the number of neighbours will force UMAP to preserve immediate local structure, while large values (100-200) will allow UMAP to preserve the larger trends of the entire dataset. Small values of the minimum distance between points (0.1) force the positions of the low-dimensional embeddings to cluster more tightly together, which is useful for preserving the finer topological structure. Large minimum-distance values (e.g., 0.99) are useful for preserving the broader topological structure. For this application, it is useful to have a larger number of

neighbours (50-75) in the graph to preserve broad trends in the data, while also setting a smaller minimum distance (0.0) between those neighbours to achieve strong clustering of S2s that are alike.

In addition to these parameters, one can also enforce the metric used by UMAP to measure distances between points in each graph. The only two metrics which were used throughout this analysis were the Euclidean and Cosine metrics. The Euclidean metric is the standard distance between two points, calculated in an N-dimensional space by

$$d(X,Y) = \sqrt{\sum_{i=1}^{N}(x_i - y_i)^2}, \tag{7.1}$$

where $X = (x_1, x_2, ..., x_N)$ and $Y = (y_1, y_2, ..., y_N)$ are vectors in the N-dimensional space. The Cosine metric is used to define the similarity between two vectors, ranging from -1 to 1, defined as,

$$d(X,Y) = 1 - \frac{X \cdot Y}{\|X\|\|Y\|}. \tag{7.2}$$

This metric is very useful for clustering waveforms that are similar in shape and size.

The final parameter is the dimensionality of the low-dimensional representation. While all plots shown in this section will be 2D, the manifolds can be constructed and minimised in a higher-dimensional space to preserve more local structure and prevent oversimplification of the latent space topology.

UMAP provides the functionality to combine two maps trained on different input types for the same events, which for this study were the RQs and waveforms. These are trained on two separate maps, such that events are clustered together in the RQ space and the waveform space separately, and can then be combined afterwards. Throughout this study, many combinations of mappings were investigated. The combination operators function as graphical operators where each data point in each graph holds a coordinate and an edge to any other data points with which there is some similarity. These can be applied in three formats:

- **Union**

  – This mode acts as an `OR` operation, maintaining relations which exist in either of the two graphs.

- **Intersection**

  – This mode acts as an `AND` operation, only retaining the edges which are present between points in both graphs.

- **Contrast**

  – This mode effectively acts as a `NOT` operation, retaining edges that are connected in one map but not the other. This is the only operation that is non-commutative and is thus represented as RQ - WF or WF - RQ to denote the two possible combinations.

Alternatively, one can train on both the RQs and waveforms in a single map by inputting a single stack of all RQ variables with the 1D waveform appended. In total, this allows for 7 different mappings of the same data points, each focusing on clustering based on different criteria and hence are useful for a range of applications.

7.4.2 Clustering Algorithms

Once the positions are embedded into the low-dimensional space, they can be clustered using the Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN) algorithm [103]. This works very similarly to the DBSCAN clustering algorithm used in the full chain simulation; however, the hierarchical nature of this algorithm means that it is not necessary to specify a single locus with which to cluster. This allows for clustering of groups with varying densities. Additionally, the HDBSCAN algorithm, like DBSCAN, will identify any points that do not fit within a single cluster as noise, which can be beneficial for anomaly detection.

7.4.3 Identifying Anomalous Clusters

Primarily, this work aims to identify anomalous populations by training on the entire dataset and identifying clusters that are well separated from the rest. In practice, training on the entire dataset is challenging since there are sufficient obvious outliers that give large differences between the clusters. For example, if Mercury cannot locate the position of the S2, it will assign a default value of $999\,\mathrm{cm}$ for both the x and y values, which is much larger than the maximum $72.8\,\mathrm{cm}$ radius of the TPC. This pushes these data points apart and clusters anything in the TPC together since they appear to be more alike in this given dataset. These situations are obvious outliers in our data and can be easily removed with a few preliminary cuts. Hence, herein, the training data is all of the

S2O unblinded S2s with pulse area between 3 to 50 electrons, with several cuts applied to remove large backgrounds, like electron trains, as well as periods of instability in the detector, like hotspots and finally a fiducial volume cut to remove wall events and poorly reconstructed S2 positions.

Of the available training permutations, RQ - WF contrast mapping consistently demonstrates the largest separation of clusters, as shown in Figure 7.10. Training is carried out on the RQs with `n_neighbours` = 10, `min_dist` = 0.1 and `metric` = Euclidean and the waveforms with `n_neighbours` = 50, `min_dist` = 0.0 and `metric` = Cosine. These maps are both trained to reduce to 3 dimensions, allowing more complex structures to cluster, but for simplicity, this will only be visualised by projecting into 2D space.

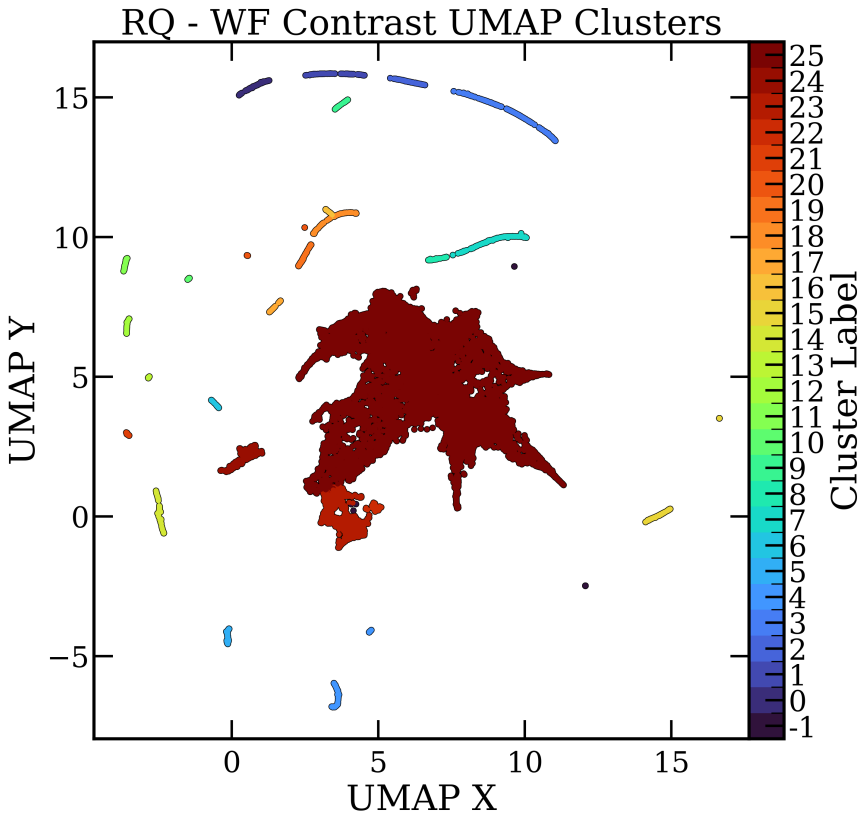


Figure 7.10: RQ-WF contrast embedding for all S2O data after the preliminary cuts are applied. 25 clusters of events are identified.

Figure 7.10 shows the outcome of this training: a large group in the central region which is split into 3 main clusters, as well as a surrounding sea of smaller clusters. To investigate the origins of these clusters, it can be useful to display the clusters coloured by the values of some RQs; a telling example of 4 RQs is shown in Figure 7.11. This figure illustrates the origins of certain clusters of interest. Firstly, the Negative Area plot shows a cluster of S2s with a large portion of the waveform

below the baseline. These are S2s that are misclassified as S2s and should be classified as Other pulses. Secondly, the S2 TBA is indicative of events that occur in the gaseous xenon. Since they are much closer to the top of the detector, they naturally have TBA close to 1. Thirdly, the S2 RMS is a measure of width and can be used to identify events from the bottom of the detector, particularly the cathode. Lastly, a hotspot index is given to any event that is within 5 cm of a known hotspot location. Given the small RMS and yet still likely due to a low-activity period of a hotspot, the events highlighted in the lower right plot are most likely gate events. The origin of the remaining events could be identified through further investigation of the remaining RQs and manual inspection of the events.

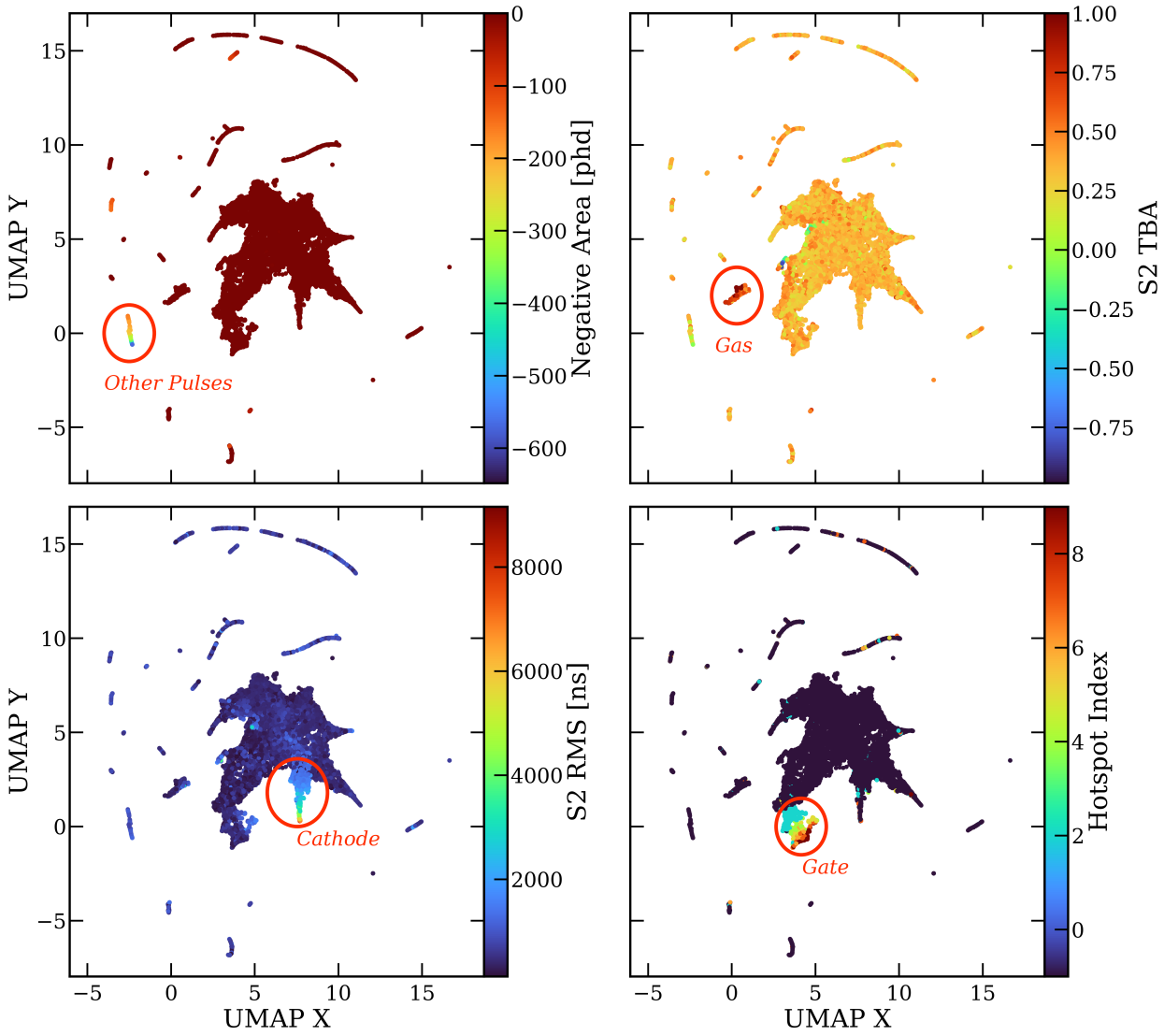


Figure 7.11: UMAP RQ-WF embeddings with colour bars showing Negative Area, TBA, RMS and hotspot index distributions in the UMAP dimensional space. Each RQ is used to identify the likely source of each cluster.

At this stage in the analysis, there already exists a suite of analysis cuts designed to remove the background populations. Hence, once training is complete in this larger parameter space, it is simple to select only the events which remain after all of the cuts have been applied, reduce the signal region to 3.5–$15.5\,\mathrm{electrons}$ and then observe the distribution of events to detect any anomalies that remain in the final dataset. This is shown in Figure 7.12, which demonstrates that the data quality cuts remove all of the clusters of events outside of the central clusters. The majority of the remaining events exist in Cluster 25, seemingly close to the region where cathode backgrounds are recognised in Figure 7.11. Clusters 22 and 23 are both in the region where gate hotspots were identified, and thus it is likely that these events are from the gate backgrounds. It is acceptable in our final dataset to have some cathode events present since this is modelled in our background model and hence will not generate an excess above the background spectrum. The gate events, however, are not currently modelled for the WS2025 dataset and hence must be removed entirely to prevent a false excess. For this reason, it is imperative to investigate the events in Clusters 22 and 23 since these are likely gate events. This can be confirmed by inspection. Many of the events in Cluster 22 are MSE-like events, with 1-3 SE pulses in the same position and much further than $5\,\mu\mathrm{s}$ away. Additionally, the waveforms themselves do not resemble point-like Gaussian sources, and appear more like a continuous stream of SE pulses stretched out to $\sim 6\,\mu\mathrm{s}$ in width. Furthermore, all of the events in Cluster 23 originate from the same (x, y) position as a known remaining hotspot in the detector. The hotspot veto is currently being updated in parallel to this analysis to include this position as another hotspot that will eventually be removed.

On the other side of Figure 7.12, 3 points can be seen in Cluster 25 with UMAP X > 10, one at UMAP Y = 5 and two in the peak at around UMAP Y = 2. The two branches can also be identified from the RQs. The lone position at Y = 5 is in a region with a high SE rate. Manual inspection of this position reveals it contains 3 events, all from the same run and all containing upwards of 9 SE pulses in the event window. However, all of these SE pulses are further than $10\,\mathrm{cm}$ away from the S2 and originate from a known hotspot location, whereas the S2 in the event is not close to the hotspot location. These are valid S2s to consider for analysis. The two events near Y = 2 are in a region of events with some MPE pulses in the event window. LZAP defines an S1 to be a 3-fold coincidence pulse; MPE pulses fall short of this condition and hence are labelled their own pulse type. However, for low-energy nuclear recoil searches, when expanding to 2-fold S1s they will include MPEs as an S1-like signal, hence these two events could be single scatters with

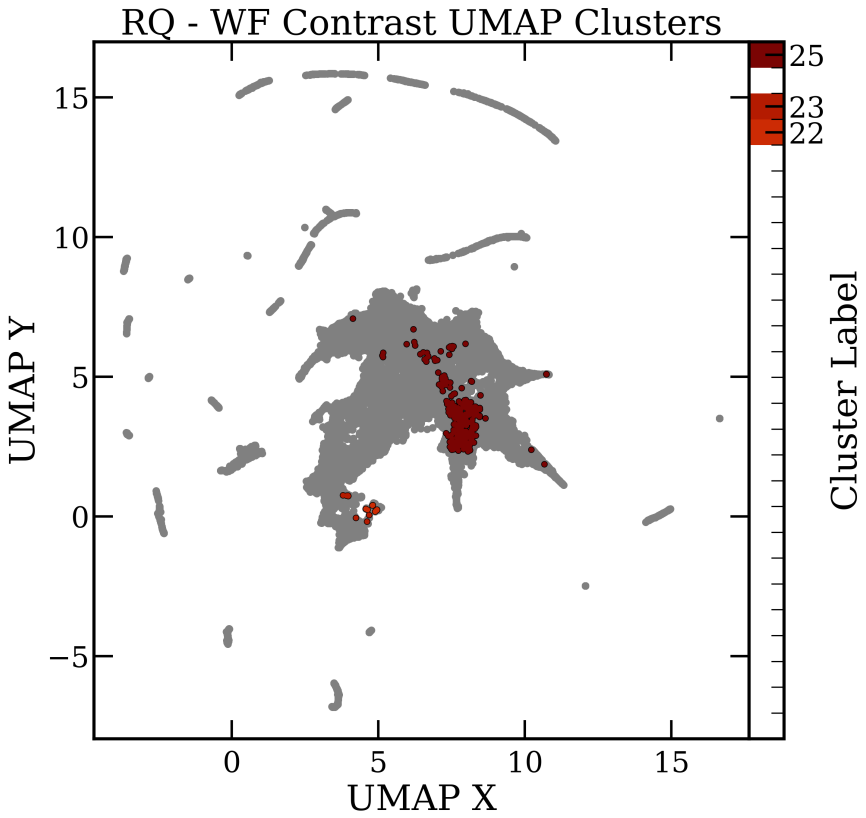


Figure 7.12: RQ - WF mapping from Figure 7.10 highlighting only the events that pass all the current analysis cuts and the clusters to which they belong.

very small S1 pulse areas. Upon inspection, it is observed that while these MPEs do exist, they are not from a genuine SS in the LXe since one of the MPEs occurs 1700 µs after the S2 and the other 2000 µs before the S2; both are unphysical drift times. While these are not considered backgrounds for the S2O searches, they will be treated as accidental backgrounds in the 2-fold low-energy NR searches, which use the same unblinded dataset.

It is expected that this final dataset includes some cathode events, which is acceptable, as this is a modelled background. Nevertheless, any reduction in the number of background events is beneficial for rare-event searches such as this and hence an effort can be made to further separate cathode backgrounds from the remaining data. This can be achieved by selecting the data that passes all of the cuts and re-training UMAP using just this limited dataset, in an identical manner to that described above. However, in this instance, the contrast mapper from WF - RQ provides a cleaner cluster selection, shown in Figure 7.13. This figure shows that there are 3 main clusters, with anomalous waveforms pushed to the very top, right, and lower left of the UMAP latent space.

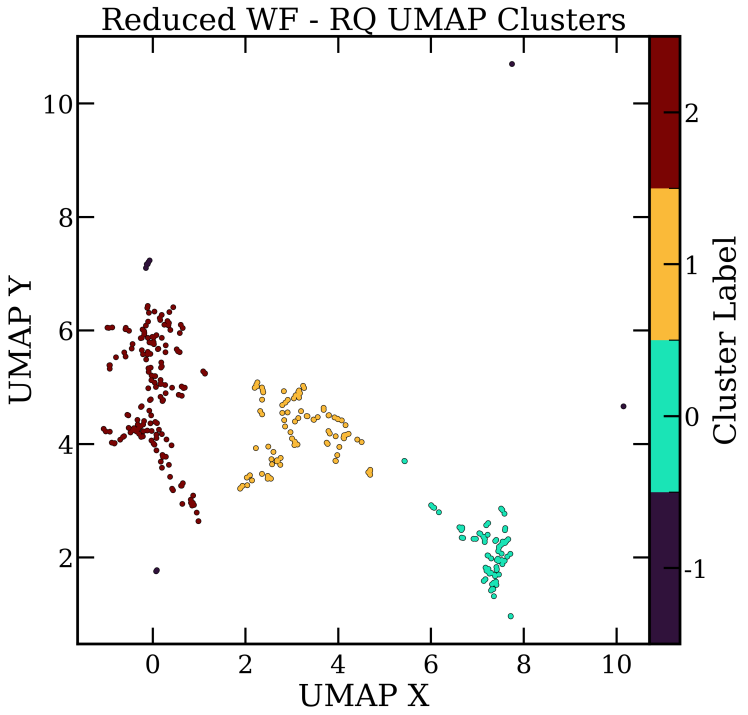


Figure 7.13: WF - RQ contrast mapping for all events passing all analysis cuts. Three anomalous regions are separated from the 3 main clustered regions.

Despite its clear separation from the bulk of the events in this space, this plot reveals one drawback of using the UMAP technique: it provides no information to the user about how the mappings were generated. Thus, the origin of these anomalous events remains unclear even after investigating the RQs and manually inspecting each event and its waveform. However, it is evident that these outliers, each containing 2 events, are self-consistent with the waveforms being similar. The 2 events at the very top of Figure 7.13 are both very wide but have small pulse areas, while the 2 events to the far right of the plot seemingly have a large initial rise with a tail of an SE merged. However, with just two events in each cluster, it is impossible to identify the physical origin of these events; more statistics are needed.

This consolidates the strength of training on a larger selection of data and then narrowing the selection to the events of interest. In this example, we are able to separate the gate events in Clusters 22 and 23. In the future, any events transformed with this trained map and residing in these clusters are likely to be gate events, providing us with stronger rejection power than is achievable from the cuts alone. However, we also note from the examples of the high SE rate and MPE-tagged events in Figure 7.13 that the central cluster contains some substructure that is

compacted together by the larger outliers in the surrounding space. In principle, one could apply any selection of analysis cuts or cluster selection and re-train, performing the iterative process of identifying anomalies in each training space. However, the primary motivation of this technique is to reduce the dimensionality of the data. A simpler secondary approach is devised.

7.4.4 Background Rejection

The anomaly detection approach in the previous section is a powerful tool to separate anomalous populations in the final dataset, as is evident from Figure 7.10. This is a valuable study to perform on the data with all cuts applied to ensure that no anomalies remain. However, in Figure 7.13, this identifies 3 anomalous populations with no information as to why they are anomalous. An alternative approach utilises the pre-existing knowledge of the backgrounds present within the data; grid events, for example. These are mostly removed by enforcing the RMS width of the S2 to reside in $RMS \in [700, 1500]$ ns. A useful approach is to apply all analysis cuts except the RMS width cut, so that most other backgrounds are removed and only the signal S2s and grid events should remain. Training on this dataset using only the waveforms results in Figure 7.14. This can be repeated with any individual analysis cut or combination of cuts to target specific event topologies. Here, we show only the RMS width, as it has the greatest impact on S2O sensitivity.

Figure 7.14 demonstrates that the RMS width cut removes the majority of Clusters 0 and 2. Hence, these are the likely populations of gate and cathode events, while Cluster 1 almost entirely remains. With just 3 clusters, it is simple to inspect each cluster's properties individually. S2s in Cluster 0 are much wider than those in the other two clusters, supporting the claim that they may be cathode backgrounds. A clear method of validating this hypothesis is to compare the average waveform of all clusters with a selection of S1-tagged cathode S2s. This comparison is shown in Figure 7.15 and demonstrates a good level of agreement between cathode S2s and Cluster 0.

Out of the 322 events that are present in Cluster 0, 129 remain after applying the RMS cut and reducing the signal region to 3.5–$15.5\,\mathrm{electrons}$; 18% of the total remaining events. Therefore, the rejection power of cathode backgrounds will improve by removing the events in Cluster 0, but it is not certain that 100% of the cathode events have been removed by this since there is still a possibility that cathode events remain in the other two clusters and the quantity of these is unknowable at this stage.

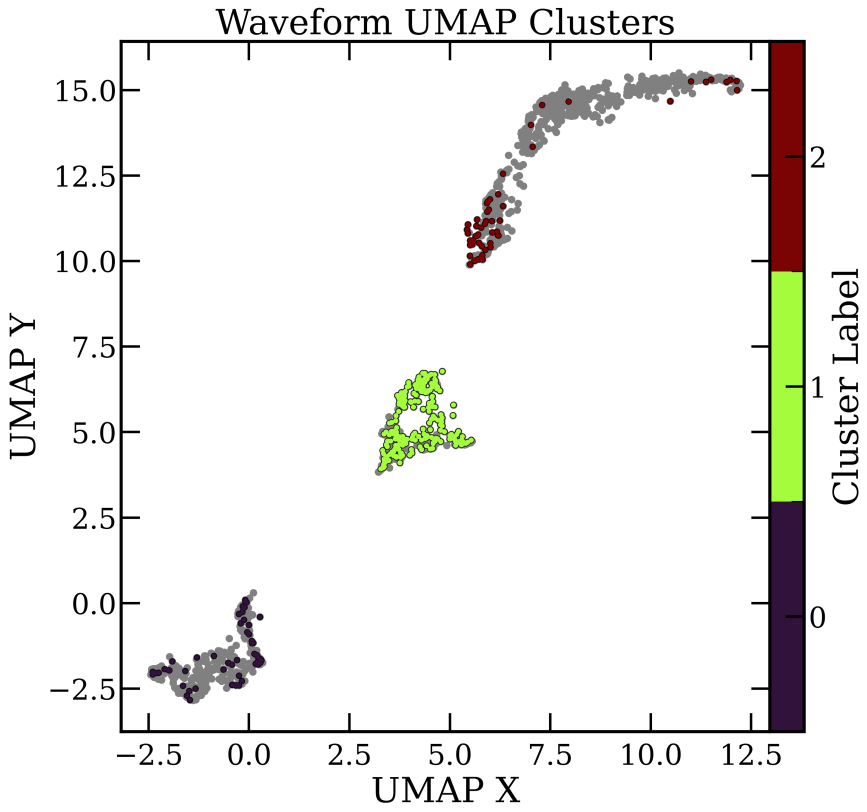


Figure 7.14: UMAP trained on S2 waveforms with all analysis cuts applied except RMS width. The grey data points show the output of all training data, while the highlighted points show the events that pass the RMS cut, coloured by cluster.

It is also clear from Figure 7.15 that the mean Cluster 2 waveform is much shorter than the other clusters, leading to the gate hypothesis. A direct comparison to the gate events is more difficult to achieve since they are not located at a clean drift time like the cathode events. The gate events will inevitably contain leakage from gas events and glue ring echoes[1] which can also be found at drift times less than $5\,\mu s$. A direct comparison of the two waveforms from a simple drift time selection from $2.0\,\mu s$ to $2.8\,\mu s$ drift time is displayed in Figure 7.16. This comparison shows that the Cluster 2 waveforms are slightly wider in time than the glue-ring-contaminated gate events. This could be caused by the contamination of gas and glue ring events in the gate event sample or by some substructure present in Cluster 2.

This substructure is shown in Figure 7.17, with the larger UMAP X positions correlating to wider average pulses. From this, it can be seen that the central region ($6.8 < UMAP\ X < 9.4$) most closely resembles the gate waveforms, while the high UMAP X values are closer to the gas wave-

[1]Glue ring echoes are caused by radioactive decays occurring in the glue ring around the top flange of the LZ detector, which can produce S2-like pulses often following a large Other pulse.

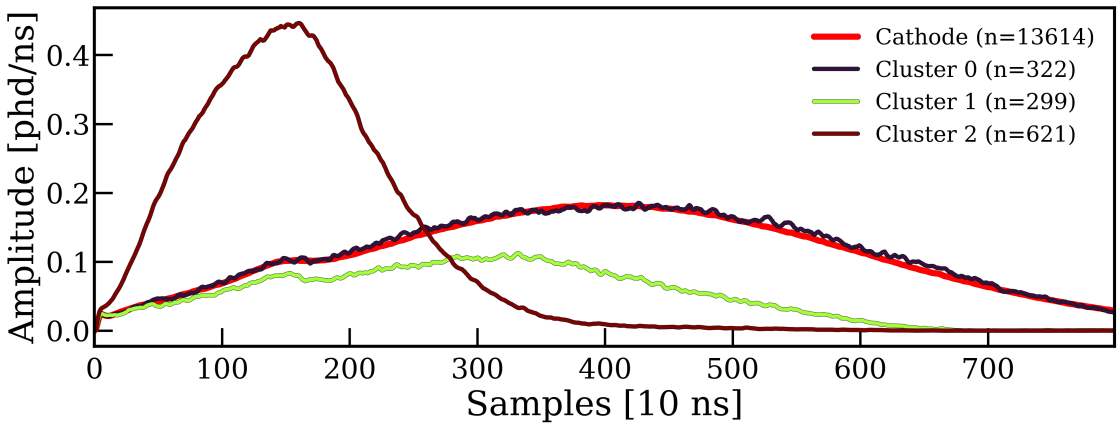


Figure 7.15: Direct comparison of the mean waveform of S1-tagged cathode S2 selection against the mean waveform of each of the clusters in Figure 7.14. Cluster 0 aligns very well with the cathode S2s, whereas the other two clusters disagree.

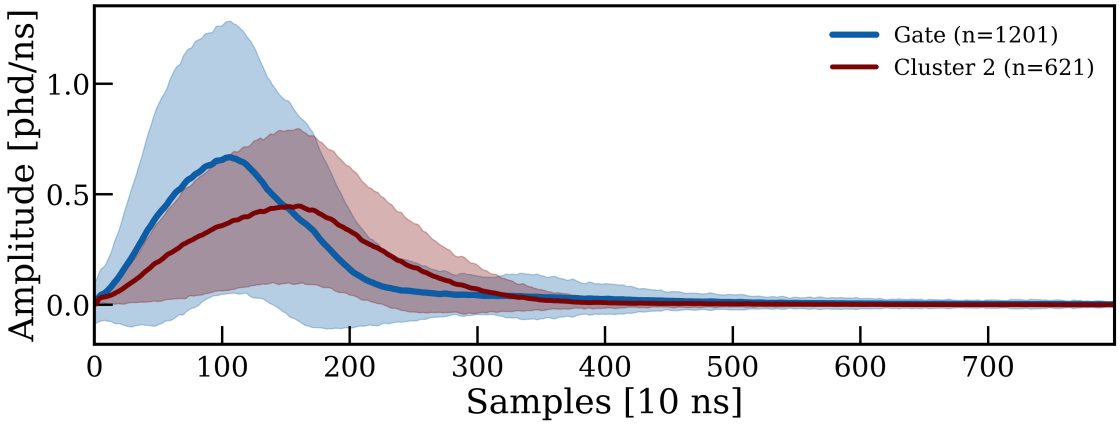


Figure 7.16: Direct comparison of the mean waveform of Cluster 2 and a glue-ring-contaminated selection of gate S2s. The 1σ bands of the mean waveforms are shaded in the corresponding colour.

forms. This gas selection also contains some signal events since it is a simple drift time selection from $2.8\,\mu s$ to $60\,\mu s$. In this instance, this allows us to determine that the origin of Cluster 2 waveforms lies in the top $60\,\mu s$ of the detector, which is removed in S1-S2 analyses by fiducialisation with the drift time. For this reason, in addition to the gate events not being modelled, it is strongly motivated that the events in this cluster should be removed from the final analysis since we must be very strict with accepting un-modelled backgrounds.

A final check one can make to identify the origin of these events is to investigate the relationship with the RQs, with the understanding that this map was trained entirely on waveform information and hence had no access to certain RQs like (x, y) position. To confirm, there are certainly some gate events in Cluster 2; Figure 7.18 shows the (x, y) position distribution and demonstrates that

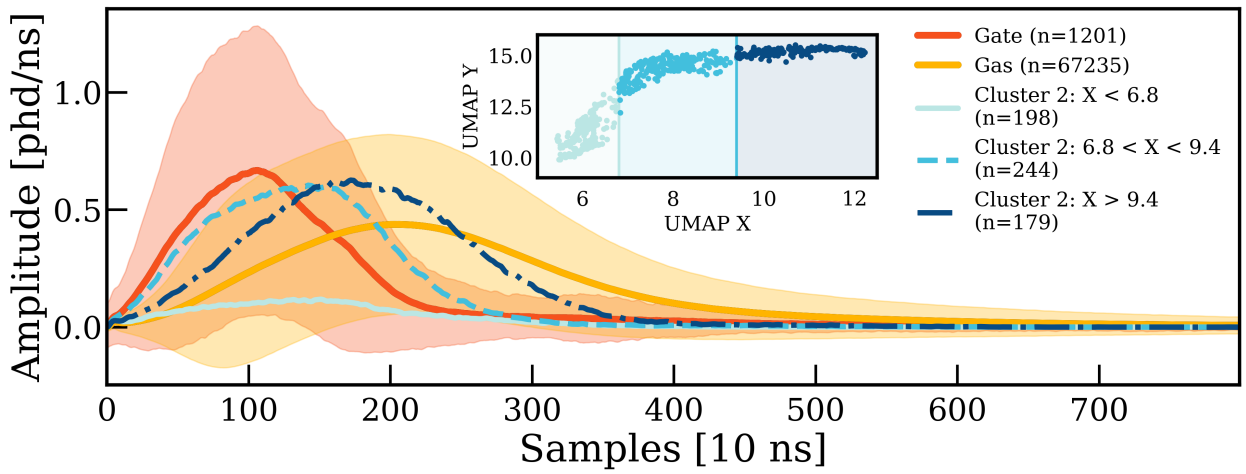


Figure 7.17: Comparison of manual selections of Cluster 2 waveforms with the mean gate and gas S2 waveforms. The Cluster 2 waveforms show substructure within the cluster, suggesting a potential mixture of gate and gas events near the top of the detector.

UMAP has clustered together events that are part of the remaining hotspot near the top of the detector. Given that UMAP had no position information, this serves as a nice confirmation that at least a fraction of the events in Cluster 2 originate from the gate, given their narrow widths and clustered positions. Additionally, comparing waveforms suggests this cluster also contains gas-event contamination, which is un-modelled in the S2O background model and should therefore be strictly removed from the dataset.

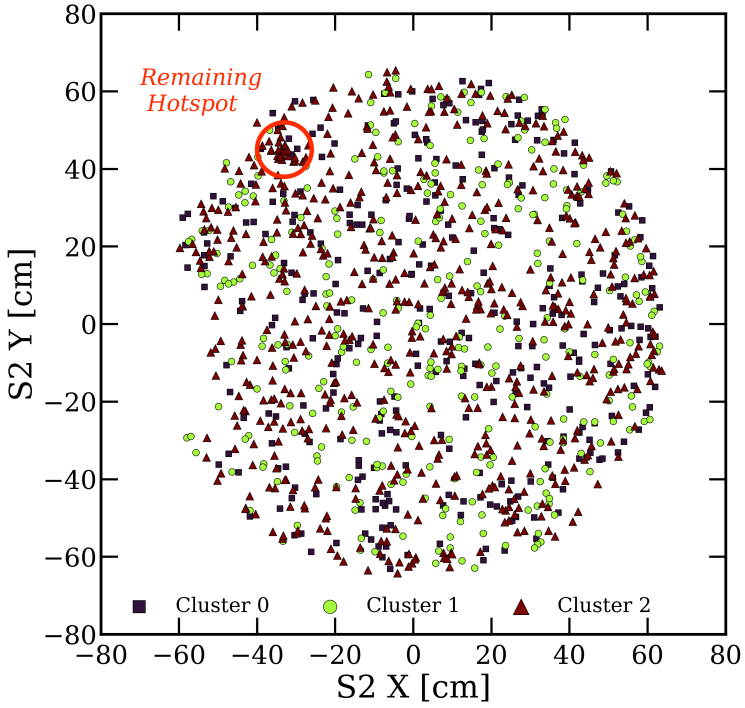


Figure 7.18: Position distribution of each of the clusters from Figure 7.14 with the position of the remaining hotspot highlighted in the red circle.

The final validation step is to investigate Cluster 1 in the RQ and waveform space to ensure this is a signal-like cluster as it is being accepted here. The result of this investigation is that all RQs and mean waveforms look consistent with the signal model. Although this does not confirm that all of the events in this cluster are signal-like, there is no substantial deviation from expectation to significantly remove any further events from this cluster. Therefore, with this approach training with just the waveforms from S2s, we identify two background-enriched clusters containing 449 of the 714 events which pass all data analysis cuts. 63% of the remaining events. Deploying this selection as a cut to remove backgrounds requires validation of a waveform-processed signal model, which is not currently available but will be revisited by analysers to complete the implementation of this work.

In the future, with access to a larger dataset of waveforms, analysers may continue to iteratively select clusters and probe the substructure to identify persistent anomalous populations and to design cuts and selection criteria to improve background rejection. Similarly, here we only discuss the improvements made to the RMS width cut; this technique can equally be applied to any other cut to increase background rejection power. Additionally, when calibration data is processed through

LZAP to store the S2 waveforms, this could be used as a supervised learning tool to distinguish between signal and background populations by training on S2 waveforms from calibration data and S1-tagged S2s. Furthermore, this technique can be applied to any analysis within LZ, since all searches utilise S2 pulses, and it has been shown that background rejection can be significantly improved with this technique.

Overall, as a result of this work, we have explored a variety of ML approaches to background rejection in the S2O regime, specifically targeting grid backgrounds. We conclude that the MSE CNN and the drift time NN are ambitious projects that, if possible, would require access to larger datasets and greater computing power. However, the implementation of UMAP as a dimensionality reduction tool proves to be powerful and versatile, applicable in many ways to aid analyses within LZ. In the near future, we endeavour to wrap this into an LZ-specific tool for analysers to easily utilise within any future analysis. Most importantly, this has improved background rejection in the S2O regime by up to 63% pending signal waveform validation for cathode, gas, and gate events within a single trained map. In future, this will be pushed for use in the 2-fold ^{8}B CEνNS analyses and further at high energies for high-mass WIMPs and $0\nu\beta\beta$ searches.

8 | Conclusion

Throughout this work, we have presented numerous results and methodologies to improve the low-energy S2 threshold in LZ analysis using full chain simulation. This has been achieved through enhanced detector modelling leading to improved waveform simulation, characterising and correcting for new low-energy detector effects, and pushing the limits of LZ analysis to exploit the full waveform information in addition to the standard RQ analysis.

8.1 Waveform Simulations

The improvements to the waveform simulation for the SE and S2 pulses, spanning Chapters 4 and 5, prove a challenging procedure due to the many degeneracies in the large number of models and parameters necessary to reach the levels of accuracy required for S2 waveforms. Nevertheless, through utilising the existing LZ data, the literature in the field and external libraries such as NEST, we demonstrated improvements to the modelling implementation within BACCARAT and DER resulting in the classification accuracy from a NN to be 50.9% and hence barely differentiable.

This work introduced a new criterion for selecting high-quality SE pulses, using TBA to discriminate between genuine SE pulses and Other pulses. We then corrected the electric field and electron drift speed modelling in the gaseous xenon, providing an improvement in the drift time of the electrons travelling in the gas from $226\,\mathrm{ns}$ to $1605\,\mathrm{ns}$. A new photon-production model is introduced to account for the exponential increase in the electric field near the grid wires. This was derived using a dedicated GARFIELD++ implementation. The SPE modelling within the DER software is implemented to correct for remnants of outdated fitting parameters from commissioning campaigns. Finally, a dedicated tuning of material parameters within the GEANT4 geometry was presented, requiring an updated electroluminescence light yield model to account for unphysical reflectivity parameters. Consequently, this led to a strong SE pulse simulation, laying the foundations for the S2 pulses to follow.

For the S2 pulses, this work identified and accounted for an irregularity in the diffusion modelling: the discrepancy between the measured longitudinal diffusion coefficient and the simulated input. The origin of this discrepancy remains an unsolved problem. Nevertheless, this work resulted in a

new, manually tuned value for the input to simulations, achieving unprecedented agreement with calibrated S2 pulse widths in LZ simulations. Additionally, we demonstrate the necessity of migrating from the LZ-tuned extraction efficiency model to the more versatile NEST model, which includes the temperature dependence, and we show that it agrees very well with the LZ measurements. By identifying weaknesses in the implementation of the clustering and yields models, we present a newly created GEANT4 event generator, `ParticleBomb`, specially designed to generate S1, S2, SS or double S2 events which can be used to efficiently generate large statistics to be studied without relying on the clustering or yields models at all. Overall, this work yields the most accurate waveform simulation currently available in LZ.

8.1.1 Alternative Developments

Alongside the developments discussed in this thesis, improvements to the full chain simulation have benefited LZ in many ways. This has enabled many groups and analysers to use waveform information generated by these pulses to advance their research more effectively. A non-exhaustive list of other advancements within LZ that this work has facilitated includes:

* Measuring single scatter detection efficiencies for the WS2024 result, performed jointly with Dan Hunt

* Training a Boosted Decision Tree using the simulated S2 waveforms as signal to discriminate between cathode, gas and gate backgrounds in the S2O search, performed by collaborators using the ParticleBomb S2 waveforms

* Validating new methodologies for high-energy searches to improve SS:MS discrimination, particularly for the $0\nu\beta\beta$ searches, performed jointly with Elisa Jacquet

- Training a NN to perform electron counting for low-energy S2s as opposed to estimations. This ultimately increases our energy resolution at low energies, performed by Tea Hall using the ParticleBomb S2 waveforms

- Testing LZAP Pulse Finder upgrades, performed by Noah Rowe using the full chain simulation

- Investigating new data storage formats for LZ data, performed by Benjamin Almquist using the full chain simulation

For this list, any items noted with $*$ are validated and incorporated into the analysis, while all others are still in progress.

8.1.2 Future Outlook

While the waveform simulation now achieves unprecedented accuracy for the S2 pulses, further work is needed to resolve the remaining inaccuracies. For example, we have noted in this thesis that two separate algorithms are applied to low- and high-energy clustering and yield models. Work is currently underway to unify both regions into a single, well-motivated algorithm that is reliable across all energies. Additionally, this work presented a new tuning of the gaseous xenon material parameters without correcting the existing liquid xenon parameters. A full optical model will need to be tuned to accurately simulate the light collection efficiency for S1 and S2 photons. With these two main updates, the remaining work to upgrade this simulation framework should focus on improving simulation speed by implementing an S1 light map similar to the S2 light map, as well as additional detector-effect models such as electron trains and photoionisation.

8.2 Fragmentation

In Chapter 6, we introduced the concept of fragmented S2 pulses, which were entirely characterised and corrected for using the waveform simulations. This work involved 3 targeted approaches to address this pathology of events, the strongest of which was to reconstruct the fragmented S2s. Reconstructing the pulses in this way negates the necessity of a long LZAP upgrade and data re-processing by using the RQ information to either reconstruct or predict the true values of the RQs for the S2 pulse. Reconstructed pulse areas are obtained exactly, since the fragment areas are additive, whereas widths are predicted using the AFTs, with a mean error of 0.7%, as validated by waveform simulation. Implementing these corrections in this manner improves the observed ^{8}B CEνNS significance from 4.5σ to 4.9σ, approaching the discovery threshold. Prospects for this work will either lie in improving tagging efficiency, which currently stands at 96.4%, improving the accuracy of the width reconstruction algorithm, or in using what was learnt throughout this study to directly upgrade LZAP to account for these effects.

8.3 S2-Only Analysis

The final results presented in this thesis are attempting to utilise machine learning algorithms in different forms to improve the low-energy background discrimination in the S2O search. This is presented in 3 approaches. Firstly, we introduce the concept of MSE events as a potential background topology with a substantial peak lying just below the current analysis threshold of 3.5 electrons. This background must be removed if LZ is to advance to lower S2 thresholds. For now, this is utilised as a portal to high-statistics datasets for grid events. We demonstrate that a simple cut on the number of SE pulses within the event window is insufficient to effectively remove the grid backgrounds, motivating the utilisation of waveforms to differentiate between background and signal with a CNN. We identify the main limitation of using the waveform simulation to do this is that the CNN trains to distinguish between simulation and data as opposed to signal against background. In future, analysers should attempt to retrain this network with a high-statistics sample of calibration S2 waveforms. Secondly, we use simulated S2 waveforms and RQs to train a FiLM ResNet to predict the S2 drift time from the waveform structure. Although this drift time regression provides insufficient accuracy for any level of background rejection of the grid backgrounds, this study introduces the complexities required for complete waveform analysis.

The final proposed approach uses the UMAP package to perform anomaly detection via contrastive clustering with waveforms and RQ-trained maps, and to improve background rejection via waveform-only clustering. Anomaly identification proves useful when trained on large datasets to identify common, larger topologies, but we observe that UMAP is limited when investigating small numbers of events within clusters containing anomalies. Background rejection of gate, gas, and cathode events using this method proves effective, isolating 63% of events that are likely background. The UMAP approach is extremely versatile and applicable to any LZ analysis for anomaly detection or background rejection in any search. This work demonstrates this well for the S2O search, but prospects extend to future LZ searches, since they will all include S2 pulses.

This thesis has demonstrated that improvements to detector simulation, waveform reconstruction and waveform-based analysis can significantly enhance the sensitivity of low-energy searches in liquid xenon TPCs. Together, these developments provide both practical tools for current LZ analyses and a framework for future rare-event searches in next-generation experiments.

Nothing. Lasts. Forever.

— Vessel

References

[1] D. S. Akerib et al. “Searches for Light Dark Matter and Evidence of Coherent Elastic Neutrino-Nucleus Scattering of Solar Neutrinos with the LUX-ZEPLIN (LZ) Experiment”. In: (Dec. 2025). arXiv: 2512.08065 [hep-ex].

[2] Gianfranco Bertone and Dan Hooper. “History of dark matter”. In: *Rev. Mod. Phys.* 90 (4 Oct. 2018), p. 045002. DOI: 10.1103/RevModPhys.90.045002. URL: https://link.aps.org/doi/10.1103/RevModPhys.90.045002.

[3] F. Zwicky. “Republication of: The redshift of extragalactic nebulae”. In: *General Relativity and Gravitation* 41 (2009), pp. 207–224. DOI: 10.1007/s10714-008-0707-4.

[4] Vera C. Rubin and Jr. W. Kent Ford. “Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions”. In: *The Astrophysical Journal* 159 (Feb. 1970), pp. 379–403. DOI: 10.1086/150317.

[5] Douglas Clowe et al. “A Direct Empirical Proof of the Existence of Dark Matter”. In: *The Astrophysical Journal* 648.2 (Aug. 2006), p. L109. DOI: 10.1086/508162. URL: https://doi.org/10.1086/508162.

[6] N. Aghanim et al. “Planck 2018 Results. VI. Cosmological Parameters”. In: *Astronomy & Astrophysics* 641 (2020), A6. DOI: 10.1051/0004-6361/201833910. arXiv: 1807.06209 [astro-ph.CO].

[7] George R. Blumenthal et al. “Formation of Galaxies and Large-Scale Structure with Cold Dark Matter”. In: *Nature* 311 (1984), pp. 517–525. DOI: 10.1038/311517a0.

[8] Ciaran O’Hare. *How To Make A Plot*. GitHub repository, accessed 3 July 2026. 2023. URL: https://github.com/cajohare/HowToMakeAPlot.

[9] Mordehai Milgrom. “A Modification of the Newtonian Dynamics as a Possible Alternative to the Hidden Mass Hypothesis”. In: *The Astrophysical Journal* 270 (1983), pp. 365–370. DOI: 10.1086/161130.

[10] Mordehai Milgrom. “A Modification of the Newtonian Dynamics: Implications for Galaxies”. In: *The Astrophysical Journal* 270 (1983), pp. 371–383. DOI: 10.1086/161131.

[11] P. Tisserand et al. “Limits on the Macho Content of the Galactic Halo from the EROS-2 Survey of the Magellanic Clouds”. In: *Astronomy & Astrophysics* 469.2 (2007), pp. 387–404. DOI: 10.1051/0004-6361:20066017. arXiv: astro-ph/0607207.

[12] C. Alcock et al. “The MACHO Project: Microlensing Results from 5.7 Years of Large Magellanic Cloud Observations”. In: *The Astrophysical Journal* 542.1 (2000), pp. 281–307. DOI: 10.1086/309512. arXiv: astro-ph/0001272.

[13] L. Wyrzykowski et al. “The OGLE view of microlensing towards the Magellanic Clouds – IV. OGLE-III SMC data and final conclusions on MACHOs”. In: *Monthly Notices of the Royal Astronomical Society* 416.4 (2011), pp. 2949–2961. DOI: 10.1111/j.1365-2966.2011.19243.x. arXiv: 1106.2925 [astro-ph.GA].

[14] Edward W. Kolb and Michael S. Turner. *The Early Universe*. Vol. 69. Frontiers in Physics. Redwood City, CA: Addison-Wesley, 1990. ISBN: 9780201116038.

[15] Jonathan L. Feng. “Dark Matter Candidates from Particle Physics and Methods of Detection”. In: *Annual Review of Astronomy and Astrophysics* 48.1 (Aug. 2010), pp. 495–545. ISSN: 1545-4282. DOI: 10.1146/annurev-astro-082708-101659. URL: http://dx.doi.org/10.1146/annurev-astro-082708-101659.

[16] Qing Xia and Lucia Canonica. “Progress and prospects in the underground laboratories’ search for dark matter”. In: *Communications Physics* 9 (Mar. 2026). DOI: 10.1038/s42005-026-02563-1.

[17] M. Ackermann et al. “Searching for Dark Matter Annihilation from Milky Way Dwarf Spheroidal Galaxies with Six Years of Fermi Large Area Telescope Data”. In: *Physical Review Letters* 115 (2015), p. 231301. DOI: 10.1103/PhysRevLett.115.231301. arXiv: 1503.02641 [astro-ph.HE].

[18] H. Abdalla et al. “Search for Dark Matter Annihilation Signals in the H.E.S.S. Inner Galaxy Survey”. In: *Physical Review Letters* 129 (2022), p. 111101. DOI: 10.1103/PhysRevLett.129.111101. arXiv: 2207.10471 [astro-ph.HE].

[19] O. Adriani et al. “An Anomalous Positron Abundance in Cosmic Rays with Energies 1.5–100 GeV”. In: *Nature* 458 (2009), pp. 607–609. DOI: 10.1038/nature07942. arXiv: 0810.4995 [astro-ph].

[20] M. G. Aartsen et al. "Search for Neutrinos from Dark Matter Self-Annihilations in the Center of the Milky Way with 3 Years of IceCube/DeepCore". In: *European Physical Journal C* 77 (2017), p. 627. DOI: 10.1140/epjc/s10052-017-5213-y. arXiv: 1705.08103 [hep-ex].

[21] K. Abe et al. "Indirect Search for Dark Matter from the Galactic Center and Halo with the Super-Kamiokande Detector". In: *Physical Review D* 102 (2020), p. 072002. DOI: 10.1103/PhysRevD.102.072002. arXiv: 2005.05109 [hep-ex].

[22] G. Aad et al. "Combination of searches for invisible decays of the Higgs boson using 139 fb^{-1} of proton-proton collision data at $\sqrt{s} = 13$ TeV collected with the ATLAS experiment". In: *Physics Letters B* 842 (July 2023), p. 137963. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2023.137963. URL: http://dx.doi.org/10.1016/j.physletb.2023.137963.

[23] G. Aad et al. "Combination and summary of ATLAS dark matter searches interpreted in a 2HDM with a pseudo-scalar mediator using 139 fb^{-1} of TeV pp collision data". In: *Science Bulletin* 69.19 (Oct. 2024), pp. 3005–3035. ISSN: 2095-9273. DOI: 10.1016/j.scib.2024.06.003. URL: http://dx.doi.org/10.1016/j.scib.2024.06.003.

[24] G. Aad et al. "Search for dark matter particles in W+W- events with transverse momentum imbalance in proton-proton collisions at $\sqrt{s}$ = 13 TeV". In: *Journal of High Energy Physics* 2024.3 (Mar. 2024). ISSN: 1029-8479. DOI: 10.1007/jhep03(2024)134. URL: http://dx.doi.org/10.1007/JHEP03(2024)134.

[25] J.D. Lewin and P.F. Smith. "Review of mathematics, numerical factors, and corrections for dark matter experiments based on elastic nuclear recoil". In: *Astroparticle Physics* 6.1 (1996), pp. 87–112. ISSN: 0927-6505. DOI: https://doi.org/10.1016/S0927-6505(96)00047-3. URL: https://www.sciencedirect.com/science/article/pii/S0927650596000473.

[26] R. Bernabei et al. "First model independent results from DAMA/LIBRA-phase2". In: *Nucl. Phys. At. Energy* 19 (2018), pp. 307–325. DOI: 10.15407/jnpae2018.04.307. arXiv: 1805.10486 [hep-ex].

[27] J. Amaré et al. “First results on dark matter annual modulation from ANAIS-112 experiment”. In: *Physical Review Letters* 123 (2019), p. 031301. DOI: 10.1103/PhysRevLett.123.031301. arXiv: 1903.03973 [astro-ph.IM].

[28] COSINE-100 Collaboration. “Three-year annual modulation search with COSINE-100”. In: *Physical Review D* 106 (2022), p. 052005. DOI: 10.1103/PhysRevD.106.052005. arXiv: 2111.08863 [hep-ex].

[29] Giulia D’Imperio et al. *Dark matter search with the SABRE experiment*. Status report and concept paper. 2018. DOI: 10.48550/arXiv.1807.00584. arXiv: 1807.00584 [physics.ins-det].

[30] F. D. Amaro et al. *The CYGNO experiment: a gaseous TPC with optical readout for rare events searches*. 2026. arXiv: 2603.22595 [physics.ins-det]. URL: https://arxiv.org/abs/2603.22595.

[31] D. S. Akerib et al. “The LUX-ZEPLIN (LZ) Experiment”. In: *Nuclear Instruments and Methods in Physics Research Section A* 953 (2020), p. 163047. DOI: 10.1016/j.nima.2019.163047. arXiv: 1910.09124 [physics.ins-det].

[32] XENON Collaboration. “Dark Matter Search Results from a One Ton-Year Exposure of XENON1T”. In: *Physical Review Letters* 121 (2018), p. 111302. DOI: 10.1103/PhysRevLett.121.111302. arXiv: 1805.12562 [astro-ph.CO].

[33] X. Cui et al. “Dark Matter Results From 54-Ton-Day Exposure of PandaX-II Experiment”. In: *Phys. Rev. Lett.* 119 (2017), p. 181302.

[34] DarkSide Collaboration. “Dark Matter Search Results from a One-Ton-Year Exposure of DarkSide-50”. In: *Physical Review Letters* 121 (2018), p. 081307. DOI: 10.1103/PhysRevLett.121.081307. arXiv: 1802.06994 [hep-ex].

[35] Ciaran O’Hare. *Direct Detection Plots*. GitHub Repository. 2026. URL: https://github.com/cajohare/DirectDetectionPlots.

[36] J. Aalbers et al. “Dark Matter Search Results from 4.2 Tonne-Years of Exposure of the LUX-ZEPLIN (LZ) Experiment”. In: *Physical Review Letters* 135.1 (July 2025). ISSN: 1079-7114. DOI: 10.1103/4dyc-z8zf. URL: http://dx.doi.org/10.1103/4dyc-z8zf.

[37] D. S. Akerib et al. "Projected sensitivity of the LUX-ZEPLIN experiment to the $0\nu\beta\beta$ decay of ^{136}Xe". In: *Physical Review C* 102.1 (July 2020). ISSN: 2469-9993. DOI: 10.1103/physrevc.102.014602. URL: http://dx.doi.org/10.1103/PhysRevC.102.014602.

[38] S. Abe et al. "Search for the Majorana Nature of Neutrinos in the Inverted Mass Ordering Region with KamLAND-Zen". In: *Phys. Rev. Lett.* 130 (2023), p. 051801. DOI: 10.1103/PhysRevLett.130.051801. arXiv: 2203.02139 [hep-ex].

[39] J. B. Albert et al. "Sensitivity and Discovery Potential of nEXO to Neutrinoless Double Beta Decay". In: *Phys. Rev. C* 97 (2018), p. 065503. DOI: 10.1103/PhysRevC.97.065503. arXiv: 1710.05075 [nucl-ex].

[40] B. J. Mount et al. *LUX-ZEPLIN (LZ) Technical Design Report*. 2017. arXiv: 1703.09144 [physics.ins-det]. URL: https://arxiv.org/abs/1703.09144.

[41] Jeremy Mock et al. "Modeling Pulse Characteristics in Xenon with NEST". In: *JINST* 9 (2014), T04002. DOI: 10.1088/1748-0221/9/04/T04002. arXiv: 1310.1117 [physics.ins-det].

[42] J. Aalbers et al. "The design, implementation, and performance of the LZ calibration systems". In: *Journal of Instrumentation* 19.08 (Aug. 2024), P08027. ISSN: 1748-0221. DOI: 10.1088/1748-0221/19/08/p08027. URL: http://dx.doi.org/10.1088/1748-0221/19/08/P08027.

[43] Peter Sorensen and Carl Eric Dahl. "Nuclear recoil energy scale in liquid xenon with application to the direct detection of dark matter". In: *Phys. Rev. D* 83 (6 Mar. 2011), p. 063501. DOI: 10.1103/PhysRevD.83.063501. URL: https://link.aps.org/doi/10.1103/PhysRevD.83.063501.

[44] Laura Baudis, Patricia Sanchez-Lucas, and Kevin Thieme. "A measurement of the mean electronic excitation energy of liquid xenon". In: *Eur. Phys. J. C* 81 (2021), p. 1060. DOI: 10.1140/epjc/s10052-021-09834-x. arXiv: 2109.07151 [physics.ins-det].

[45] Matthew Szydagis et al. "A Review of Basic Energy Reconstruction Techniques in Liquid Xenon and Argon Detectors for Dark Matter and Neutrino Physics Using NEST". In:

Instruments 5.1 (2021). ISSN: 2410-390X. DOI: 10.3390/instruments5010013. URL: https://www.mdpi.com/2410-390X/5/1/13.

[46] J. Aalbers et al. "First Dark Matter Search Results from the LUX-ZEPLIN (LZ) Experiment". In: *Phys. Rev. Lett.* 131 (2023), p. 041002. eprint: 2207.03764.

[47] G. Pereira, C. Silva, and V. N. Solovov. "Energy resolution of the LZ detector for high-energy electronic recoils". In: *JINST* 18.04 (2023), p. C04007. DOI: 10.1088/1748-0221/18/04/C04007.

[48] J. Aalbers et al. "Background determination for the LUX-ZEPLIN dark matter experiment". In: *Physical Review D* 108.1 (July 2023). ISSN: 2470-0029. DOI: 10.1103/physrevd.108.012010. URL: http://dx.doi.org/10.1103/PhysRevD.108.012010.

[49] J. Aalbers et al. "The XLZD Design Book: towards the next-generation liquid xenon observatory for dark matter and neutrino physics". In: *The European Physical Journal C* 85.10 (Oct. 2025). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-025-14810-w. URL: http://dx.doi.org/10.1140/epjc/s10052-025-14810-w.

[50] D. S. Akerib et al. "Simulations of Events for the LUX-ZEPLIN (LZ) Dark Matter Experiment". In: *Astropart. Phys.* 125 (2021), p. 102480. DOI: 10.1016/j.astropartphys.2020.102480. arXiv: 2001.09363 [physics.ins-det].

[51] D. S. Akerib et al. "The Large Underground Xenon (LUX) Experiment". In: *Nucl. Instrum. Meth. A* 704 (2013), pp. 111–126. DOI: 10.1016/j.nima.2012.11.135. arXiv: 1211.3788 [physics.ins-det].

[52] David Woodward. "Latest results of the LUX dark matter experiment". In: *Int. J. Mod. Phys. Conf. Ser.* 50 (2020). Ed. by Anna Erickson and Marianne Hamm, p. 2060002. DOI: 10.1142/S2010194520600022.

[53] D. S. Akerib et al. "Investigation of background electron emission in the LUX detector". In: *Phys. Rev. D* 102.9 (2020), p. 092004. DOI: 10.1103/PhysRevD.102.092004. arXiv: 2004.07791 [physics.ins-det].

[54] D. S. Akerib et al. “Nuclear Recoil Calibration at Sub-keV Energies in LUX and Its Impact on Dark Matter Search Sensitivity”. In: *Phys. Rev. Lett.* 134.6 (2025), p. 061002. DOI: 10.1103/PhysRevLett.134.061002. arXiv: 2210.05859 [physics.ins-det].

[55] Michael Hahsler and Matthew Piekenbrock. *dbscan: Density-Based Spatial Clustering of Applications with Noise (DBSCAN) and Related Algorithms*. R package version 1.2.4. 2025. DOI: 10.32614/CRAN.package.dbscan. URL: https://CRAN.R-project.org/package=dbscan.

[56] Michael Hahsler, Matthew Piekenbrock, and Derek Doran. “dbscan: Fast Density-Based Clustering with R”. In: *Journal of Statistical Software* 91.1 (2019), pp. 1–30. DOI: 10.18637/jss.v091.i01.

[57] Anna David. “Background Mitigation in the LUX-ZEPLIN Experiment Using Statistical Analysis of Waveforms”. PhD thesis. University College London, 2025.

[58] J. Aalbers et al. *The Data Acquisition System of the LZ Dark Matter Detector: FADR*. 2024. arXiv: 2405.14732 [physics.ins-det]. URL: https://arxiv.org/abs/2405.14732.

[59] B. Lopez Paredes et al. “Response of photomultiplier tubes to xenon scintillation light”. In: *Astroparticle Physics* 102 (Nov. 2018), pp. 56–66. ISSN: 0927-6505. DOI: 10.1016/j.astropartphys.2018.04.006. URL: http://dx.doi.org/10.1016/j.astropartphys.2018.04.006.

[60] D.S. Akerib et al. “Data acquisition and readout system for the LUX dark matter experiment”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 668 (Mar. 2012), pp. 1–8. ISSN: 0168-9002. DOI: 10.1016/j.nima.2011.11.063. URL: http://dx.doi.org/10.1016/j.nima.2011.11.063.

[61] P. Brás et al. “A machine learning-based methodology for pulse classification in dual-phase xenon time projection chambers”. In: *The European Physical Journal C* 82.6 (June 2022). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-022-10502-x. URL: http://dx.doi.org/10.1140/epjc/s10052-022-10502-x.

[62] M. Szydagis et al. “A review of NEST models for liquid xenon and an exhaustive comparison with other approaches”. In: *Frontiers in Detector Science and Technology* 2 (Jan. 2025). ISSN: 2813-8031. DOI: 10.3389/fdest.2024.1480975. URL: http://dx.doi.org/10.3389/fdest.2024.1480975.

[63] S. Dey. “Electric Fields and their Implications in the LUX-ZEPLIN Time Projection Chamber”. PhD thesis. University of Oxford, 2025.

[64] Matthew W. Scroggs et al. “Construction of arbitrary order finite element degree-of-freedom maps on polygonal and polyhedral cell meshes”. In: *ACM Transactions on Mathematical Software* 48.2 (2022), 18:1–18:23. DOI: 10.1145/3524456.

[65] C M B Monteiro et al. “Secondary scintillation yield in pure xenon”. In: *Journal of Instrumentation* 2.05 (May 2007), P05001. ISSN: 1748-0221. DOI: 10.1088/1748-0221/2/05/p05001. URL: http://dx.doi.org/10.1088/1748-0221/2/05/P05001.

[66] D Yu Akimov et al. “A setup for studying electroluminescence in gaseous xenon”. en. In: *Instrum. Exp. Tech.* 65.5 (Oct. 2022), pp. 849–857.

[67] Ryan Edward Linehan. “High voltage electrode development and the LZ experiment’s WIMP search”. PhD thesis. Stanford U., 2022.

[68] R. Linehan et al. “Design and production of the high voltage electrode grids and electron extraction region for the LZ dual-phase xenon time projection chamber”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 1031 (May 2022), p. 165955. ISSN: 0168-9002. DOI: 10.1016/j.nima.2021.165955. URL: http://dx.doi.org/10.1016/j.nima.2021.165955.

[69] S. F. Biagi. “Monte Carlo simulation of electron drift and diffusion in counting gases under the influence of electric and magnetic fields”. In: *Nucl. Instrum. Meth. A* 421 (1999), pp. 234–240. DOI: 10.1016/S0168-9002(98)01233-9.

[70] R. Veenhof et al. “Garfield++: Simulation of tracking detectors”. In: *CERN* (2010). https://garfieldpp.web.cern.ch/.

[71] D. S. Akerib et al. "Low-energy calibration of the LUX dark matter experiment using Kr-83m". In: *Phys. Rev. D* 96 (2017), p. 112011. eprint: 1705.03380.

[72] A. Simón et al. "Electron drift properties in high pressure gaseous xenon". In: *JINST* 13 (2018), P07013. eprint: 1804.01680.

[73] K. Abe et al. "A measurement of the scintillation decay time constant of nuclear recoils in liquid xenon with the XMASS-I detector". In: *Journal of Instrumentation* 13.12 (Dec. 2018), P12032–P12032. ISSN: 1748-0221. DOI: 10.1088/1748-0221/13/12/p12032. URL: http://dx.doi.org/10.1088/1748-0221/13/12/P12032.

[74] K. C. Oliver-Mallory et al. *Spectroscopy of VUV luminescence in dual-phase xenon detectors*. 2025. arXiv: 2512.20935 [physics.ins-det]. URL: https://arxiv.org/abs/2512.20935.

[75] Brais López Paredes. *Optical Simulations of the LZ Detector*. Presentation at the 8th International Conference on Low Energy Particle Physics (LIDINE 2019), Rochester, NY, USA. Conference presentation. Sept. 2019. URL: https://lz.lbl.gov/wp-content/uploads/sites/6/2019/09/BraisLopezParedes_LIDINE2019.pdf.

[76] V Chepel and H Araújo. "Liquid noble gas detectors for low energy particle physics". In: *Journal of Instrumentation* 8.04 (Apr. 2013), R04001–R04001. ISSN: 1748-0221. DOI: 10.1088/1748-0221/8/04/r04001. URL: http://dx.doi.org/10.1088/1748-0221/8/04/R04001.

[77] D. Yu. Akimov. "Scintillation in xenon at pressures up to several atmospheres". In: *Soviet Physics JETP* 39.1 (1974). English translation of Zh. Eksp. Teor. Fiz. 66 (1), 195 (1974).

[78] A. Parsons et al. "High pressure gas scintillation drift chambers with wave shifter fiber readout". In: *IEEE Transactions on Nuclear Science* 36.1 (Feb. 1989). URL: https://www.osti.gov/biblio/6285947.

[79] A. C. Fonseca et al. "Study of secondary scintillation in xenon vapour". In: *IEEE Nuclear Science Symposium Conference Record*. Vol. 1. IEEE. 2004, pp. 572–576.

[80] C. Silva et al. "Reflectance of Polytetrafluoroethylene (PTFE) for Xenon Scintillation Light". In: *IEEE Transactions on Nuclear Science* (2009). arXiv: 0910.1056 [physics.ins-det].

[81] W. Lorenzon et al. “PTFE Reflectance Measurements for Xenon Scintillation Light”. In: *LIDINE 2017 Conference*. 2017.

[82] D. Bajpai et al. “Validation of the VUV-reflective coating for next-generation liquid xenon detectors”. In: *arXiv* (2024). arXiv: 2402.15177 [physics.ins-det].

[83] J. Crank. *The Mathematics of Diffusion*. 2nd ed. Oxford: Oxford University Press, 1975. ISBN: 9780198534112.

[84] O. Njoya et al. “Measurements of electron transport in liquid and gas Xenon using a laser-driven photocathode”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 972 (2020), p. 163965. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2020.163965. URL: https://www.sciencedirect.com/science/article/pii/S0168900220304320.

[85] W. -T. Chen et al. *Measurement of the transverse diffusion coefficient of charge in liquid xenon*. 2011. arXiv: 1109.3300 [physics.ins-det]. URL: https://arxiv.org/abs/1109.3300.

[86] J. B. Albert et al. “Measurement of the drift velocity and transverse diffusion of electrons in liquid xenon with the EXO-200 detector”. In: *Physical Review C* 95.2 (2017), p. 025502. DOI: 10.1103/PhysRevC.95.025502. arXiv: 1609.04467 [physics.ins-det].

[87] Hamamatsu Photonics K.K. *Photomultiplier Tube R11410-20A Datasheet*. Available from Hamamatsu Photonics. URL: https://www.hamamatsu.com/us/en/product/optical-sensors/pmt/pmt_tube-alone/head-on-type/R11410-20A.html.

[88] Dan Hunt. “Optimising Low-Mass WIMP Searches in the LUX-ZEPLIN Experiment”. PhD thesis. University of Oxford, 2024.

[89] D. S. Akerib et al. “3D Modeling of Electric Fields in the LUX Detector”. In: *Physical Review D* 97 (2018), p. 102008. DOI: 10.1103/PhysRevD.97.102008. arXiv: 1709.00095 [physics.ins-det].

[90] Peter Sorensen. *Electron train backgrounds in liquid xenon dark matter search detectors are indeed due to thermalization and trapping*. 2017. arXiv: 1702.04805 [physics.ins-det]. URL: https://arxiv.org/abs/1702.04805.

[91] M. Gushchin E. et al. "Electron emission from condensed noble gases". In: *Soviet Physics JETP* 49.5 (1979). English transl. of Zh. Eksp. Teor. Fiz. 76(5), 1685 (1979), pp. 856–862.

[92] V. N. Solovov et al. "Position Reconstruction in a Dual Phase Xenon Scintillation Detector". In: *IEEE Transactions on Nuclear Science* 59.6 (2012), pp. 3286–3293. DOI: 10.1109/TNS.2012.2218668.

[93] Austin Vaitkus. "Characterization of the R11410-22 Photomultiplier Tube and the Search for a Neutron Induced Migdal Effect in the LUX-ZEPLIN Experiment". PhD thesis. Brown University, 2025.

[94] National Energy Research Scientific Computing Center. *Perlmutter System Documentation*. 2025. DOI: 10.25344/S4GC7R. URL: https://docs.nersc.gov/systems/perlmutter/.

[95] E. Aprile et al. "Light Dark Matter Search with 7.8 Tonne-Year of Ionization-Only Data in XENONnT". In: *arXiv* (2026). arXiv: 2601.11296 [hep-ex].

[96] Meng Wang et al. "Search for Light Dark Matter with Ionization Signals in the PandaX-4T Experiment". In: *Physical Review Letters* 130.26 (2023), p. 261001. DOI: 10.1103/PhysRevLett.130.261001. arXiv: 2212.10067 [hep-ex].

[97] D. S. Akerib et al. "Improving sensitivity to low-mass dark matter in LUX using a novel electrode background mitigation technique". In: *Phys. Rev. D* 104 (1 July 2021), p. 012011. DOI: 10.1103/PhysRevD.104.012011. URL: https://link.aps.org/doi/10.1103/PhysRevD.104.012011.

[98] D. S. Akerib et al. *Study of few-electron backgrounds in the LUX-ZEPLIN detector*. 2026. arXiv: 2510.06500 [physics.ins-det]. URL: https://arxiv.org/abs/2510.06500.

[99] Glen Cowan et al. "Asymptotic Formulae for Likelihood-Based Tests of New Physics". In: *European Physical Journal C* 71 (2011), p. 1554. DOI: 10.1140/epjc/s10052-011-1554-0. arXiv: 1007.1727 [physics.data-an].

[100] Leland McInnes, John Healy, and contributors. *umap-learn*. Version 0.5.9. 2025. DOI: 10.5281/zenodo.15794118. URL: https://doi.org/10.5281/zenodo.15794118.

[101] John Winnicki et al. "Unsupervised Learning Techniques for Identification of Anomalous LZ Waveform Data". In: *EPJ Web Conf.* 337 (2025), p. 01122. DOI: 10.1051/epjconf/202533701122.

[102] Leland McInnes, John Healy, and James Melville. *UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction*. 2020. arXiv: 1802.03426 [stat.ML]. URL: https://arxiv.org/abs/1802.03426.

[103] Ricardo J. G. B. Campello, Davoud Moulavi, and Joerg Sander. "Density-Based Clustering Based on Hierarchical Density Estimates". In: *Advances in Knowledge Discovery and Data Mining*. Ed. by Jian Pei et al. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 160–172. ISBN: 978-3-642-37456-2.