Optimising Low-Mass WIMP Searches in the LUX-ZEPLIN Experiment

Dan Hunt

St. Peter's College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy
Trinity 2024

Dedicated to my parents, without whom I would not be the man I am today.
Dedicated to my granddad, who helped me to ask the right questions now that I'm here.

Acknowledgements

My time in Oxford has taught me many lessons, about both the real world and how to detect particles: all of these lessons have been essential parts of my D.Phil. While this work has helped us to see further into the world of dark matter, it has done so on the shoulders of giants and peers alike. This page is dedicated to the people without whom I would not have been able to produce this work.

I would like to firstly thank my supervisor Hans Kraus for his guidance and advice throughout this DPhil. I hope I have done you proud through this work, and I wish you the best during your (very well-deserved!) retirement. I would also like to thank my peers in the Oxford LZ group: Kim, Daria, Sparshita, Josh, Nick, and Anthony. It's been a wonderful four years, and I'm looking forward to many more in the future.

For their input to the work presented in this thesis, I would like to thank: Paulo, for his insight into the mysteries of LZap; Ibles, for his insight into accidentals and for generally being lovely; and both Teal and Ann, for their joint leadership of the S2-Only analysis thus far.

For more general advice in academia, I would like to extend special thanks to both Amy and Aiham, without whom I would have floundered long ago.

I should also thank Scott Kravitz for his insights into machine learning and its applications in LZ. I look forward to working with you more in the coming years.

There are many more I should thank for their contributions to LZ (and also my life) - Kevin, Pawel, Theresa, Kelsey, Nicolas, Eli, Ewan, Alberto, Elisa, Albert, Andrew, Ryan, Kelly, Ishan - but to name everyone might stretch the patience of anyone reading this, as well as my printer ink budget when I try to bind this thesis.

Not to mention that I still have my family to thank: particularly my mum and dad, my grandparents, and my auntie in particular, for keeping me sane during the COVID-19 pandemic. And also my friends: those from home - Laurence, Elliot, Leonie, Adam, Owen, Jack, James, Sam - those in Oxford - Cameron, Juliet, Rob, Sam, Ed, Ali, Maxence, John, Jamie, Sally - and those from my time in Birmingham - Torin, Sam, Daria, Taylor, James, Evan, Izzy, Shark, and Lex. Special mention goes out to Joe, who left us under tragic circumstances, but whose spirit I hope I am continuing. I love you, buddy.

And last of all, I am indebted to my partner Foley, for her unwavering love and encouragement through my time in Oxford.

Statement of Originality

I, Dan Hunt, declare that this thesis has been composed solely by myself and that it has not been submitted, in whole or in part, in any previous application for a degree.

It is acknowledged that the work described here relies on or is otherwise connected to work performed by others in the LUX-ZEPLIN collaboration. Such works are acknowledged below. Except where stated otherwise, the work presented is entirely my own.

- Chapter 3 details the design and operation of LZ, as well as the first WIMP search results. This is work carried out by the wider LZ collaboration.
- Chapter 4 discusses methods of modelling accidental coincidence events. These datasets were acquired and generated by the wider LZ collaboration.
- Chapter 5 uses data from simulations of LZ calibration runs. The framework for simulation was tuned and developed by Nicholas Fieldhouse at the University of Oxford. The author worked on validation of these simulations, although this work is not presented in the thesis.
- Chapter 6 discusses the selection criterion for S2-Only events. These criterion were developed by the wider LZ collaboration.

Abstract

The nature of dark matter is an ongoing mystery in the fields of both astrophysics and particle physics. Evidence over a wide range of length scales indicates that dark matter accounts for 84% of the matter density of the universe. However, detection has remained elusive.

The LUX-ZEPLIN (LZ) experiment is a dual-phase LXe time projection chamber designed for the direct detection of dark matter. In its first science run, LZ achieved sensitivity to spin-independent WIMP-nucleon cross sections of $9.2 \times 10^{-48}\,\mathrm{cm}^2$ for a WIMP of mass $36\,\mathrm{GeV/c}^2$ at the 90% confidence level. While LZ is designed for the detection of high-mass WIMPs, there is also ample motivation to maximise LZ's sensitivity to dark matter of lower masses.

The contributions of this thesis to dark matter searches improve the sensitivity of LZ to low-mass dark matter candidates: both in standard analysis, where sensitivity is limited by low-energy backgrounds and signal efficiency, as well as through the use of electroluminescence-only (S2-Only) data, where this signal efficiency is maximised.

This thesis introduces a novel method of rejecting accidental coincidence events, the lowest-energy background present in LZ data. It is possible in LZ to reject 95% of these events while preserving >99% of signal, allowing for sensitivity to both low and high-mass WIMPs.

The first measurement of S2 pulse mis-classification in LZ is also presented. This work has been used to achieve maximum signal detection efficiency at low energies. The probability of mis-classification of any electroluminescence pulse, down to detection threshold of 2 e, is below 5%. These measurements are then incorporated into statistical analysis, in order to minimise uncertainty on measurements of signal sensitivity at lower masses.

The thesis concludes with a demonstration of how the capability of LZ to reject low-energy backgrounds and accept low-energy signals can be leveraged in an analysis of electroluminescence-only (S2-Only) data in LZ, a dataset sensitive to dark matter of masses as low as 80 MeV. A Boosted Decision Tree (BDT) trained to reject grid radiogenic events in S2-Only data is presented. It is shown to be possible to reject 80% of grid background events with a signal efficiency of 80%, a significant improvement on similar results from LUX which improves the sensitivity of LZ to low-mass WIMPs by up to a factor of ten.

Contents

Appendices

List of Figures

List of Tables

Glossary

Area	The total integrated waveform across all PMTs for a given pulse
BACCARAT	Basically, A Component-Centric. Analog Response to AnyThing, an LZ simulation package built on GEANT4
CDF	Cumulative Distribution Function, the probability of an event at less than or equal to a given input
CMB	Cosmic Microwave Background, isotropic and homogenous microwave radiation. Provides information about the early universe
CSD	Calibration Source Deployment, a method of deploying non-liquid sources into LZ through the TPC walls
DAQ	Data AcQuisition, components involved in the process of recording and storing data
DD	Deuterium-Deuterium, a neutron calibration source
DER	Detector Electronics Response, a method of modelling electronics and DAQ effects on simulated events
ER	Electron Recoil, the interaction of a particle with the electrons surrounding an atom
FV	Fiducial Volume, the region of a detector used for analysis due to its reliability
GdLS	Gadolimium-doped Liquid Scintillator, a mixture of Gd and water known for its high rate of fluorescence for low-energy interactions
GXe	Gaseous Xenon
iRRS	in-line Radon Reduction System, a system designed to reduce the radon levels in LZ
ICV	Inner Cryostat Vessel, the vessel used to store the primary LXe volume of LZ
KS	Kolmogorov-Smirnov, a statistical test comparing a sample to a presumed probability distribution through their CDFs

ΛCDM	Lambda cold dark matter
LUX	Large Underground Xenon experiment, a previous dual-phase LXe TPC designed to detect dark matter
LXe	Liquid xenon
LZ	LUX-ZEPLIN, a dual-phase LXe TPC designed to detect dark matter and the focus of this thesis
MACHO	MAssive Compact Halo Object, a term used to describe a variety of astrophysical objects which may or may not comprise a fraction of dark matter
Migdal effect	The hypothesised production of a possible secondary ER following a NR, which would improve sensitivity to low-energy signals where quenching prevents the detection of a NR but not an ER
MS	Multiple Scatter, an interaction which produces more than one S1 or more than one S2 signal
NEST	Noble Element Simulation Technique, a prolific toolset used to simulate noble-element energy deposition microphysics
NR	Nuclear Recoil, the interaction of a particle with the nucleus of an atom
OCV	Outer Cryostat Vessel, the outermost vessel in LZ
OD	Outer Detector, a GdLS chamber surrounding the primary LZ TPC, used primarily to tag neutrons entering the TPC
PBH	Primordial Black Holes, a hypothesised astrophysical object which may or may not comprise some or all dark matter
PDF	Probability Density Function, a function whose output for a given input can be taken to represent the relative likelihood of sampling this value from a given probability distribution
phd	photons detected
PLR	Profile Likelihood Ratio, a method of hypothesis testing which compares a given sample of data to expectation distributions in order to accept or reject~~validate or invalidate~~ a background-only hypothesis
PMT	PhotoMultiplier Tube, a vacuum phototube very sensitive to photons with a high gain
PTFE	Polytetrafluoroethylene

Quenching	The loss of energy in an interaction to sources other than scintillation and ionization
RFR	Reverse Field Region, a region of LZ below the cathode where the ionization signal of events drifts downwards, rather than upwards
ROI	Region Of Interest, the range of a parameter or set of parameters considered for a given analysis
S1c	Corrected primary scintillation signal area
S2c	Corrected secondary scintillation signal area
SR1	Science Run 1 of LZ, with data collected between December 23rd, 2021 to May 11th, 2022
SR3	Science Run 3 of LZ, with data collected from June 8th, 2023 to March 3rd, 2024
SURF	Sanford Underground Research Facility, the location of LZ
TPC	Time Projection Chamber, a type of particle detector used to reconstruct the position and trajectory of a given interaction in 3D
VUV	Vacuum UltraViolet, photons of wavelengths 10 nm to 200 nm
WIMP	Weakly Interacting Massive Particle, a leading candidate for dark matter and the focus of this thesis
XLZD	Xenon-LUX-ZEPLIN-Darwin, a proposed future LXe TPC observatory for dark matter and neutrino signals
ZEPLIN	ZonEd Proportional scintillation in LIquid Noble gases, a series of previous dual-phase LXe TPCs designed to detect dark matter

Introduction

The Standard Model of particle physics is the successful culmination of decades of experimentation and measurements, and is capable of explaining the vast majority of experimental observations in high-energy physics [1]. However, there are several regions where it is in tension with existing data. In the field of astrophysics, there is compelling evidence [2] pointing to the existence of a non-luminous substance comprising 84% of the total mass density of the known universe, referred to as 'dark matter'. Particle dark matter would be capable of resolving tensions in the Standard Model. While a worldwide effort has taken place to search for particle dark matter, it has so far not been directly observed.

The LUX-ZEPLIN (LZ) experiment [3] is designed to search for dark matter through its hypothesised interaction with ordinary matter, using seven active tonnes of xenon in a dual-phase time projection chamber. While primarily designed to identify Weakly Interacting Massive Particles (WIMPs), it is also capable of searching for lower-mass dark matter candidates.

The work presented in this thesis improves the sensitivity of LZ to low-mass dark matter candidates. This involves minimising the level of low-energy backgrounds, and precisely measuring detection efficiency in low-energy signals.

Chapter 1 outlines the key observational evidence motivating dark matter. This is then followed by a description of the physical constraints on dark matter.

The candidates LZ is specifically designed to detect are outlined, as are the candidates it is more sensitive to as a result of this work.

Chapter 2 details the principles of direct detection searches for dark matter.

This is followed by the projected sensitivity of LZ to dark matter, and the sensitivity of existing and upcoming dark matter searches in these mass ranges.

Chapter 3 outlines the LZ experiment, detector systems, operating principles, and analysis methods. The concept of signal production in liquid xenon within a dual-phase time projection chamber is discussed. This chapter concludes with an overview of the first science run of LZ, as well as the contributions of the work presented in this thesis to achieving this sensitivity.

Chapter 4 presents two analysis cuts for reducing the rate of accidental coincidence events in LZ, a dominant low-energy background after live-time cuts are applied to data. These cuts are demonstrated to be effective in both high-energy and low-energy physics searches in LZ. This chapter concludes by demonstrating that LZ is operating at the ideal drift field for low-mass WIMP searches, where accidentals are the dominant background.

Chapter 5 presents measurements of S2 pulse mis-classification in LZ. Measurements are made of signal inefficiency and S2 pulse area resolution using both calibration data and simulations. Candidate events can be used to tune pulse classification software for low-energy searches. Where mis-classification is irreducible, measurements can instead be used in statistical inference as a nuisance parameter, to minimise uncertainty on low-mass WIMP sensitivity.

Chapter 6 presents a series of analysis cuts for the rejection of grid radiogenic events in an electroluminescence-only (S2-Only) analysis. In early science runs with low statistics, a cut in S2 pulse width is demonstrated to reject 60% of grid backgrounds while preserving 80% of signal. In later science runs with higher statistics, a Boosted Decision Tree is shown to reject 80% of grid backgrounds while preserving 80% of signal. This chapter concludes by demonstrating that the required data for discovery of a $2\,\mathrm{GeV/c^2}$ WIMP is almost an order of magnitude lower with the application of these cuts.

The final chapter provides a summary of the findings of this thesis, highlighting how the low-energy analysis cuts and signal acceptance measurements developed in this work have maximised sensitivity to low-mass WIMPs in both standard and S2-Only LZ analysis. This concludes with an outlook on LZs sensitivity to light dark matter, and on the sensitivity of future dark matter observatories.

1

Dark Matter

Since humanity first looked up to the stars, it has contemplated just how vast the universe could be, and how much of it was yet to be understood. Advancements in science have allowed for the observation of more of this unknown, although there has always remained the impression that much still remains to be learned, imperceptible to the human eye.

While many of these discoveries were made through observation of celestial bodies, researchers have long posited that non-luminous matter exists which only acts gravitationally [4–6], referred to as 'dark matter' [7]. While many agreed at the time that the presence of dark matter was minimal [8, 9], approximately 84% of the universe's mass density is composed of dark matter.

This chapter outlines the key pieces of evidence motivating a hypothesis of dark matter, and provides a formal definition of its properties and role in the ΛCDM model of cosmology [10]. The chapter concludes with a description of dark matter candidates which the LUX-ZEPLIN (LZ) experiment [3] is sensitive to.

1.1 Experimental Evidence for Dark Matter

The hypothesis of dark matter made by the ΛCDM model of cosmology is motivated by a swathe of evidence from multiple areas of astrophysics and cosmology on several length scales. This section outlines the history of this evidence for galactic scales, the scale of galaxy clusters, and at a larger cosmological scale.

1.1.1 Galactic Scale

The first suggestion of non-luminous galactic matter came from early measurements of red shift from galaxies. The linear relationship between the red shift of a galaxy, its distance from earth, and its orbital velocity relative to the earth had only recently been demonstrated [11], but significant variations in these velocities were observed in 1931 [12].

For a galaxy with mass distribution $M(r)$, the orbital velocity $v(r)$ of its constituents can be described as

$$v(r) = \sqrt{\frac{GM(r)}{r}}, \tag{1.1}$$

where G is the gravitational constant and r denotes radial distance from the galactic centre.

As the average orbital velocity depends on the mass of the galaxy cluster, which was measured from spectroscopic studies, a variation in orbital velocity could imply one of two things: either the calculation used to make these predictions was incorrect, and would have to be modified accordingly, or there existed a population of non-luminous matter in each galaxy cluster, which can vary in mass.

More refined measurements of galaxy rotation curves later in the 20th Century helped demonstrate the latter of these two hypotheses. Baryonic matter in galaxies can be approximated to take the form of a sphere of radius R and constant density ρ. Inserting these expressions into Equation 1.1 leads to the conclusion that velocity should increase linearly with radius up to a fixed point R, after which it should decay as $v(r) \propto r^{-1/2}$.

However, this was demonstrated to not be the case in measurements of M31 (The Andromeda Nebula). The first discrepancies between this prediction and real data were observed by Horace Babcock in 1939 [13], with further evidence provided by radio measurements in 1966 [14].

In 1970, Vera Rubin and Kent Ford demonstrated that the velocity distribution of M31 did not align with predictions of galaxy rotation without the existence of non-luminous matter [15]. A discrepancy was also demonstrated by the pair in other galaxies in 1978 [16], with measurements from others in the astrophysics community corroborating this [17–19]. These measurements provided a strong case for large amounts of non-luminous mass spread diffusely throughout galaxies.

Figure 1.1 illustrates this concept. Data shows the orbital velocity around M33 (the Triangulum Galaxy). The disk of baryonic matter in M33 orbits at a radius of 3 kpc, with a ’bulge’ of stars distributed more uniformly [20]. The distribution of this matter does not reflect what is observed in data without the existence of non-luminous matter in a halo surrounding the galaxy, with density $\rho \propto r^{-2}$ far greater than the density of baryonic matter.

1.1.2 Cluster Scale

Clusters of galaxies also provide evidence for the existence of dark matter. In 1933, Fritz Zwicky continued the measurements of the Coma cluster made by Hubble in 1931. These measurements used the Virial theorem [23] to estimate the kinetic energy of the cluster, then compared this to the predictions made assuming the presence of only baryonic matter. From this, Zwicky obtained a mass-to-light ratio of 500 [24]. This suggested a significantly larger proportion of non-luminous mass driving the velocity dispersion, contrary to previous predictions that the amount of non-luminous matter was subdominant in galaxies [8, 9]. It was here that Zwicky referred to this mass as ’dark matter’.

The merging of galaxy clusters can shed more light on the nature of dark matter. The most notable example of this is the merging of galaxy clusters 1E0657-558, referred to as the ’Bullet Cluster’.

When galaxy clusters collide, the most massive visible component, the hot intra-cluster gas, interacts electromagnetically. This interaction can be observed via the associated

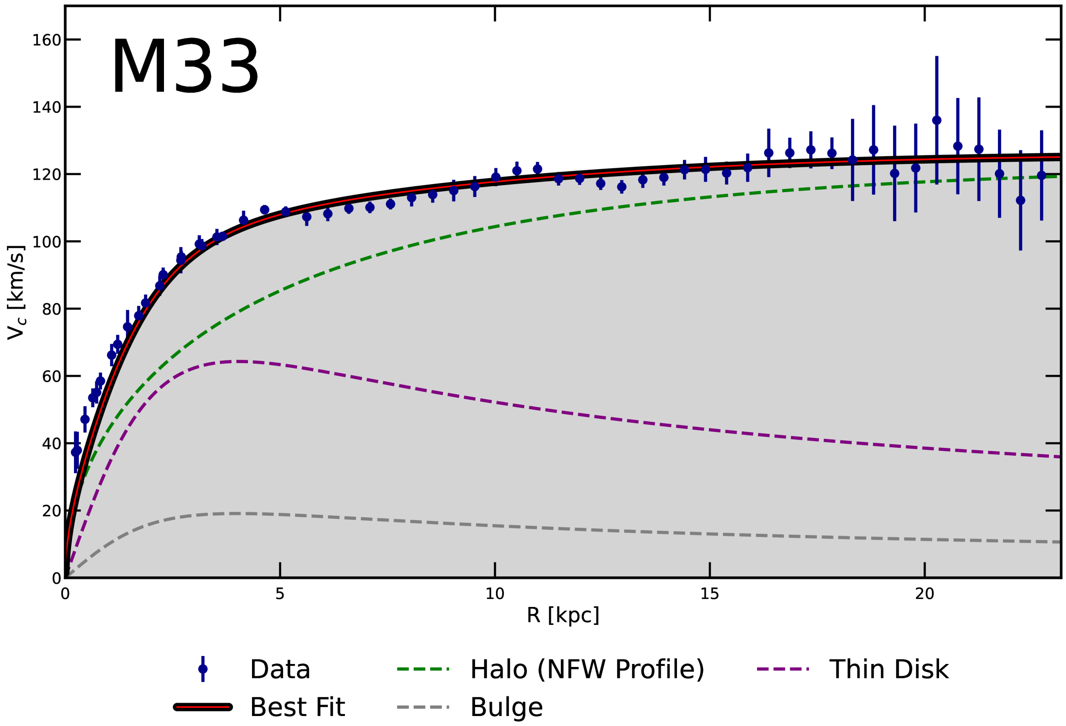


Figure 1.1: The rotational Velocity of M33 compared to models. The best fit (solid, red) is plotted alongside each contributing component, with contributions from both galactic baryonic matter (purple, grey) and dark matter (green), modelled using the NFW profile [21]. Generated using GalRotpy [22].

emission of photons in the x-ray and gamma range. The total mass, visible or otherwise, of the cluster can then me measured through gravitational lensing, a feature of general relativity where light from a background source is distorted proportionally by the presence of mass in the foreground. The results of these measurements can be seen in Figure 1.2.

Two key conclusions can be drawn from measurements of the Bullet Cluster. Firstly, the separation between the measured mass and light distributions indicates the presence of mass which behaves differently from visible matter, which is slowed due to frictional interaction, whereas the remaining mass of the galaxies continues unimpeded. Secondly, as the separation between these distributions is high, it can be assumed that the majority of the mass in a galaxy exists as a non-luminous halo, as opposed to originating in the luminous gas at the galaxy's core.

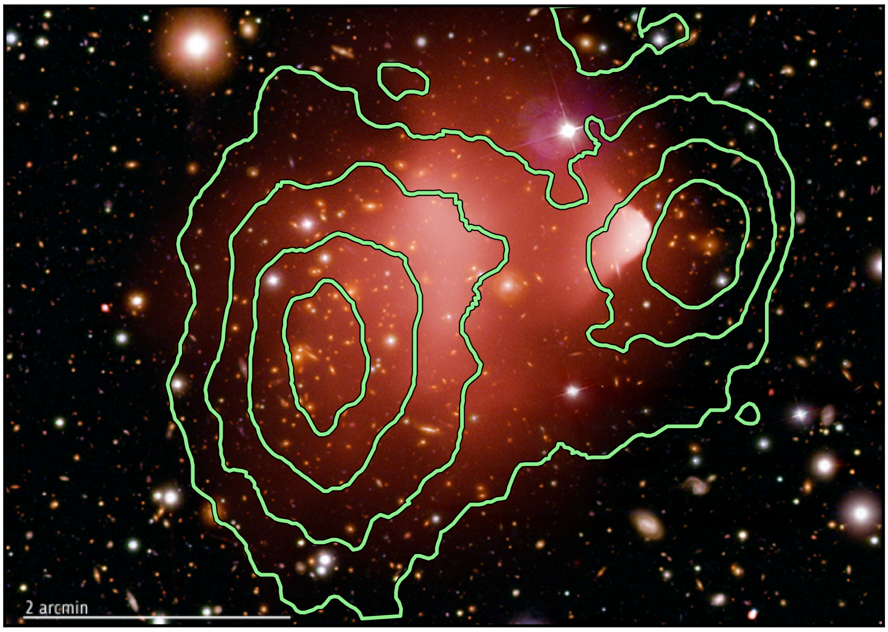


Figure 1.2: Background: composite image of the bullet cluster, combining X-Ray images from the Chandra X-Ray Observatory and Optical data from the Hubble Space Telescope [25]. Foreground: Contours mapping the mass distribution from the bullet cluster, measured through gravitational lensing. Digitized from data in Ref. [26].

1.1.3 Cosmological Scale

Approximately $1 \times 10^{-6}\,\mathrm{s}$ after the Big Bang, the universe had cooled from expansion to the point that quarks could begin to form hadrons, referred to as the hadron epoch. However, as electrons had not yet coupled to these hadrons, Thompson scattering heavily restricted the mean free path of photons, making the universe opaque.

Eventually, the universe expanded to the point that electrons could bind to hadrons to form hydrogen atoms, referred to as the epoch of recombination. At this point, the universe became transparent. The surface from which photon production at this epoch is detectable in the present day is referred to as the surface of last scattering of the universe. As the remaining photons from this surface have since red-shifted into the microwave regime, this radiation is now known as the Cosmic Microwave Background (CMB) [27].

Early measurements of the CMB found it to be uniform across the entire universe, resulting in a background temperature of $2.73\,\mathrm{K}$ [28, 29]. However, there did exist unexpected anisotropies in the measured spectra, which were measured with greater precision by Planck [30] and WMAP [31].

These $\mathcal{O}(10\,\mu\mathrm{K})$ fluctuations in temperature provide information about the surface of last scattering and, as a result, the origin of the universe. At large scales, fluctuations in temperature are caused by overdensity at the time of last scattering, causing photons to lose energy overcoming gravitational potential. Fluctuations of CMB temperature at large scales therefore provide information on the density of different regions of space.

At smaller scales, fluctuations are the result of Baryon Acoustic Oscillations (BAO), where photon-baryon systems oscillate under the forces of pressure and gravity until the epoch of decoupling [27]. Fluctuations in this regime therefore provide information on the baryon number in the early universe. These two measurements, in combination, allow for information to be extracted from the CMB on the ratio of baryonic to non-baryonic matter.

Figure 1.3 shows the distribution of CMB anisotropies in a power spectrum, where the multipole moment ℓ is inversely proportional to angle (i.e. a larger moment represents a smaller measured angle). Data is fitted to a functional form related to several cosmological parameters [32], allowing for those parameters to be measured using CMB data.

More recently, constraints on dark matter have been applied through the use of N-body simulations. Such models have helped to identify haloes of mass as low as $1 \times 10^{10}\,\mathrm{M}_{\odot}$, and match a large number of empirical trends observed in data such as the NFW profile [33, 34].

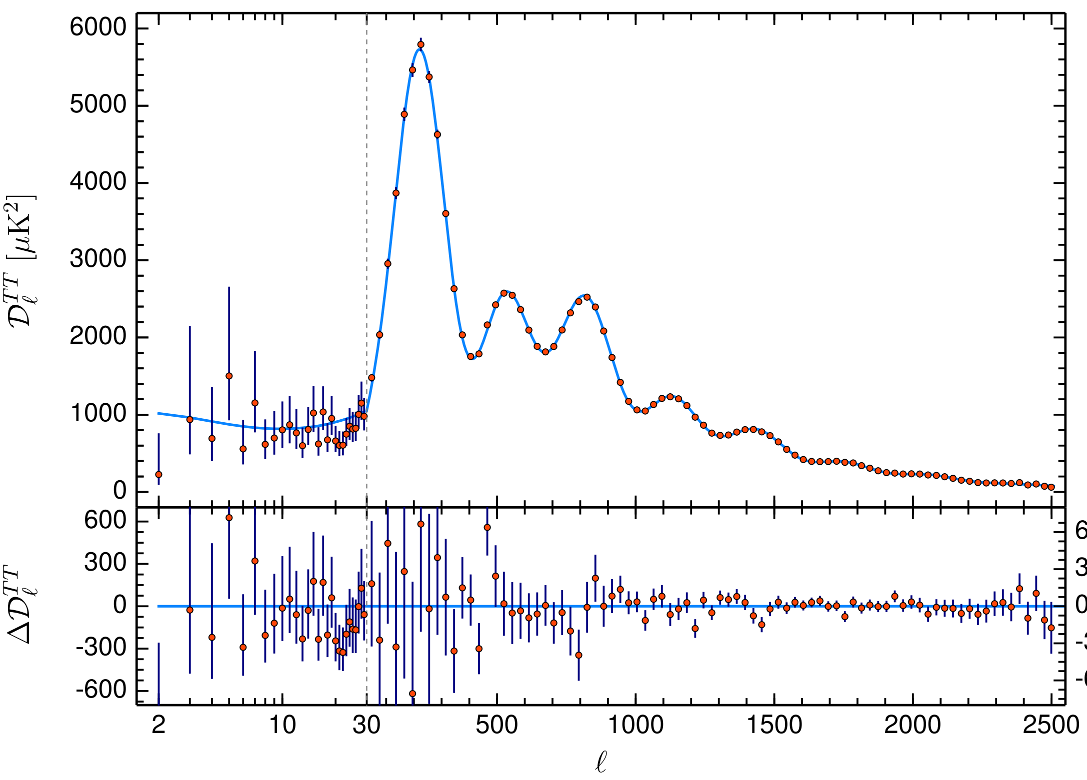


Figure 1.3: Angular power spectrum from Planck 2018 data, as a function of the spherical multipole moment ℓ. Data is fitted to parameters assuming base ΛCDM cosmology. Plot sourced from Ref.[32].

1.2 Constraints on Dark Matter

The observations provided in Section 1.1 do not constitute all evidence for dark matter. However, the plethora of evidence across multiple approaches to analysis and scales of physics provide ample motivation for a hypothesis: that there exists a form of matter, referred to as 'dark matter', that is widespread across the universe and significantly more abundant than baryonic matter. A number of properties of this dark matter can then be inferred from these observations:

- **Dark matter is dark.** It does not produce any observable signals, implying it cannot cool by radiating photons. This also implies it is electrically and chromodynamically neutral.
- **Dark matter is massive.** Its gravitational effects have been observed both

explicitly, as haloes surrounding galaxies, and implicitly, through its influence on the CMB.

- **Dark matter is non-baryonic.** It does not influence BAO in the CMB.
- **Dark matter is stable, or long-lived.** It continues to exist in large quantities, despite having been present during galaxy formation.
- **Dark matter is collisionless.** Haloes of dark matter do not visibly interact with one another or with baryonic matter. This implies the cross section for interactions is very low.
- **Dark matter is cold.** Some dark matter must travel non-relativistically, in order to form haloes around galaxies.

These assumptions, as well as the experimental constraints measured in order to make them, are referred to collectively as the Lambda Cold Dark Matter (ΛCDM) model of cosmology. This model assumes the existence of five constituents of the universe: dark matter, baryonic matter, CMB radiation, relic neutrinos decoupling shortly after the big bang, and dark energy. The latter is included to explain the accelerating expansion of the universe, through the re-introduction of the cosmological constant Λ [10].

The model also contains six independent free parameters. From these, the relative density of dark energy, dark matter, and baryonic matter $(\Omega_\Lambda, \Omega_c, \Omega_b)$ can be derived [35]. The most recent constraints from the Planck collaboration, based on the fit in Figure 1.3, are:

$$\begin{aligned}\Omega_\Lambda &= (0.688\,90 \pm 0.005\,60)\\ \Omega_c &= (0.260\,70 \pm 0.001\,99)\\ \Omega_b &= (0.048\,98 \pm 0.000\,31).\end{aligned}$$

The assumptions outlined above are typically assumed in research related to cosmology - however, they are not set in stone, and several assumptions are in tension with existing data. For example, one key assumption in the model is the Cosmological Principle, which states that the universe is spatially isotropic and homogeneous. This is in tension with studies into the CMB that have identified the presence of a dipole [36], as well as the observation of large-scale astronomic structures [37, 38], which appear to violate isotropy and homogeneity respectively. Studies to resolve these tensions are ongoing [39].

Other issues arise not from observed anomalies but from the inability to reconcile N-body simulations with observed dark matter distributions. Two examples of this are the overestimation of subhaloes and the difference in density distributions of dark matter, referred to as the 'missing satellite' problem and the 'cuspy halo' problem respectively [40, 41]. These apparent contradictions can possibly be resolved by loosening the assumptions made in ΛCDM: for example, the existence of both relativistic and non-relativistic ('warm') dark matter could resolve the cuspy halo problem [42], whereas lifting the requirement for dark matter to not self-interact can resolve issues with the number of observed satellites [43].

1.2.1 Mass Constraints

The ΛCDM model of dark matter assumes a particle candidate is non-relativistic, with a mass in the keV-TeV range. However, several methods exist of placing additional constraints on the mass of dark matter using experimental evidence, given a number of base assumptions. In order for a dark matter candidate to stray outside of these boundaries, a number of additional assumptions must be made about its properties.

Absolute Constraints

Many hot dark matter models assume the existence of a scalar field which both produced dark matter and drove the epoch of inflation. A candidate of particularly low mass would affect the noise of density perturbations after this epoch, which would be observable. The mass of such candidates therefore cannot be lower than approximately $1 \times 10^{-19}\,\mathrm{eV/c^2}$ [44].

At much higher mass ranges on the order of 1 g, dark matter would take the form of clusters of baryonic or non-baryonic particles, referred to as composite or macroscopic dark matter [45, 46]. Such particles are heavily restricted by constraints on baryon density from the CMB and thermonuclear runaway in white dwarfs [47].

Freeze-Out

If dark matter is assumed to be weakly interacting and present at the beginning of the universe, its domination over baryonic matter can be explained through what is referred to as the freeze-out mechanism [48]. Dark matter is assumed to be in thermal equilibrium with other matter in the baryonic plasma of early epochs, with equilibrium maintained through pair-production and pair-annihilation.

As the universe cools, the energy of each particle decreases. As a result, the number of dark matter particles is Boltzmann-suppressed, and dark matter pair-production drops exponentially. However, the decreasing number of dark matter particles and the expansion of the universe both lower the probability of pair-annihilation. Eventually, this causes the two processes to re-enter equilibrium, at which point number density becomes constant and dark matter is said to have 'frozen out'.

If dark matter is of a higher mass than most baryonic particles, this process would occur later than for baryonic matter, resulting in less annihilation of dark matter. This would explain the higher proportion of dark matter in the current universe.

Using the equations of thermodynamics, the Boltzmann equation, and the equations of cosmological expansion, dark matter abundance Ω_c can be expressed in this framework as

$$\Omega_c \approx \frac{T_0^3}{\rho_c M_p} \frac{x_f}{\langle \sigma v \rangle}, \tag{1.2}$$

where T_0 is the current measured temperature of 2.73 K, ρ_0 is the critical density, M_p is the Planck mass of $1.22 \times 10^{19}\,\mathrm{GeV/c^2}$, $\langle \sigma v \rangle$ is the thermally-averaged pair-annihilation cross section, and $x_f \equiv m_\chi / T_f$ is a ratio of dark matter mass and the temperature of freeze-out [48].

Figure 1.4 illustrates how number density evolves with time for a given mass of particle. The position at which comoving number density (Y) remains constant changes for a given value of $\langle \sigma v \rangle$. Measurements of Ω_c in the early and current universe can therefore place constraints on σ and v.

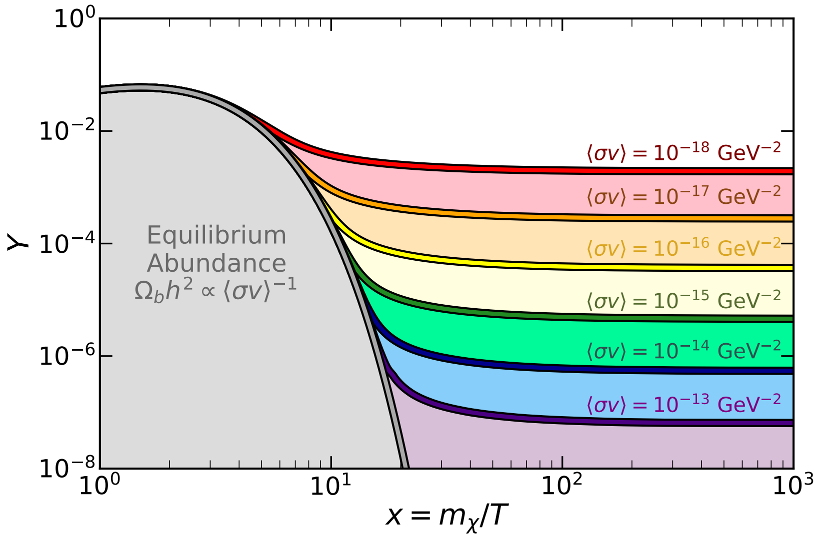


Figure 1.4: Comoving number density (Y) evolution as a function of the ratio $\frac{m_\chi}{T}$, in the context of thermal freeze-out for a 80 GeV/c^2 WIMP. The size of the annihilation cross section determines abundance, since $\Omega_{DM} \propto \langle\sigma v\rangle^{-1}$. Adapted from Ref.[49], using formulae from [50].

The lowest value of σ permitted for electroweak mediators leads to the conclusion that a neutral lepton lighter than 2 GeV/c^2 could not freeze-out in a manner compatible with existing measurements of dark matter. This is referred to as the Lee-Weinberg limit [51]. On the opposing end of the mass spectrum, constraints on the maximum value of $\langle\sigma v\rangle$ lead to a unitarity constraint at a mass between 96 TeV/c^2 and 144 TeV/c^2 [52].

The unitarity bound on freeze-out sets a hard constraint on non-composite particle dark matter. In contrast, the Lee-Weinberg limit can be circumvented with additional assumptions about the properties of dark matter. Candidates with these properties will be discussed in Section 1.3, alongside those satisfying the minimal freeze-out mechanism.

1.3 Dark Matter Candidates

A wide spectrum of dark matter candidates have been proposed, both particles between the wide mass ranges of 1×10^{-19} eV/c^2 to 1×10^{30} eV/c^2, as well as astrophysical objects of masses up to 1×10^{36} g [53]. Specific candidates are motivated by various lingering problems in astroparticle physics. Several alternative hypotheses to non-baryonic dark matter have also been proposed, such as baryonic candidates or a modification to the current understanding of gravity. These are discussed in Section 1.3.1, while particle dark matter candidates are discussed in Section 1.3.2.

1.3.1 Alternative Explanations to Particle Dark Matter

Baryonic Explanations for Dark Matter

Several astrophysical objects are known to exist which do not emit light, such as white dwarfs, brown dwarfs, and baryonic black holes. As these objects will contribute to mass distributions of galaxies without being observable, they fall under the broad category of dark matter. These are collectively referred to as 'MAssive Compact Halo Objects' (MACHOs). While MACHOs are known to exist, constraints from the CMB require that they compose only a small fraction of the total dark matter density [54].

Primordial Black Holes

During the epoch of inflation, density perturbations existed such that large scale structures began to form. If a region of space was sufficiently overdense at the epoch of the surface of last scattering, such that its density surpassed the Chandrasekhar limit, gravity would have overcome internal forces and a black hole would be formed. As this black hole formed from non-baryonic matter, it would therefore satisfy the conditions for dark matter. Such black holes are referred to as Primordial Black Holes (PBHs).

Limits on the fraction of dark matter which can be composed of PBHs have been imposed through microlensing, gravitational waves, and accretion disks. Constraints from current measurements require that PBHs cannot form the entirety of dark matter in the universe, unless they exist in the mass range 1×10^{18} g to 1×10^{21} g, referred to as the 'asteroid gap' [53].

Modified Newtonian Dynamics

Some researchers have proposed that, on large length scales, gravity may exhibit slightly different properties. This theory is referred to as 'MOdified Newtonian Dynamics'

(MOND). MOND is capable of creating flat galaxy rotation curves matching astrophysical observations. However, it is incapable of explaining CMB anisotropies [55].

1.3.2 Particle Dark Matter

While alternative hypotheses to particle dark matter can resolve some questions formed from astrophysical observations, none have been capable of resolving all of them. In contrast, several as-yet-undiscovered particle species could resolve these issues. Such particles are well-motivated not only from astrophysical observations, but also in the field of high-energy physics. Figure 1.5 displays the broad spectrum of particle candidates for dark matter. LZ is designed to be sensitive to WIMPs, but is also sensitive to lighter dark matter candidates.

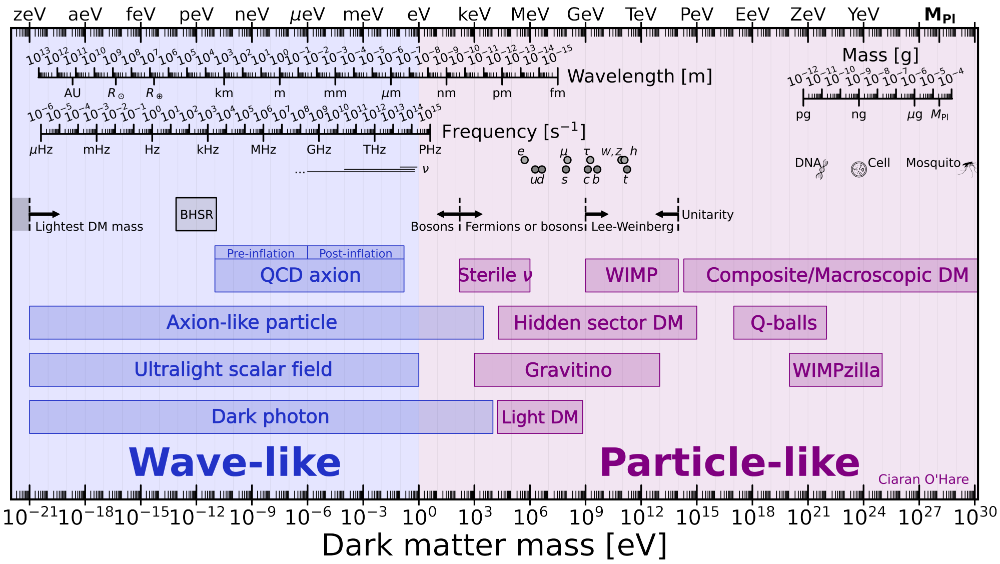


Figure 1.5: Illustration of the mass range of dark matter candidates, both wave-like and particle-like. Several points of reference are annotated, such as the Lee-Weinberg limit [51], SM particle masses, black hole superradiance [56], and a scale of wavelengths and frequencies for wave-like particles. Adapted from Ref.[57].

WIMPs

The freeze-out mechanism outlined in Section 1.2 predicts that a dark matter particle which weakly interacts, self-interacts, and exhibits no other anomalous properties, must have a mass between 2 GeV/c^2 and 130 TeV/c^2. Particles exhibiting these properties have also been proposed as solutions to several issues within the Standard Model, a coincidence referred to within the community as the WIMP miracle [50].

Solutions to the Hierarchy Problem [58] in the Standard Model [1] typically propose the existence of a partner for every particle, with a greater mass and alternate spin. This extension to the Standard Model is referred to as Supersymmetry [58]. The Hierarchy Problem can also be resolved through the introduction of extra dimensions to the Standard Model Lagrangian. These extra dimensions would result in the production of particles in higher dimensions observable as a unique non-baryonic form of matter, referred to as Kaluza-Klein matter [59, 60]. Another solution to the Hierarchy Problem proposes that the Higgs field contains an accidental symmetry: such a symmetry would also result in the existence of a second scalar particle coupling to the Higgs field, referred to as the ’little higgs’ [61, 62]. Each of these solutions to the hierarchy problem proposes the existence of particles which satisfy the conditions required of a WIMP.

The term ’WIMP’ is often defined very broadly in literature, with varying scope. For this thesis, ’WIMP’ refers to any particle satisfying the conditions of freeze-out with no additional properties. While the Lee-Weinberg limit formally constrains particle masses to above $2\,\mathrm{GeV/c^2}$, it is generally accepted by the community that a particle above $1\,\mathrm{GeV/c^2}$ will satisfy these conditions [63]. A particle of mass 1 GeV may still weakly interact and freeze out, but would not be considered a WIMP, as additional assumptions are needed.

Asymmetric Dark Matter

The WIMP model of dark matter assumes a candidate particle does not display any interaction asymmetries. This is unlike baryonic matter, which both violates CP symmetry and contains an asymmetry in baryon-lepton interactions. Were such an asymmetry to exist in dark matter also, it would result in an earlier freeze-out, allowing for particles of lower mass to exist in the early universe [64]. Many models of light dark matter particles exist that would satisfy the conditions of asymmetric dark matter, particularly in the mass range $5\,\mathrm{GeV/c^2}$ to $15\,\mathrm{GeV/c^2}$: these include mirror dark matter, darkogenesis, and pangenesis [65].

Light Dark Matter

Dark matter which couples to a fermion or a new dark gauge boson would be capable of freeze-out at masses below 2 GeV, as this coupling would affect the equilibrium in the early universe. Such particles would be constrained to masses below 100 MeV/c^2, above which the rate of gamma production from this coupling would be in tension with astrophysical observations [66]. Elastic scattering between baryonic and non-baryonic matter would also allow for particles in this mass range of 1 MeV/c^2 to 100 MeV/c^2 [67].

Alternatively, it is possible that dark matter instead froze in. Particles of a sufficiently low cross section would be effectively decoupled from the early plasma, and would increase in number over time rather than decreasing as in the freeze-out mechanism. Freeze-in models allow for the existence of non-baryonic particles well below the Lee-Weinberg limit [68].

Hidden Sector

While it is widely accepted that the dominant form of dark matter is cold, due to the evolution of cosmic structure, it is possible for relativistic dark matter to also exist in smaller quantities in a warm dark matter model. Such candidates are generally referred to as the hidden sector. Particles in this region are not only capable of resolving lingering issues in the Standard Model, such as the strong CP problem [1], but may also have helped to drive cosmological inflation. Popular candidates include Axion-Like Particles (ALPs) [69], Hidden Photons [70], and Sterile Neutrinos [71].

1.4 Sensitivity of LZ to Dark Matter Candidates

The LUX-ZEPLIN experiment, the focus of this thesis, is primarily designed for the detection of WIMPs. As standard analysis extends from 8 GeV/c^2 to 1×10^4 GeV/c^2, this encompasses the majority of the range of potential WIMP masses, and probes some of the mass range of asymmetric dark matter.

The sensitivity of LZ to masses above WIMP-scale is constrained by the active

LXe mass. At lower energies, however, LZ can detect dark matter candidates through alternative analysis methods. For example, through analysis of the ER band of S1-S2 events, LZ is sensitive to candidates from the hidden sector [72].

Many of these are only accessible through the use of alternative analysis methods, such as the exploitation of the double photo-electron (DPE) and Migdal effect, or the use of only electroluminescence pulses. Such analyses are capable of searching for weakly interacting dark matter candidates of mass as low as $80\,\mathrm{MeV/c^2}$ [73]. LZ is therefore sensitive to all of the light dark matter candidates outlined in this chapter.

2

Direct Dark Matter Detection

Although many models of dark matter have been proposed, observational evidence remains elusive. However, several methods exist of searching for candidates, through the comparison of data to a background-only hypothesis.

This chapter reviews the detection methods of direct, indirect, and collider dark matter searches. The Profile Likelihood Ratio is discussed, a method of comparing data to a background-only hypothesis in order to assess the probability of having observed a dark matter signal. This chapter then concludes with an overview of the sensitivity of LZ to dark matter, with comparison to previous, current, and upcoming direct dark matter detectors.

2.1 Detection Methods

Chapter 1 highlights how astrophysical and cosmological data present evidence of dark matter and constrain its properties. However, many constraints have been placed by experiments in the field of particle physics, through standard model particles. Figure 2.1 illustrates this concept. The different directions in which this diagram can be interpreted allow for three different approaches to detecting dark matter.

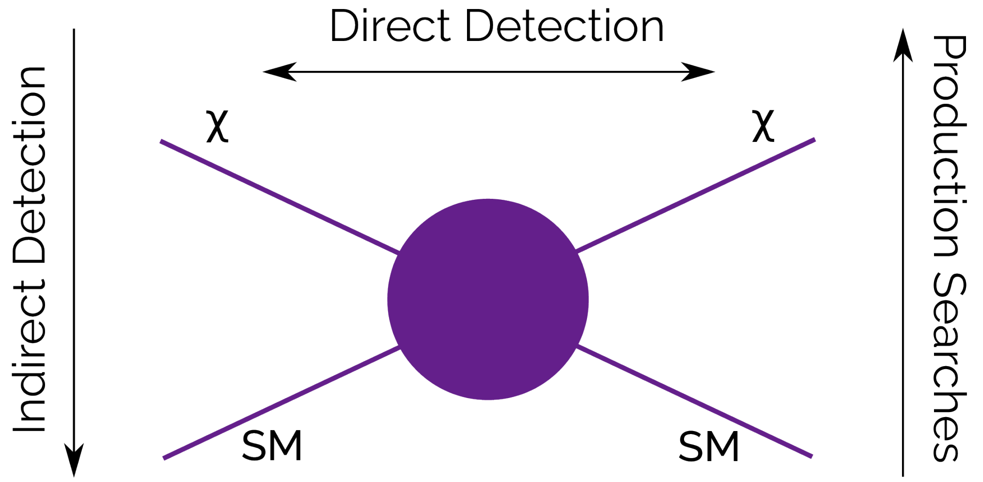


Figure 2.1: Illustration of the interaction of dark matter with baryonic matter. The direction in which this diagram is read illustrates the channels through which a dark matter search can be conducted.

One route to detecting dark matter is its production from the interaction of SM particles, seen when reading Figure 2.1 upwards. Any experiment looking into the interaction of SM particles at a high energy is therefore capable of searching for dark matter candidates. Many examples of this can be found at the Large Hadron Collider, with WIMP searches conducted in ATLAS [74], dark sector searches in CMS [75], dark photon searches in LHCb [76], axion-like particle searches in FASER [77], and long-lived particle searches in CODEX-b [78].

Alternatively, the annihilation of dark matter may result in an observable baryonic signal, seen when reading Figure 2.1 downwards. Many experiments in astroparticle physics are dedicated to searching for such processes, through the detection of cosmic photons, cosmic rays, and neutrinos [79].

Finally, dark matter can be searched for through its direct interaction with baryonic matter, seen when reading Figure 2.1 horizontally. Experimental searches for signals from these interactions are categorised as direct detection experiments. The direct detection of WIMPs and light dark matter candidates, particularly in liquid xenon in the LUX-ZEPLIN experiment, is the focus of this thesis.

2.2 Concept of Direct Detection

2.2.1 WIMP-nucleus Scattering

The following discussion of WIMP direct detection follows Refs. [80–82].

The rate of interaction of dark matter with a detector can be expressed as the differential

$$\frac{dR}{dE_{nr}} = \frac{\rho_0 M}{m_N m_\chi} \int_{v_{min}}^{v_{esc}} v f(v) \frac{d\sigma}{dE_{nr}} dv, \tag{2.1}$$

where dark matter of mass m_χ, density ρ_0, and velocity distribution $f(v)$ deposits energy E_{nr} in a target medium of nuclear mass m_N and total mass M. The details of dark matter's interaction with standard model particles are represented through the differential cross section $\frac{d\sigma}{dE_{nr}}$.

When fully integrated between the minimum and escape velocities v_{min} and v_{esc}, this expression takes the form

$$\frac{dR}{dE_{nr}} \propto exp\left(-\frac{E_{nr}}{E_0}\frac{4m_\chi m_N}{(m_\chi + m_N)^2}\right), \tag{2.2}$$

where E_0 is the expectation energy of the recoil. This interaction rate decreases exponentially for decreasing values of $\frac{-E_{nr}}{E_0}$, highlighting how the energy threshold of an experiment is the limiting factor on its overall sensitivity at low masses.

Equation 2.1 can alternatively be expressed as a sum of spin-dependent (SD) and spin-independent (SI) components,

$$\frac{d\sigma(E_{nr})}{dE_{nr}} = \frac{m_N}{2v^2\mu^2}\left[\sigma_{SI} F_{SI}^2(E_{nr}) + \sigma_{SD} F_{SD}^2(E_{nr})\right], \tag{2.3}$$

where μ denotes the reduced mass of the WIMP-nucleon system, σ denotes cross sections, and F denotes form factors.

The ΛCDM model of cosmology and the dark matter candidates outlined in Chapter 1 assume a spin-independent elastic scattering dark matter particle. As a result, this discussion will not consider the spin-dependent component of Equation 2.3, although it is detailed in Refs.[80–82].

The spin-independent form factor typically chosen for WIMP searches is the analytical approximation of Helm [81, 83]. This takes the form

$$F(x) = \frac{3j_i(x)}{x} \exp \frac{-(xs)^2}{2R_N^2}, \tag{2.4}$$

where x is the dimensionless parameter $qR_N/\hbar$, $j_i(x) = \sin x/x^2 - \cos x/x$ is the spherical Bessel function of the first kind, R_N denotes the nuclear radius and s the skin thickness. R_N is typically $\approx 1.2A^{1/3}$fm and s $\sim 0.5\,$fm, although these parameters depend on the chosen nucleus.

The spin-independent cross section is given by

$$\sigma_{SI} = \sigma_n \frac{\mu^2}{\mu_n^2} \frac{(f_p Z + f_n(A - Z))^2}{f_n^2}, \tag{2.5}$$

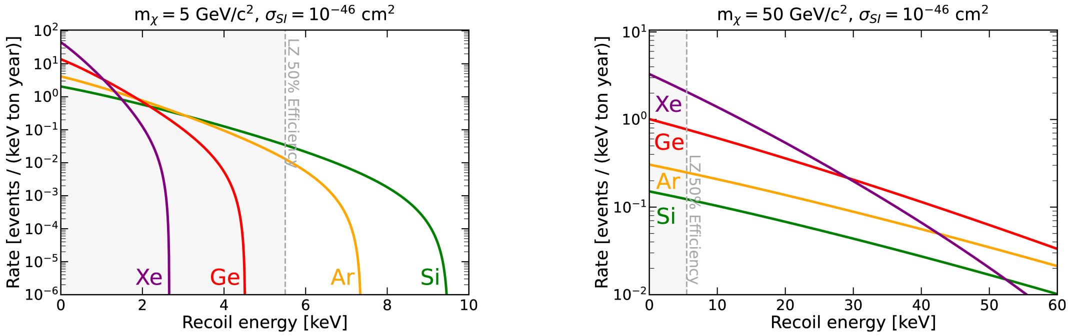
where μ_n and σ_n are the reduced mass and cross section of the WIMP-nucleon system, and f_p and f_n are coupling constants to protons and neutrons for a nucleus of atomic number Z and atomic mass A.

If it is assumed that $f_p \approxeq f_n$, this expression can be reduced to the form

$$\sigma_{SI} = \sigma_n \frac{\mu^2}{\mu_n^2} A^2. \tag{2.6}$$

This expression is proportional to A^2. A heavier target nucleus will therefore increase dark matter interaction rates. The proportionality to $\frac{\mu^2}{\mu_n^2}$ results in this rate being greatest when the mass of a dark matter candidate is similar to the target mass. However, the recoil energy of an interaction also decreases for increasing A. While the rate of events is greater in heavier nuclei, the rate of observed events is dependent on the energy threshold of a detector.

Figure 2.2 illustrates this concept, through the recoil spectra of WIMPs of mass 5 GeV/c^2 and 50 GeV/c^2. In the case of the 5 GeV/c^2 WIMP, the event rate in Xe is high overall. However, the majority of this integrated rate is well below the detection threshold of LZ. For WIMPs of this mass, sensitivity is dependent on signal acceptance at low energies. In contrast, in the case of the 50 GeV/c^2 WIMP, the majority of the integrated rate is significantly above threshold. At this mass, the limiting factor to sensitivity is leakage from background events.



(a) 5 GeV/c^2 WIMP **(b)** 50 GeV/c^2 WIMP

Figure 2.2: NR spectra for an assumed spin-independent WIMP-nucleon cross section of 1×10^{-46} cm^2, generated using wimprates [84]. A vertical line denotes the energy at which the detection efficiency of LZ to NRs is 50% [85]. Calculations assume the Helm form factor from Lewin & Smith [81]. Left: A 5 GeV WIMP. The energy of interactions is below the detection threshold of LZ. Sensitivity is dependent on maximising detection efficiency. Right: A 50 GeV WIMP. Many interactions are above the detection threshold of LZ. Sensitivity is dependent on minimising background leakage.

2.2.2 Statistical Inference in Direct Dark Matter Detectors

Once the rate of dark matter interactions for a given nucleus is determined, a method must be decided on with which to measure the probability of a dataset containing an excess of events compared to simulated backgrounds. This statistical analysis can then be used to measure the likelihood of a dataset containing a dark matter signal of a given cross section and mass, based on these models. This section reviews material from Refs. [86–88].

The Profile Likelihood Ratio

In order to assess the significance of any evidence (or lack thereof) of dark matter, direct detectors record interactions over a period of time, then compare this data to a background-only hypothesis, in order to assess the probability of a signal. In previous experiments with limited statistics, this has been achieved with a simple "cut-and-count" approach, where aggressive cuts are applied to data in order to reject all background events. However, in larger modern detectors, it has been agreed to use a Profile Likelihood Ratio (PLR) approach, in order to make it easier to directly compare results from multiple direct dark matter searches.

Figure 2.3 shows a simplified illustration of a direct dark matter detector dataset. Data is distributed in two metrics: the energy of an interaction, and a separate discrimination parameter. This discrimination parameter separates background and signal events, and signal is predicted at lower energies: however, the rate of signal events is low enough that it is hard to discern by eye if events in the signal region are signal events, as opposed to simply background leakage.

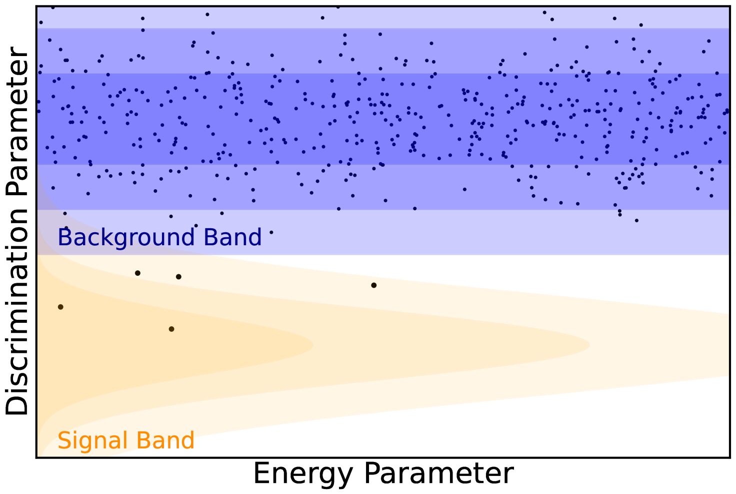


Figure 2.3: A simple illustration of data from direct detection experiments. A contour of expected background events (blue) and signal events (orange) are well separated by a discrimination parameter. Background events are evenly distributed in an energy parameter, while signal events tend towards lower energies. Black points indicate the position of detected events.

The majority of direct dark matter detectors use forms of energy deposition as a discrimination parameter. However, some experiments instead use time as a parameter, as an annual modulation is expected of dark matter due to the motion of the earth [89]. Some proposed experiments use directionality as an input, which would be capable of discriminating signal from neutrino backgrounds [90–92].

The parameters of interest μ, and measured experimental uncertainties (nuisance parameters) θ, can be inserted into a profile likelihood ratio. This is defined as

$$\lambda(\mu) = \frac{\mathcal{L}(\mu, \hat{\hat{\theta}})}{\mathcal{L}(\hat{\mu}, \hat{\theta})}, \tag{2.7}$$

with the likelihood function $\mathcal{L}$, maximum likelihood at $(\hat{\mu}, \hat{\theta})$, and likelihood for a given μ at $\hat{\hat{\theta}}$, the set of parameters of interest and nuisance parameters which maximize the likelihood function. λ takes a value between 0 and 1, with a value close to 1 signifying agreement between data and μ.

Given that the number of dark matter counts in a dataset can only be additive, the statistic $\tilde{t}_\mu$ can be defined as

$$\tilde{t}_\mu = \begin{cases} -2\log\frac{\mathcal{L}(\mu,\hat{\theta})}{\mathcal{L}(\hat{\mu},\hat{\theta})} & \hat{\mu} \geq 0, \\ -2\log\frac{\mathcal{L}(\mu,\hat{\theta})}{\mathcal{L}(0,\hat{\theta}(0))} & \hat{\mu} < 0. \end{cases} \tag{2.8}$$

The degree of disagreement between observed data and a hypothesis can be tested using the p-value, which corresponds to the probability of getting a value t_μ for a given μ as large or larger than that observed:

$$p(t_\mu|\mu) = P(t_\mu \geq t_{\mu,obs}|\mu) = \int_{t_{obs}}^{\infty} f(t_\mu|\mu) dt_\mu, \tag{2.9}$$

where $f(t_\mu|\mu)$ is the probability density function for t_μ.

From Equation 2.9, a distribution p_μ can be calculated, varying only the number of predicted dark matter counts μ for a given dark matter mass. This distribution follows a beta distribution.

Figure 2.4 illustrates this concept with two examples. In the first, the highest value of p_μ is at zero, and drops below 10% at approximately 2 counts: above this upper limit (UL), it can be said that were a background-only dataset to be acquired, a count above this would be measured in less than 10% of instances. In the second example, the mean is not centred around zero. In this instance, a lower limit (LL) is also capable of being set, where the probability of a dataset being generated with a rate of signal-like counts below this is less than 10%.

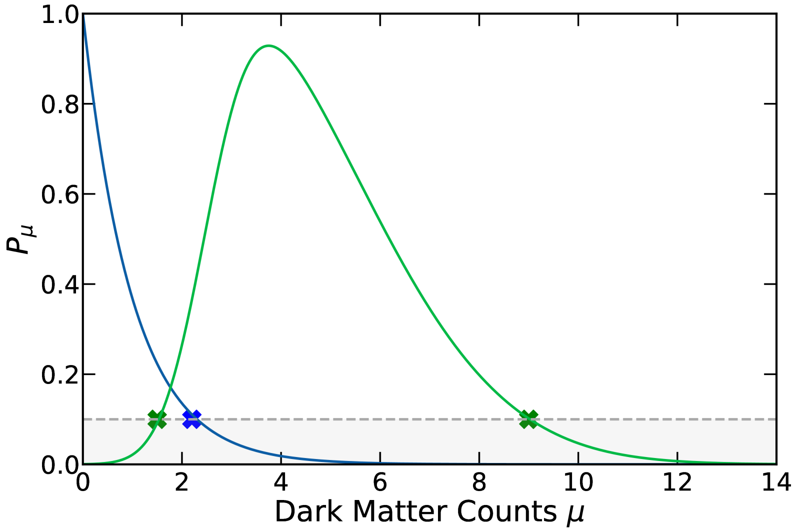


Figure 2.4: Simple illustration of the methodology for measuring the number of dark matter counts μ in a given dataset, using the p-value curve p_μ. Two examples are given. In the first example (blue), the most probable number of counts is given at around 0, with this decreasing below 10% at 2, allowing an upper limit on signal counts to be established. In the second example (green), the most probable number of counts is non-zero: as a result, both an upper limit and a lower limit can be established.

These distributions can be used to measure the number of dark matter interactions of a given mass in a dataset, and to project the sigma-level confidence of such a signal [86–88]. However, until such a discovery is achieved, direct dark matter searches instead prioritise the calculation of sensitivity to signal.

Dark Matter Sensitivity

The sensitivity of a direct detection experiment can also be defined using a PLR test and the above value of p_μ. For a given mass of dark matter, toy model data can be generated as in Figure 2.3, forming a distribution as in Figure 2.4, from which an upper limit can be extracted. This can be repeated for many toy models to generate a cumulative distribution function (CDF) of upper limits. From this, the mean and percentile ranges of limits can be calculated: the cross section at which a dark matter signal could be rejected from a dataset of this size with 90% confidence.

At low values of μ, a distribution with and without signal become very similar to one another. At this limit, it is possible for values of μ to be excluded to which an experiment is not sensitive, a concept referred to as spurious exclusion. Two distinct approaches exist of addressing this, both involving the comparison of the distribution of p_μ under the signal hypothesis μ and background-only hypothesis $\mu = 0$ for varying $\tilde{t_\mu}$.

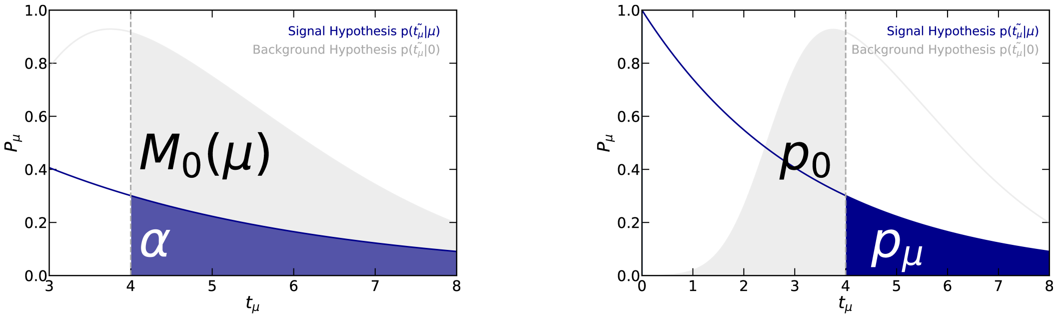
The primary method suggested for addressing the issue of spurious exclusion by Ref. [86] is that of power constrained limits. In this approach, the value of p_μ is integrated in $\tilde{t_\mu}$ between a chosen value and infinity for both a predicted dark matter count μ and zero, giving parameters α and M_0 respectively, corresponding to the rate of false positives from fluctuations in data and false negatives. Values of μ are then excluded from distributions if the value p_μ is lower than α, or if M_0 exceeds a chosen threshold [93], chosen to mitigate false positives while maintaining power to detect true positives.

An alternative method to avoid spurious exclusion is the CL_S approach, used widely in the field of particle physics [94]. In this approach, the parameter α is defined as before (though instead labelled p_μ, the p-value for a signal hypothesis): however, instead of integrating from a value upwards for M_0, the integral from 0 up to that value is taken as p_0, the p-value for a null hypothesis. Confidence intervals are then defined as CL_S,

$$CL_S = \frac{p_\mu}{1 - p_0}, \tag{2.10}$$

such that the likelihood of excluding a signal hypothesis is reduced to account for statistical fluctuations.

Illustrations of these concepts are shown in Figure 2.5.



(a) Power Constraint

(b) CL_S Method

Figure 2.5: Illustrations of two methods of preventing spurious exclusion in PLR tests of direct detection data. The details of these approaches are outlined in Section 2.2.2.

LZ currently produces limits using both power constraints and the CL_S method, for comparison with all other direct dark matter detection experiments.

The first science run of LZ measured WIMP sensitivity using this approach. In future science runs of LZ, profile likelihoods will be calculated through the convolution of analytic probability elements, using the FlameNEST framework, instead of using computationally expensive toy models [87, 95].

The design of the PLR method of statistical inference provides insight into the requirements of a direct dark matter detector. To maximise sensitivity of an experiment, the number of background events must be minimised, such that the distribution of p_μ is consistently close to zero. The width of this distribution of p_μ is dependent on nuisance parameters, such as the signal detection efficiency and precision with which the energy of an event is measured. Accurately measuring signal efficiency and energy uncertainty in data therefore improves both the mean sensitivity and the uncertainty of this sensitivity by constraining $\hat{\theta}$, as does lowering the energy threshold of an experiment.

2.3 Detector Technologies

2.3.1 History of Detector Operation

The field of direct detection of dark matter was first conceived in the 1980s, and has evolved over time with improving technology. Figure 2.6 displays the sensitivity of direct detectors over time, separated by detection medium. Initially, direct detectors were crystalline, using NaI and Ge as a target nucleus. These were quickly surpassed by cryogenic detectors, using phonon readout at temperatures $\mathcal{O}(\text{mK})$. These detectors have excellent signal discrimination: however, the rate of background events is difficult to reduce, and became a limiting factor to sensitivity. In contrast, liquid noble detectors have poorer discrimination, but can benefit at large scales from the introduction of a fiducial volume. Once it became possible to maintain an electric field at large scales and fully purify noble liquids, the majority of the community pivoted to the use of liquid noble TPCs.

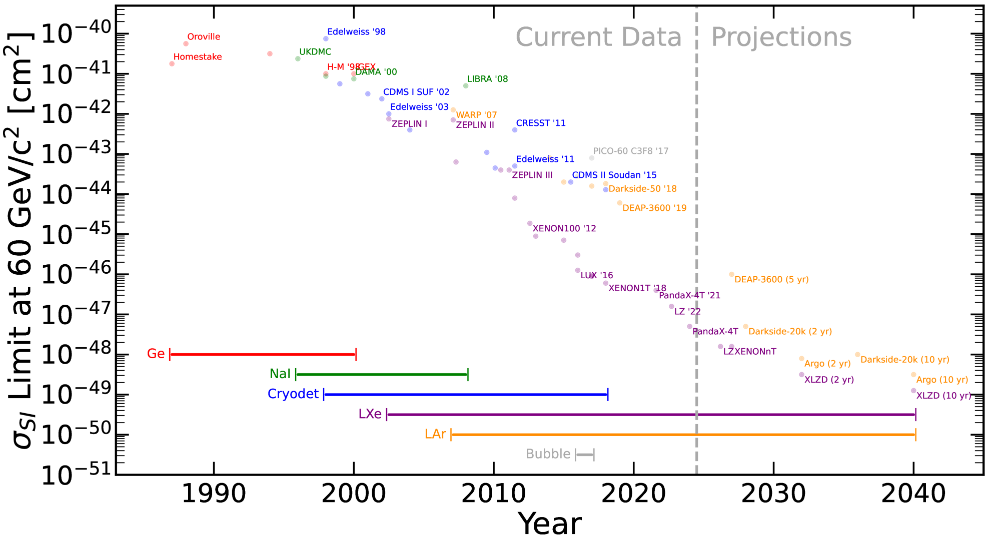


Figure 2.6: Limits set on a spin-independent WIMP-nucleon cross section for a 60 GeV/c^2 WIMP over time, from different direct detection experiments. Several experiments are annotated. Data provided by the LZ collaboration.

2.3.2 Sub-TeV Mass

Several liquid noble experiments search for WIMPs in the GeV-TeV scale, through the detection of scintillation (S1) and electroluminescence (S2) light. Figure 2.7 shows the sensitivity of current leading experiments in the liquid noble TPC regime.

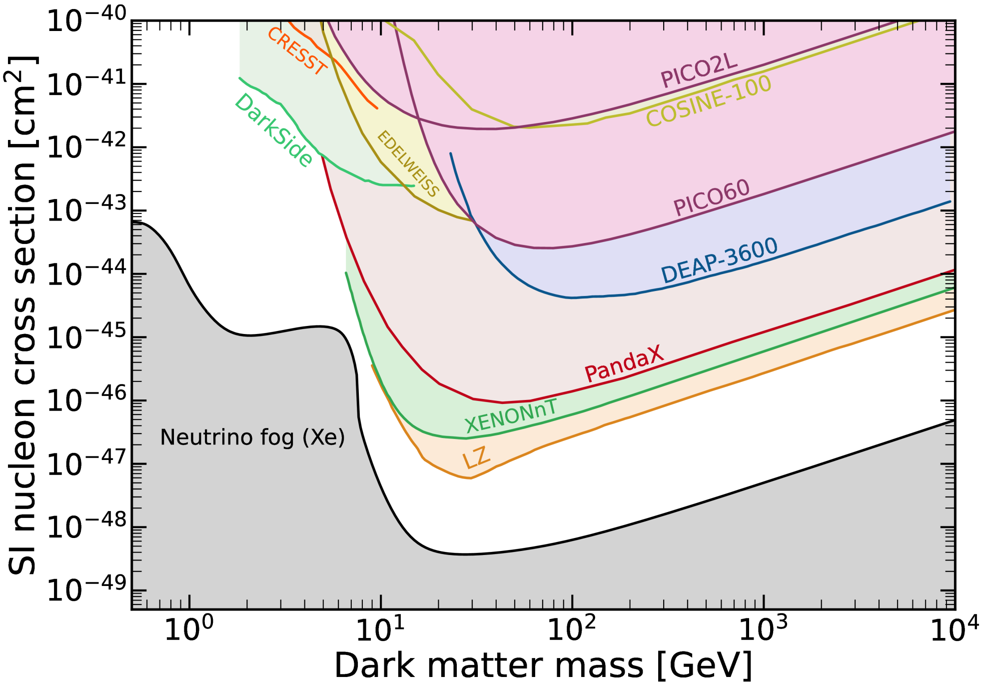


Figure 2.7: Limits placed on a spin-independent WIMP-nucleon cross section for a spectrum of dark matter masses. Leading limits are placed by results from the LZ collaboration [85] and DarkSide [96]. Other results are included for reference, as is the neutrino fog. Using data from [57].

Mid-2024, the most stringent limits on TeV-mass dark matter are set by LZ [85], with competitive results from XENONnT [97] and PandaX [98]. As illustrated in Figure 2.2, the Xe nucleus is well-positioned for the detection of dark matter in this mass range. Developments in recent years have led to significant reduction in radiogenic backgrounds, due to improvements in radioassays, improved purification of LXe, and the larger fiducial volume of modern TPCs.

At masses of around 10 GeV/c^2, best sensitivity has currently been achieved by the DarkSide collaboration [96]. In this regime, the rate of events in LAr above threshold exceeds that in LXe.

These experiments and their successors will continue to iterate on these sensitivities for the coming decades [99, 100]. Eventually, they will achieve sensitivity to neutrinos, which are for the most part indistinguishable from dark matter interactions in TPCs. This point, referred to as 'the neutrino fog', will be the limit of dual-phase TPCs in the direct detection of dark matter, but allow for future direct dark matter detectors to operate as neutrino observatories [101]. Current liquid noble TPCs are approaching sensitivity to ^{8}B coherent elastic neutrino-nucleon scattering (CEνNS)[102, 103], with more advanced observatories likely to achieve measurements of hep neutrinos at higher energies and lower cross sections.

2.3.3 GeV Mass

At the GeV scale, dual-phase TPCs begin to lose sensitivity to dark matter, due to the energy threshold imposed by the requirement of scintillation light from an interaction. However, an electroluminescence-only (S2-Only) analysis in these detectors removes the requirement of a scintillation signal entirely, allowing for increased sensitivity to low-mass dark matter [73]. The nature of this analysis in LZ will be discussed in detail in Chapter 6.

Figure 2.8 shows the leading dark matter sensitivities in the GeV regime, alongside the projected sensitivity of LZ using an S2-Only analysis.

Current leading limits are set by previous LXe and LAr experiments in S2-Only analysis. At lower masses, successors to past cryogenic experiments have maximised their sensitivity to low-mass candidates by decreasing their active mass, allowing for the lowering of detection thresholds. Current leading limits are established for cryogenic detectors by CRESST [109].

Sensitivity in this regime has also been achieved through the use of light gases such as He, Ne, or H as a detector medium by the NEWS-G collaboration. Current limits are set from preliminary science runs at Modane [104] and SNOLAB [110], with greater sensitivity upcoming through improved shielding and the proposed DarkSPHERE detector

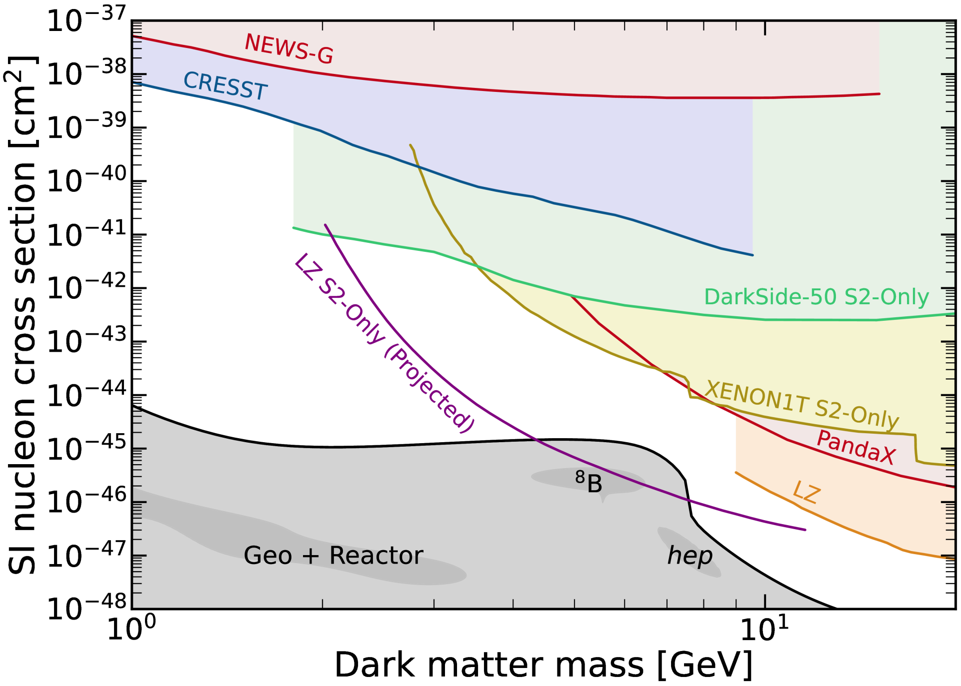


Figure 2.8: Projected WIMP sensitivity of LZ for an S2-Only analysis, from [73]. For reference, sensitivities from NEWS-G [104], CRESST [105], DarkSide-50 [106], XENON1T [107, 108], PandaX [98], LZ's S1-S2 analysis [85], and the neutrino fog [101] are overlain. Contours in the neutrino fog indicate the presence of ^{8}B, hep, and geothermal/reactor neutrinos.

[111]. An S2-Only analysis in LZ is projected to achieve sensitivity to WIMPs in this mass range beyond that currently established.

2.3.4 Sub-GeV Mass

At MeV-scale energies, standard S2-Only analyses begin to lose sensitivity to dark matter, as a result of the detection efficiency of electroluminescence. However, it may be possible for a nuclear recoil to induce a change in atomic energy levels in an atom, resulting in the secondary emission of an electron. This secondary electron recoil signal would occur in $\lesssim 10^{-5}$ of interactions, but would be above detection threshold in liquid noble TPCs. This effect, referred to as the Migdal effect, would improve the sensitivity of liquid noble TPCs to low-energy signals in the MeV-GeV range [112].

The Migdal effect has not yet been directly observed in nuclear recoils, although efforts to directly observe this effect are ongoing [113]. Current S2-Only Migdal searches assume the probabilities derived in Ref.[112].

Figure 2.9 shows the sensitivity of existing dark matter detectors to dark matter in the MeV-GeV regime, alongside the projected sensitivity of an LZ S2-Only Migdal search.

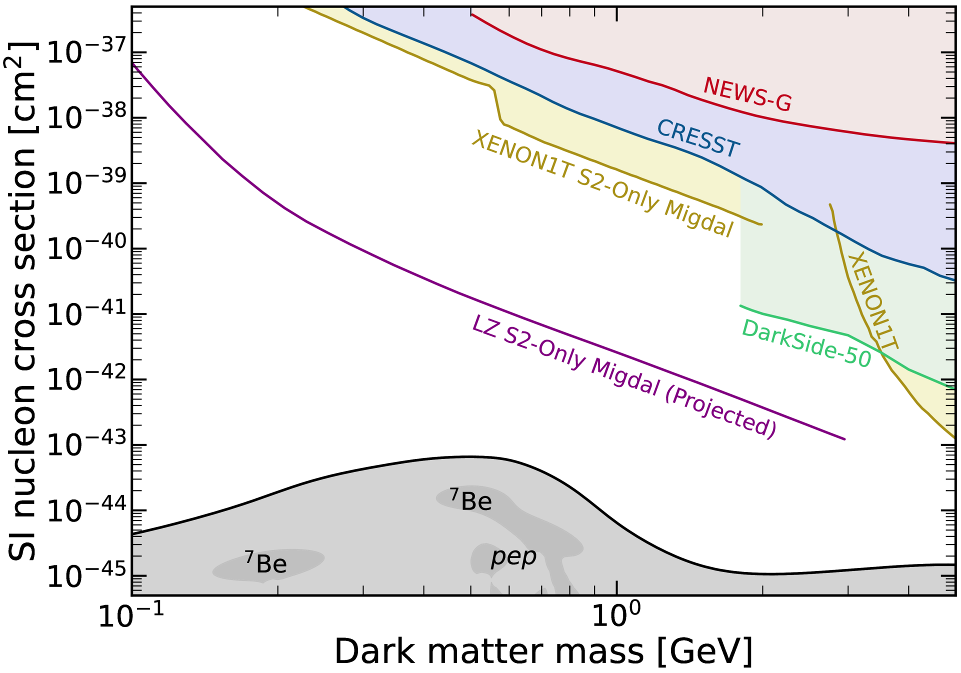


Figure 2.9: Projected WIMP sensitivity of LZ for an S2-Only Migdal analysis, from [73]. For reference, sensitivities from NEWS-G [104], CRESST [105], DarkSide-50 [106], XENON1T [107, 108], and the neutrino fog [101] are overlain. Contours in the neutrino fog indicate the presence of ^7Be and pep solar neutrinos.

The current leading sensitivity in this region is established by S2-Only Migdal analysis of XENON1T data [107]. Future data from NEWS-G is set to improve considerably on its current established sensitivity, as is data from the proposed cryogenic experiment SuperCDMS [114]. Early results from SENSEI have achieved sensitivity at these masses in Si through the use of Skipper CCDs, an alternative detection apparatus sensitive to energy depositions $\mathcal{O}(1\,\text{eV})$ [115]. These physics searches will likely produce results competitive with S2-Only Migdal analysis in LZ.

3

The LUX-ZEPLIN Experiment

LUX-ZEPLIN is a direct dark matter detection experiment, achieving sensitivity to a spin-independent WIMP-nucleon cross section of $9.2 \times 10^{-48}\,\mathrm{cm}^2$ for a WIMP of mass $36\,\mathrm{GeV/c}^2$ in its first science run [85]. In this chapter, the structure and operating principle of LZ are outlined, with an overview of the detector subsystems, a discussion of xenon as a detector medium, and a summary of the backgrounds, calibration sources, and analysis techniques. The sensitivity of LZ to dark matter candidates is then discussed, with an emphasis on how the work presented in this thesis has contributed to these results.

3.1 The LUX-ZEPLIN Time Projection Chamber

3.1.1 Overview

The LUX-ZEPLIN (LZ) experiment is a direct dark matter detector, 4850 ft below ground at the Sanford Underground Research Facility (SURF), in South Dakota. At the centre of LZ is a dual-phase LXe Time Projection Chamber (TPC). Surrounding this is a Xe Skin volume, which prevents electrical breakdown of components in the TPC and complications from thermal expansion, but also contributes to shielding from cavern gammas. This is housed in an inner cryostat vessel (ICV), which is itself suspended in an outer cryostat vessel (OCV), both of which support and protect the inner layers. These are surrounded by an outer detector (OD) made from gadolinium-doped liquid scintillator (GdLS), for the detection of neutrons. This is surrounded by the Davis Campus water tank, which is lined with 8" PMTs, in order to shield the TPC from both neutrons and cavern gammas. An overview of the detector can be seen in Figure 3.1.

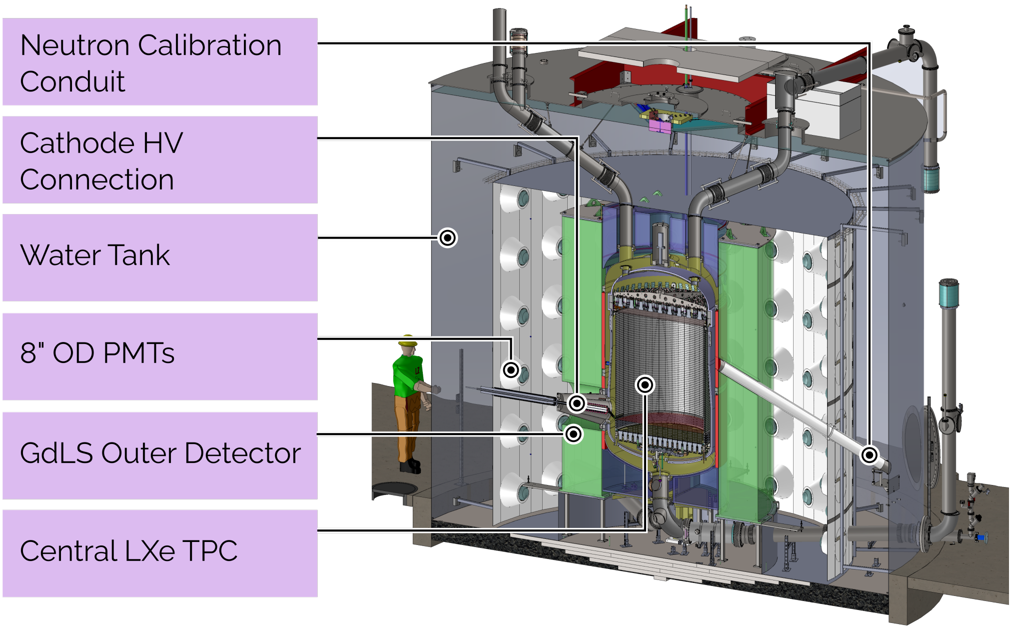


Figure 3.1: A rendering of the LZ experiment. Major detector subsystems are annotated. Rendering provided by the LZ collaboration.

Figure 3.2 shows a cross section of the LZ TPC, with each region and grid labelled. The LZ TPC contains a forward field 'drift' region of height and diameter 145.6 cm. A vertical electric field is imposed in this region by a cathode and gate grid. Above the gate is an ≈1.3 cm extraction region, containing an anode grid.

The field imposed between the cathode and gate draws electrons towards the extraction region, in which they produce an electroluminescence signal. The details of this signal, and how it can be used to reconstruct the position of an event, are outlined in Sections 3.1.2 and 3.1.3 respectively.

Above the anode wire, a GXe gap of 7 cm minimises field strength near to the top PMT array. For the bottom PMT array, this is achieved using an electrode grid at 0 V, referred to as the bottom grid. This produces a reverse field region (RFR) at the very bottom of the TPC.

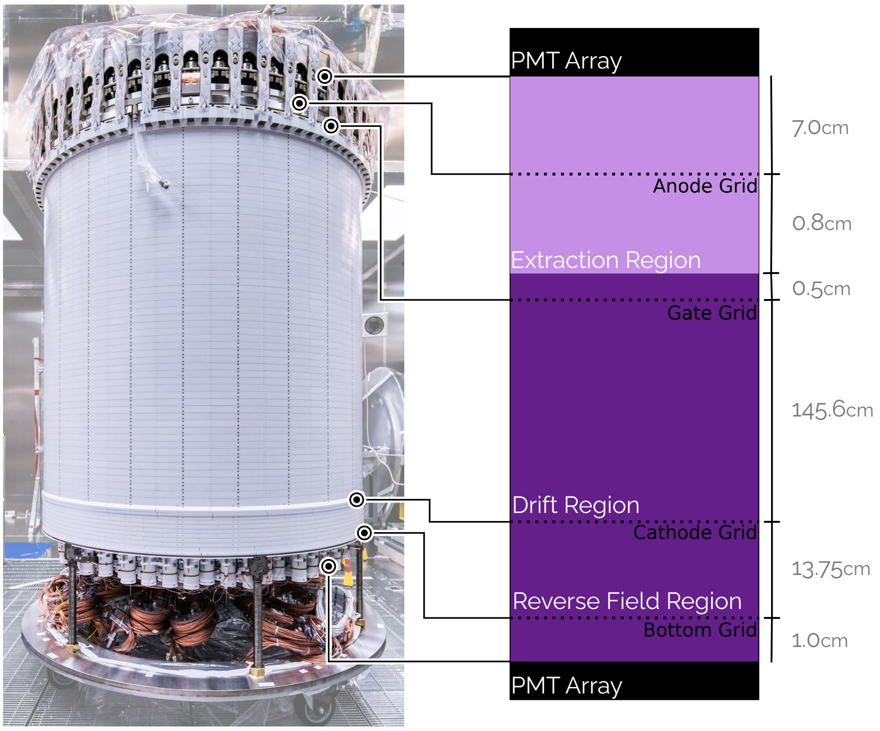


Figure 3.2: The LZ TPC, shown with a diagram of the relative positions of PMT arrays and field regions (not-to-scale). Photograph taken by Matthew Kapust, Sanford Underground Research Facility.

3.1.2 Interactions in LXe

Particles interact with a Xe atom through recoil with either its nucleus, in a nuclear recoil (NR), or outer electron shells, as an electron recoil (ER). This leads to either the ionisation or excitation of the atom, which leads to the release of heat, electrons, and vacuum ultraviolet (VUV) photons. This process is illustrated in Figure 3.3.

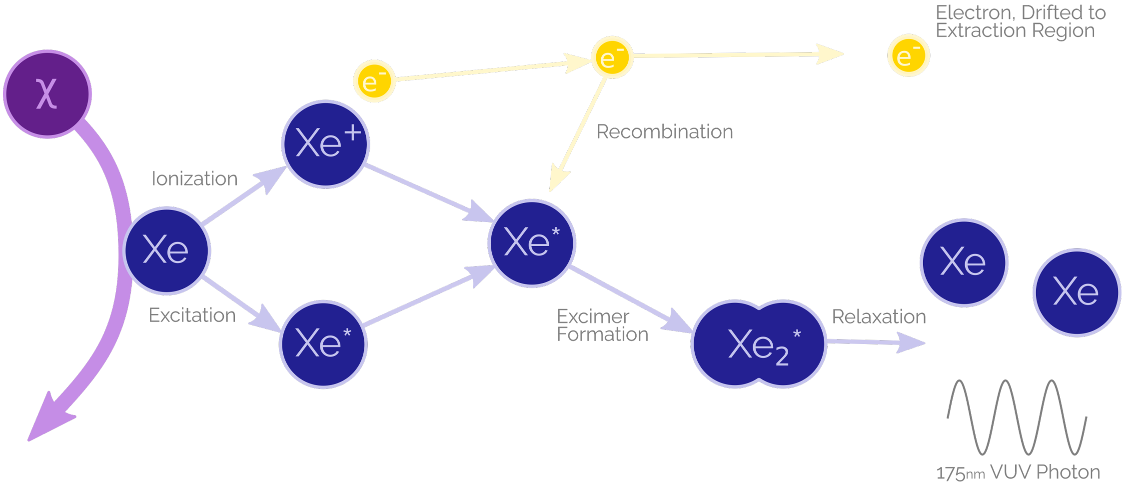


Figure 3.3: Illustration of the interaction of a particle with a Xe atom. Interactions produce VUV scintillation, as well as ionisation electrons. Steps taken from Ref. [116].

Primary Scintillation - The S1 Signal

Two of the processes illustrated in Figure 3.3 can lead to the emission of VUV photons. The first and simplest of these excites the Xe atom, resulting in the formation of strongly-bound Xenon dimers in an excited state (excimers). The relaxation of the excimer results in the production of VUV light (S1 light). This process follows the steps of:

$$\mathrm{X} + \mathrm{Xe} \longrightarrow \mathrm{Xe}^* + \mathrm{X} \qquad \text{Recoil and Excitation} \qquad (3.1)$$

$$\mathrm{Xe}^* + \mathrm{Xe} \longrightarrow \mathrm{Xe}_2^{*,v} \qquad \text{Excimer Formation} \qquad (3.2)$$

$$\mathrm{Xe}_2^{*,v} \longrightarrow \mathrm{Xe}_2^{*} + \mathrm{X} \qquad \text{Relaxation} \tag{3.3}$$

$$\mathrm{Xe}_2^{*} \longrightarrow \mathrm{Xe} + \mathrm{Xe} + \gamma \qquad \text{VUV Emission,} \tag{3.4}$$

where Xe^* indicates an excited state and $\mathrm{Xe}^{*,v}$ indicates that this excitation is in the form of both electron and vibrational excitation. Ionisation can also lead to the emission of primary scintillation light, through the process of recombination of ionised electrons with Xe_2^+ molecules. This follows the steps of:

$$\mathrm{X} + \mathrm{Xe} \longrightarrow \mathrm{Xe}^{+} + e^{-} + \mathrm{X} \qquad \text{Ionisation} \tag{3.5}$$

$$\mathrm{Xe}^{+} + \mathrm{Xe} + \mathrm{Xe} \longrightarrow \mathrm{Xe}_2^{+} + \mathrm{Xe} \qquad \text{Diatomic Formation} \tag{3.6}$$

$$e^{-} + \mathrm{Xe}_2^{+} \longrightarrow \mathrm{Xe}^{**} + \mathrm{Xe} \qquad \text{Excimer Formation} \tag{3.7}$$

$$\mathrm{Xe}^{**} + \mathrm{Xe} \longrightarrow \mathrm{Xe}^{*} + \mathrm{Xe} \qquad \text{Relaxation} \tag{3.8}$$

$$\mathrm{Xe}^{*} + \mathrm{Xe} \longrightarrow \mathrm{Xe}_2^{*,v} \qquad \text{Excimer Formation} \tag{3.9}$$

$$\mathrm{Xe}_2^{*,v} + \mathrm{Xe} \longrightarrow \mathrm{Xe}_2^{*} + \mathrm{Xe} \qquad \text{Relaxation} \tag{3.10}$$

$$\mathrm{Xe}_2^{*} \longrightarrow \mathrm{Xe} + \mathrm{Xe} + \gamma \qquad \text{VUV Emission.} \tag{3.11}$$

In both channels, VUV photons are emitted from the singlet and triplet state of the Xe dimer. As recombination is the rate-determining step in this process, the S1 pulse has a characteristic shape of a sharp increase in charge, followed by an exponential decay. This decay contains three time constants, as a result of the singlet lifetime, triplet lifetime, and recombination time.

Secondary Electroluminescence - The S2 Signal

Free electrons that do not recombine are drifted towards the gas phase of the detector through an applied drift field, and are subsequently extracted into the gas phase, where they are accelerated through a higher-voltage extraction field. This acceleration produces electroluminescence light (S2 light).

Light from a single electron during electroluminescence is measured by a PMT uniformly in time, creating a uniform pulse of duration $\mathcal{O}(1\,\mu\text{s})$. However, the electron cloud producing an S2 diffuses while travelling through the detector. Most S2 pulses therefore have a characteristic Gaussian-like shape. Some electrons are delayed during diffusion, due to field fluctuations at the TPC walls, thermalisation of electrons at the liquid surface, or the absorption and re-emission of electrons by oxygen impurities in LXe [117]. This produces a small 'tail' of charge on the end of an S2, producing a slightly asymmetric Gaussian waveform.

3.1.3 Position and Energy Reconstruction

The S1 and S2 pulse from a single scatter can be used to reconstruct the position and energy of an interaction. An example event in LZ can be seen in Figure 3.4.

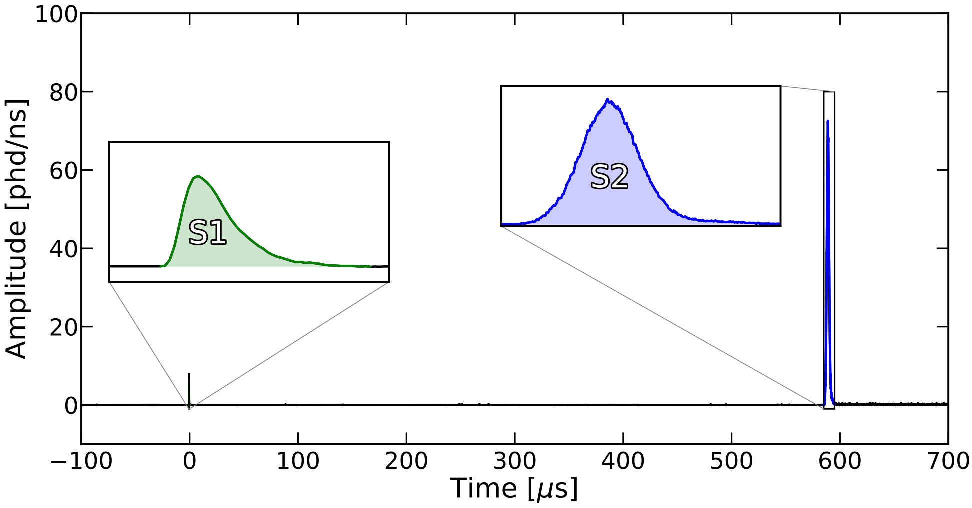


Figure 3.4: An example single scatter in LZ data, with the S1 and S2 pulse highlighted in green and blue respectively.

The (x, y) position of an event can be calculated from the charge collected by each PMT in the top PMT array [118]. This can be seen in Figure 3.5, which shows the S2 pulse area recorded by each PMT on the top array, as well as the reconstructed (x, y) position. The time for electrons to travel to the extraction region - the drift time of the event - can be calculated from the timing difference between an S1 and an S2 pulse. As an event's drift time is proportional to its vertical position in the detector, this can be used to reconstruct the (z) coordinate of an event.

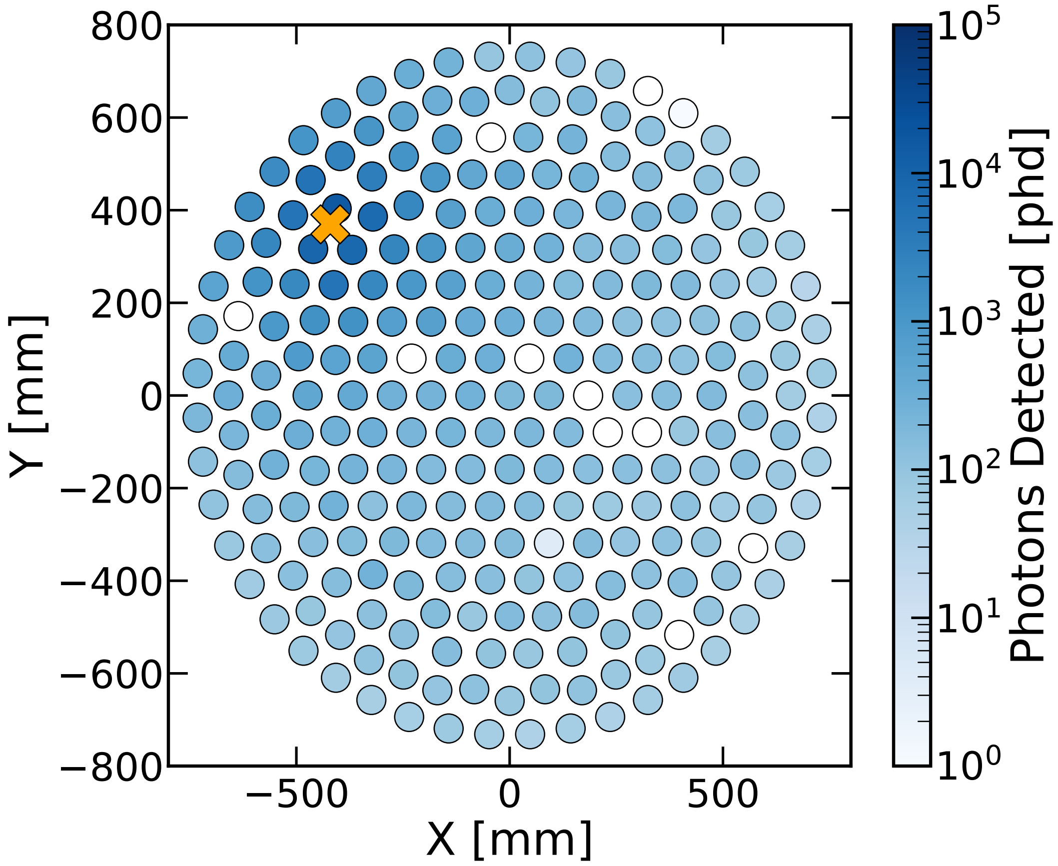


Figure 3.5: PMT hitmap for the event displayed in Figure 3.4, for the top PMT array. A cross is annotated to show the reconstructed (x, y) position of the event.

As an electron recoil occurs further from the Xe atom than a nuclear recoil, the probability of recombination is lower. This results in ER events exhibiting a larger S2 pulse area and smaller S1 pulse area than NR events of the same energy. Figure 3.6 shows the yield of electrons and photons from an ER and NR event for varying energies, generated using NEST [119]. The light yield from NR events is considerably higher at most energies than charge yield, resulting in events with a smaller S2 and larger S1. This relationship can also be seen in LZ calibration data, in Figure 3.8.

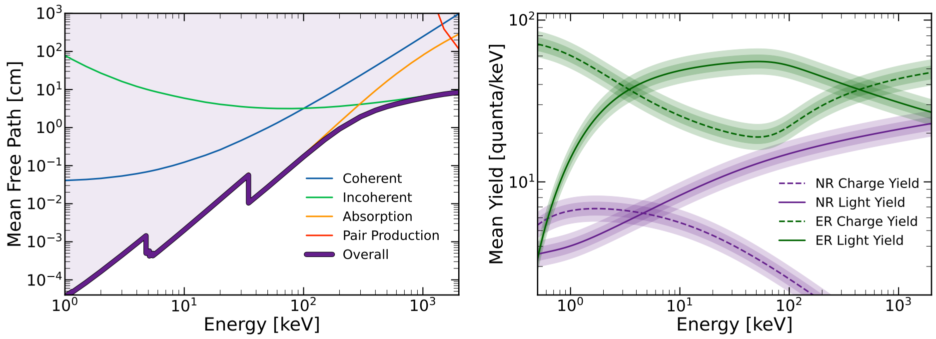


Figure 3.6: Left: Approximate mean free path of photons in LZ, calculated using XCOM [120]. Contributing factors (elastic collisions, inelastic collisions, absorption, and pair production) are highlighted. The probability of scintillation light reaching the TPC PMTs is high. Right: Approximate scintillation and ionisation yields of electron and nuclear recoils in LZ, generated using NEST [121]. Bands indicate 1σ and 3σ confidence. The ratio of charge to light in interactions can be used to discriminate between ER and NR events. Data assumes a nominal drift field of 100 V/cm.

Some scintillation and ionisation from an interaction is not detected. For example, electrons are absorbed by oxygen impurities in the TPC, resulting in up to 20% of electrons in an interaction being lost during diffusion. The probability of photon loss from travel through the TPC is minimal, as illustrated in Figure 3.6. The majority of light loss is instead a result of absorption at the liquid surface, by polytetrafluoroethane (PTFE) surrounding PMTs, or is a result of a PMT's Quantum Efficiency (QE).

The measured light and charge loss from calibration data can be used to correct an event's S1 and S2 pulse area, to account for these losses. This results in the parameters $\mathrm{S1}_c$ and $\mathrm{S2}_c$, the corrected S1 and S2 pulse area of an event.

The energy of a recoil can be reconstructed from the number of electrons n_e and photons n_γ observed. For electron recoils, this is given by

$$E = W(n_e + n_\gamma), \tag{3.12}$$

where W is the average energy required for photon emission [121], measured to be $11.5^{+0.2}_{-0.3}$eV in LXe [122].

Measured values of S1$_c$ and S2$_c$ can be converted to energy E using the parameters g_1 and g_2:

$$E = \frac{W}{\mathcal{L}}(\frac{S1_c}{g_1} + \frac{S2_c}{g_2}). \tag{3.13}$$

g_1 and g_2 describe the average energy of an interaction in LZ per detected electron and photon, which can be determined using calibration data [123].

$\mathcal{L}$ denotes the Lindhard factor. For ERs, this parameter is near-unity, as the energy lost to atomic motion in an interaction is negligible. For NRs, however, this factor is energy-dependent, estimated to be 0.17 at an energy of 10 keV [119]. A common convention when displaying reconstructed energies from LXe WIMP searches is to plot energy with the notation keV$_{ee}$, or keV-electron-equivalent, as well as keV$_{nr}$, which takes into account the Lindhard factor and assumes all data are nuclear recoils.

3.1.4 TPC PMT Arrays

The PMTs used in the top and bottom array of the TPC were specifically designed with Hamamatsu for high radiopurity, high QE in the VUV photon range at low temperatures, and low rates of spurious light emission (dark counts). For a signal to be generated in a PMT, an incident photon must be transmitted through a PMT window and strike the photocathode, producing free electrons through the photoelectric effect. These are then accelerated onto the first dynode, where they are multiplied through the process of secondary electron emission. This is repeated through successive dynodes, until electrons strike the anode and are collected by an amplifier [124].

The majority of S1 light produced in the TPC is collected by the bottom PMT array, as light reflects off of the liquid surface. To maximise light collection, the bottom PMT array is ordered in a tightly-packed hexagonal structure. The majority of S2 light is collected by the top PMT array, due to the close proximity of the extraction region. As S2 light is essential to the reconstruction of the (x, y) position of an event, the top array is constructed in a hybrid layout, where the outermost two rings of PMTs are arranged non-hexagonally, in order to maximise reconstruction efficiency [125]. This can be seen in Figure 3.5

3.2 Calibrations

To measure the values of g_1, g_2, and W, as well as to measure the properties of NR and ER events, LZ requires a rigorous in-situ calibrations programme. Table 3.1 presents an overview of the calibration sources used in LZ.

The majority of calibration sources for LZ are external to the detector - these are deployed around the wall of the TPC through the calibration source deployment (CSD) system, with the exception of deuterium-deuterium neutrons, which are injected into the TPC through a separate D-D generator, seen in Figure 3.1. However, the self-shielding properties of LXe prevent calibration of the central regions of the TPC using external sources, prompting the use of internal calibration sources as part of this programme. Such sources are injected into the LXe as it flows into the TPC, and are selected such that they decay away or can be filtered out of the LXe by a heated zirconium getter prior to collecting data for a WIMP search [125].

3.2.1 Tritium

Tritium (^{3}H) is the primary source used for modelling the response of the TPC to ERs. β decay of tritium has a half life of 12.3 years, with a continuous energy spectrum up to 18.6 keV. Tritiated methane (CH_3T) is injected into the TPC for

Table 3.1: An overview of calibration sources used in LZ. The type of source, decay channel, energy range, half life, and purpose of each source are listed [125].

Deployment	Isotope	Type	Energy [keV]	$\tau_{1/2}$	Purpose
Internal	^{3}H	β	18.6 endpoint	12.5 y	ER Band
	^{14}C	β	156 endpoint	5730 y	ER Band
	^{83}Kr	γ	9.4, 32.1	1.83 h	TPC (x, y, z)
	^{131m}Xe	γ	164	11.8 d	TPC (x, y, z)
	^{220}Rn	(α, β, γ)	various	10.6 h	Xe Skin
External Deployed	^{57}Co	γ	122	0.74 y	Xe Skin
	^{22}Na	γ	511, 1275	2.61 y	TPC & OD Sync
	^{133}Ba	γ	356	10.5 y	ER Response
	^{54}Mn	γ	835	312 d	ER Response
	^{60}Co	γ	1173, 1333	5.27 y	ER Response
	^{228}Th	γ	2615	1.91 y	ER Response
	^{124}AmLi	(α, n)	1500 endpoint	432 y	NR Band
	^{124}AmBe	(α, n)	11,000 endpoint	432 y	NR Band
	^{252}Cf	n	Watt Spectrum	2.65 y	NR Response
External Gamma	^{88}YBe	(γ, n)	152	107 d	NR Response
	^{124}SbBe	(γ, n)	22.5	60.2 d	NR Response
	^{205}BiBe	(γ, n)	88.5	15.3 d	NR Response
	^{206}BiBe	(γ, n)	47	6.24 d	NR Yields
DD		n	272-400	-	NR Yields
			2450	-	NR Yields

tritium calibrations. The size of the methane molecule ensures that isotopes can be removed from the TPC through the getter. Like other injected sources, CH_3T diffuses throughout the detector, and is therefore valuable when evaluating the positional dependence of TPC physics.

3.2.2 AmLi

^{241}Am produces α radiation, which can trigger (α, n) reactions with Li. This produces neutrons of continuous energy up to an endpoint of 1.5 MeV, resulting in AmLi producing a spectrum of NR events across all energies. AmLi calibration sources are deployed through the CSDs. Data is collected with the sources near to the top, middle, and bottom of the TPC, in order to ensure AmLi events take place over a range of (z) positions.

3.2.3 Deuterium-deuterium Generator

The detector response to low-energy NRs is calibrated through an external deuterium - deuterium (DD) generator. This accelerates deuterium ions into a fixed deuterium target, resulting in the process:

$$D + D \longrightarrow {}^{3}\mathrm{He} + n. \tag{3.14}$$

A conduit penetrating the water tank allows these produced neutrons at energy 2.45 MeV to enter the TPC.

In order to ensure interactions are single scatters, the DD conduit is aimed towards the top of the TPC, such that neutrons escape through the gas phase after an average of one scatter. This leads to a distribution of DD events that is closely localised in (x, y, z) relative to other calibration sources.

3.2.4 Krypton

^{83m}Kr is injected into the TPC in order to introduce positional corrections to data. This source has a half life of 1.83 hours, with an intermediate state of half life of 154 ns. Both decays produce mono energetic ERs, with the separation between decays being useful to tag events for use in calibration.

3.2.5 Radon

^{220}Rn is used as a calibration source for the skin region of the TPC. This is not to be mistaken for ^{222}Rn, a dominant background in LZ with many radioactive daughters. ^{220}Rn decays through alpha decay at 6.3 MeV, with a second alpha produced $\mathcal{O}(150\,\mathrm{ms})$ after at 6.8 MeV, allowing events to be easily tagged.

3.3 Backgrounds

In order to evaluate the sensitivity of LZ to WIMPs, the backgrounds present must be modelled. The known physics effects which contaminate the region of interest for a WIMP search can be sorted into three distinct categories. The first of these, accidental coincidence events, are rejected using analysis cuts discussed in detail in Chapter 4. This leaves two sources that can be predicted in advance of operation: leakage of ER events into the NR band, and true NRs, from neutrons, internal decays, or Coherent Elastic ν-Nucleon Scattering (CEνNS).

3.3.1 Radiogenic Backgrounds

Radionuclides of noble elements dominate the single scatter data collected in LZ. These contribute to the ER band in data, and can leak into the NR band, as illustrated in Figure 2.3.

The largest obstacle to WIMP sensitivity is ^{222}Rn. This originates from the decay chain of ^{238}U in detector materials [126]. ^{222}Rn is long-lived, and as such can emanate into the bulk xenon before decaying. While the decay of ^{222}Rn itself is of a high enough energy to be rejected from a WIMP search, its progeny can be more problematic, in particular the β decay of ^{214}Pb and ^{212}Pb. The decay chain of ^{222}Rn can be seen in Figure 3.7.

3.3.2 Cosmogenic Backgrounds

Most cosmogenic backgrounds in LZ come from the activation of ^{127}Xe and ^{37}Ar. The former produces X-Rays and Auger electrons at 1.2 keV, 5 keV, and 33 keV. Many of these events can be tagged through secondary de-excitation gammas from ^{127}I, although the probability of gammas from this decay escaping the TPC is high. The decay of ^{37}Ar produces Auger electrons and X-Rays of a total energy of 2.82 keV, directly within the region of interest for a WIMP search.

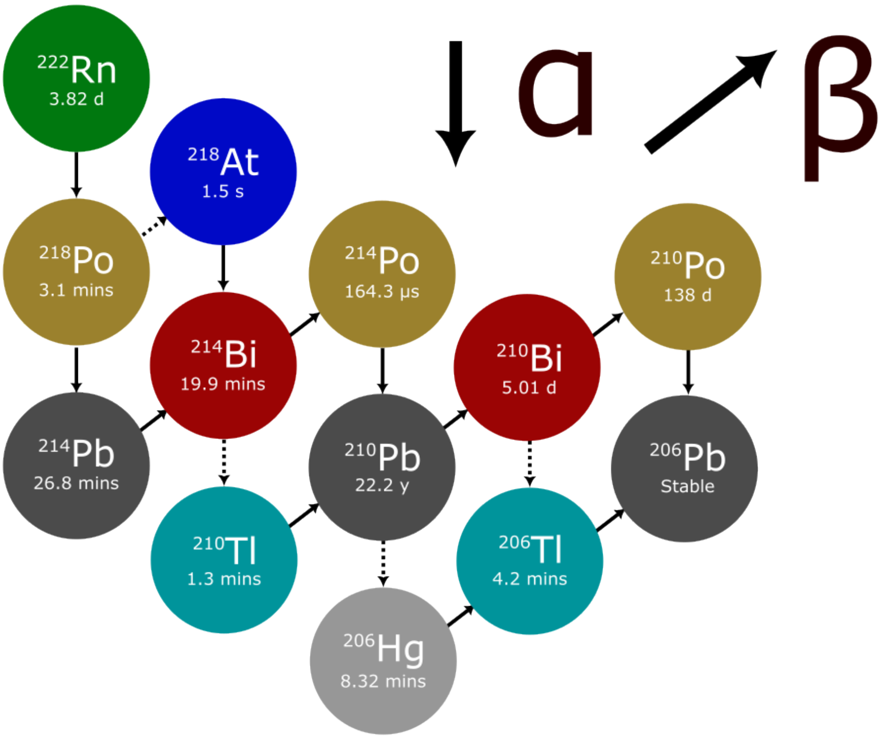


Figure 3.7: Decay chain of ^{222}Rn. Dotted lines indicate secondary decay channels. The decay of ^{222}Rn produces several sources of β emission which contribute to ER leakage in LZ. Data taken from [127].

Cosmogenic muons can interact with rock in the Davis Cavern, producing neutrons which produce a TPC background. Most neutrons are thermalized in LZ's outer water tank, with the remaining events tagged by the OD [128].

3.3.3 Physics Backgrounds

Three known physics backgrounds generate single-scatter events with no corresponding veto signal. ν-e scattering produces electron recoils, which contribute to low-energy ER leakage. $2\nu\beta\beta$ ^{136}Xe decay produces NR events, at a rate measured by EXO-200 [129] and KamLAND-Zen [130]. CEνNS has been measured for the first time by the COHERENT collaboration[131]. These measurements indicate that leakage from radiogenic backgrounds remains the dominant background for high-mass WIMP searches in LZ.

3.4 First Science Run Data from LZ

The results of LZ's first science run were published on 8th July 2022 on arχiv [85]. 60 days of live data were collected with an active mass of 5.82 t. The most stringent limit is set for spin-independent scattering at 36 GeV/c^2, rejecting cross-sections above $\sigma_{SI} = 9.2 \times 10^{-48}\,\mathrm{cm}^2$ at the 90% confidence level.

3.4.1 Detector Conditions

Data were collected from the 23rd December 2021 to the 11th May 2022, totalling 89 live days after removing periods of maintenance, calibration, dead time, and anomalous trigger periods. The detector was operated at a drift field of 193 V/cm, established by the cathode and gate electrodes, and an extraction field of 7.3 kV/cm. LXe temperature and pressure were maintained at 174.1 K and 1.791 bar(a). Operating conditions were stable to within 0.2% [85].

A detector response model was developed using the calibration data detailed in Table 3.1. The S1 and S2 pulse areas are corrected using this model to give a value of S1$_c$ and S2$_c$ for each event, given in phd (photons detected). S1$_c$ and S2$_c$ are used to reconstruct the energy of an event, with $g_1 = (0.115 \pm 0.002)$ phd/photon and $g_2 = (46.4 \pm 1.5)$ phd/electron [85].

The ER and NR bands are defined using CH$_3$T and DD calibration data. The distribution of this data in S1$_c$-S2$_c$ space can be seen in Figure 3.8.

3.4.2 Analysis Cuts

After application of a series of cuts to live time, removing periods of low data quality, a suite of event-level cuts removes populations that are not WIMP-like, or are difficult to precisely model. These are:

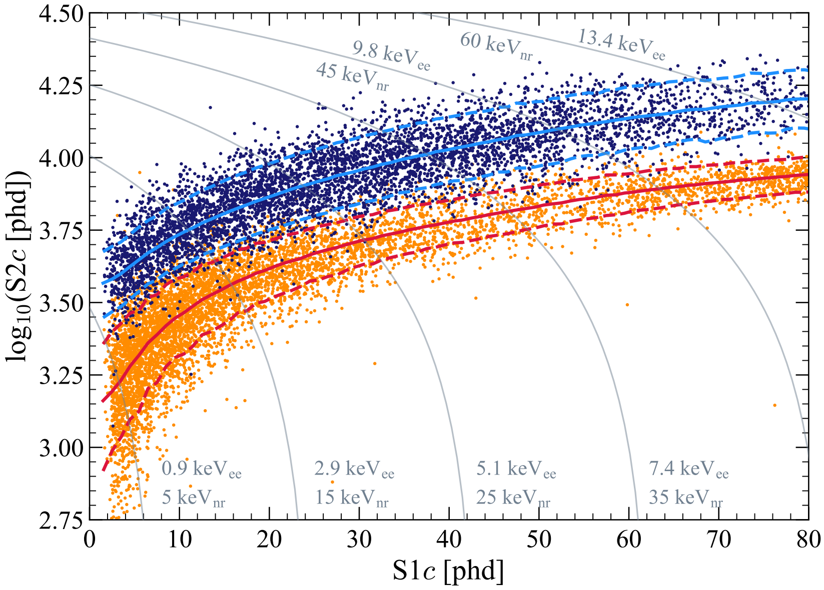


Figure 3.8: Calibration events in $\log_{10}$S2$_c$-S1$_c$ space from CH$_3$T calibration (blue points) and D-D neutrons (orange points). Blue and red lines indicate the median of simulated ER and NR events respectively, with dotted lines indicating the 3σ boundaries of these bands. Grey lines illustrate contours of constant keV$_{ee}$ and keV$_{nr}$. Taken from Ref.[85].

Single Scatter Cut: Only events of exactly one prominent S1 and S2 are selected, as the probability of a WIMP multiply scattering is negligible.

Fiducial Volume: The self-shielding properties of LXe result in a lower rate of background events in the TPC centre. Events at the TPC boundaries are also of a lower quality. A fiducial cut therefore removes events near to the TPC walls, as well as those near to the liquid surface and cathode.

WIMP ROI: Single scatter events above or below a given size of S1$_c$ and S2$_c$ are ignored. A minimum S1$_c$ and S2$_c$ of 3 phd and 600 phd respectively allow a large fraction of accidental coincidence events to be removed from analysis with minimal signal loss. The ROI then extends to S1$_c$ and S2$_c$ of 80 phd and 1×10^5 phd.

Vetos: Events with signal in either the OD or skin that occur in prompt coincidence with a TPC S1 are removed, as they are likely a product of gammas or neutrons from the Davis Cavern.

Accidental Cuts: Data quality cuts are applied to the WIMP search to reduce the number of accidental coincidence events. These cuts are discussed in Chapter 4.

3.4.3 Results

After all cuts, 336 events remain in the SR1 WIMP ROI. Around 300 of these events are within the ER band, seen in seen in Figure 3.9. Around 30 events within and beneath the grey contour are predicted to originate from NR events of lower $S2_c$, annotated as a red band, with around 30 nearer to the NR band of lower $S2_c$. Events in the ER band follow a uniform spectrum of energies up to 20 keV$_{ee}$, excluding a mono energetic peak from ^{37}Ar centred around 2.8 keV$_{ee}$.

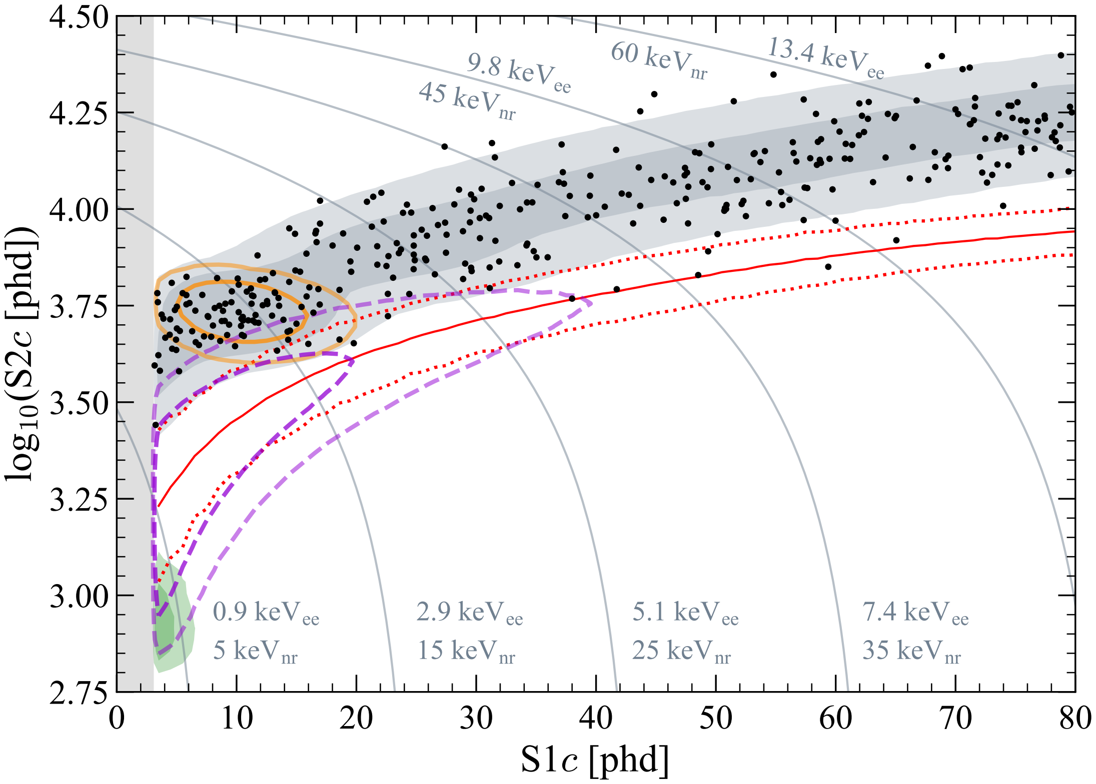


Figure 3.9: SR1 data in $S1_c$-$S2_c$ space. Contours illustrate the modelled distribution of ^{37}Ar (orange), ^{8}B (green), and the ER band (grey), as well as a 30 GeV/c^2 WIMP signal (purple). The NR band modelled in Figure 3.8 is included in red [85].

The number of events in each population can be measured as part of the statistical tests outlined in Section 2.2.2. The number of events within each population in data is within errors of those measured in simulations. Of particular note is the measurement of accidental events, identifying just (1.2 ± 0.3) events in data. 95% of accidental events are rejected by the work presented in Chapter 4, contributing to this low rate.

Across all masses tested, the observed data is consistent with a background-only hypothesis, with a measurement of $0.0^{+0.6}$ events for a $30\,\text{GeV/c}^2$ WIMP. A table of expected and observed events for each background population can be seen in Table 3.2.

Table 3.2: Number of events from various sources in the 60 d, 5.5 t exposure. The middle column shows the measured number of events in simulations, with uncertainties as described in the text. The uncertainties are used as constraint terms in a combined fit of the background model plus a $30\,\text{GeV/c}^2$ WIMP signal to the selected data, the result of which is shown in the right column. ^{37}Ar and detector neutrons have non-Gaussian prior constraints and are totalled separately. Values at zero have no lower uncertainty due to the physical boundary. Taken from [85]

Source	Expected Events	Fit Result
β decays + Det. ER	215 ± 36	222 ± 16
ν ER	27.1 ± 1.6	27.2 ± 1.6
^{127}Xe	9.2 ± 0.8	9.3 ± 0.8
^{124}Xe	5.0 ± 1.4	5.2 ± 1.4
^{136}Xe	15.1 ± 2.4	15.2 ± 2.4
^{8}B CE νNS	0.14 ± 0.01	0.15 ± 0.01
Accidentals	1.2 ± 0.3	1.2 ± 0.3
Subtotal	273 ± 36	280 ± 16
^{37}Ar	$[0, 288]$	$52.5^{+9.6}_{-8.9}$
Detector neutrons	$0.0^{+0.2}$	$0.0^{+0.2}$
30GeV/c^{2}WIMP	-	$0.0^{+0.6}$
Total	-	333 ± 17

The $S1_c$ and $S2_c$ of events passing data quality cuts, shown in Figure 3.9, was inserted into the PLR test as outlined in Section 2.2.2. The measured rate of each source of background events was accounted for as a series of nuisance parameters. A key nuisance parameter yet to be incorporated in LZ is the measured signal acceptance, which is affected by the mis-classification of small S2 pulses by pulse classification software. Chapter 5 measures the probability of mis-classification, and its dependence on pulse area and TPC position, allowing for improved WIMP sensitivity with a reduced error and minimal bias.

The distribution of WIMP sensitivity from this PLR analysis can be seen in Figure 3.10. Limits are applied using both power constraints and the CL_S method, detailed in Section 2.2.2. Sensitivity decreases linearly with increasing WIMP mass, as the number of candidates and the mass of each candidate are inversely correlated. As highlighted in Equation 2.2, sensitivity decreases exponentially at low WIMP masses, as a result of detection threshold. The dominant limiting factor for this low-mass sensitivity is the S1 detection efficiency. An analysis method to overcome this limitation, by forgoing the requirement of an S1, is presented in Chapter 6.

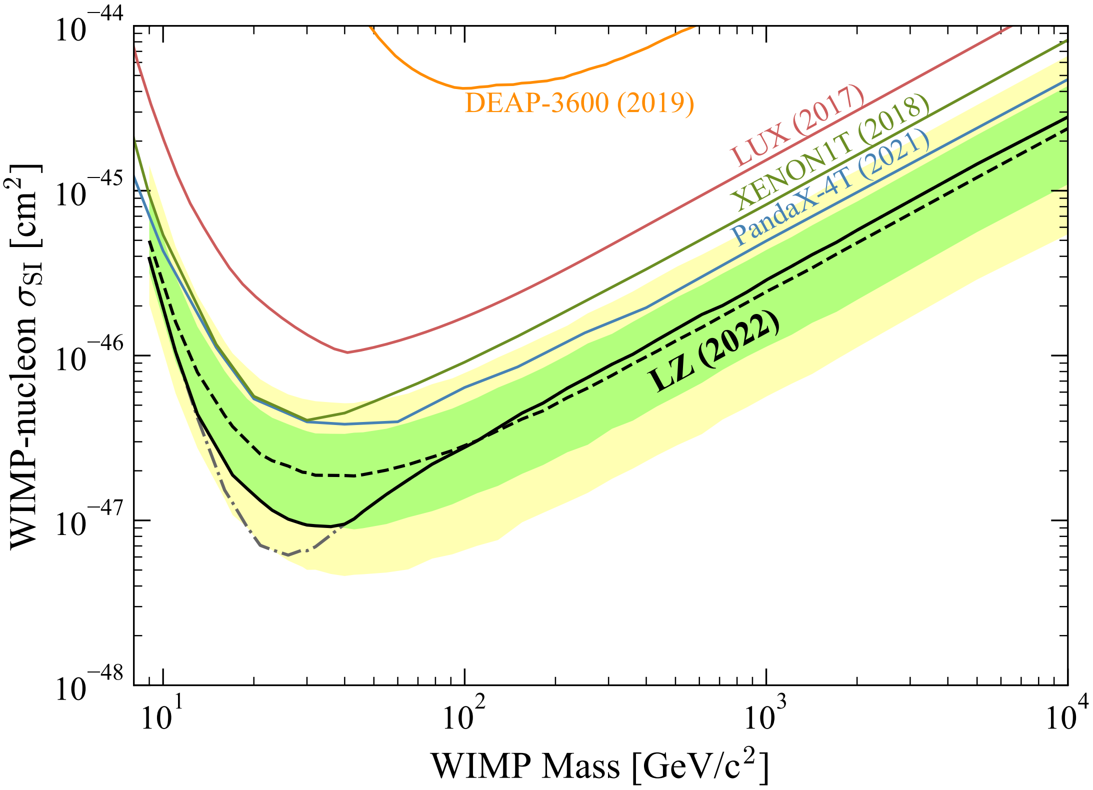


Figure 3.10: The limit obtained on the WIMP-nucleon spin independent cross section as a function of mass from LZ SR1 data, with power constraint (solid black) and CL_S (dash-dotted black) applied. The dashed line shows the expected SR1 sensitivity, while the green and yellow bands show the 1σ and 2σ uncertainties on this expectation. Competitive limits from other direct detection experiments at the time of the SR1 result publication are included and annotated [85].

4

Rejection of Accidental Coincidence Events in the LUX-ZEPLIN Experiment

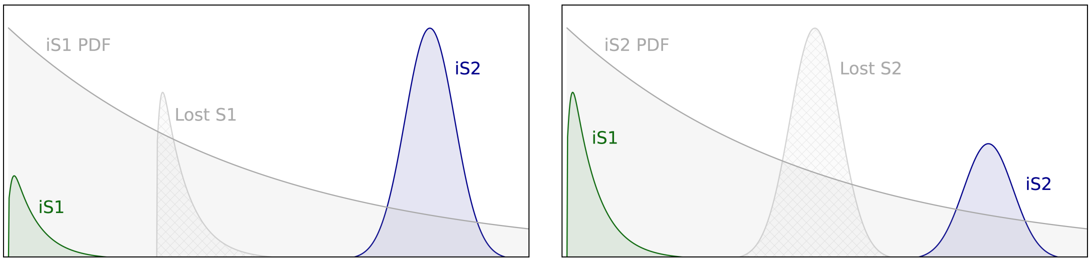
This chapter presents the development of two key analysis cuts in LZ, designed for the rejection of what are referred to as "Accidental Coincidence events". These cuts remove 95% of accidental events from analysis at the expense of less than 5% of data, following preliminary analysis cuts. The cuts are also demonstrated to remain effective outside of the WIMP ROI, and as such can be used in alternate high-energy and low-energy physics searches in LZ.

4.1 Accidental Coincidence Backgrounds

4.1.1 Overview

In order for an event to be considered in standard LZ analysis, it must contain both an S1 and S2 pulse. This mitigates the number of background events in data, many of which produce exclusively an isolated S1 (iS1) or S2 (iS2) pulse.

However, Figures 4.1(a) and 4.1(b) illustrate how isolated pulses can still produce a background in S1-S2 analysis. It is possible for an iS2 to be present within a time window corresponding to one drift length of an iS1, producing an event indistinguishable from a ’true’ single scatter (SS) event. These events, referred to as accidental coincidence events or accidentals, are an irreducible low-energy background in dual-phase TPCs.



(a) (b)

Figure 4.1: Illustrations of the formation of accidental events, relative to (a) the S2 and (b) the S1. Isolated pulses, as well as an illustration of the Poisson probability of observing an isolated pulse relative to $t = 0$, are shown.

4.1.2 Isolated S2 Pulses

An iS2 is produced in LZ when electrons are observed without the accompanying detection of an S1. In order to illustrate their physical origin, iS2 events from the first LZ science run have been passed through a Gaussian Mixture Model (GMM) and sorted into 25 clusters, taking in as inputs pulse area, (x, y) position and associated χ^2, top-bottom asymmetry (TBA), rms width, full width at half maximum, and the times at which the S2 waveform reached 1, 5, 25, 75, and 95% of its total integrated pulse area. The origin of each cluster is then identified by eye, and clusters with similar origins are combined. Details of a GMM are discussed in Appendix A. iS2 events are known to be produced in eight different regions:

Glue Ring: Events from the gate and anode glue ring. S1 light does not reach the PMT arrays.

Hotspot: Field emission from grid wires. No photons are produced.

E-Trains: Delayed electron emission from electron trains.

Liquid Surface: S1-S2 events near to or within the extraction region. The S1 and S2 cannot be reconciled in analysis.

Gate: Radiogenic events from the gate array. High fields result in no detected photons.

Cathode: Radiogenic events from the cathode array. High fields result in no detected photons.

Field Cage: Radiogenic events from the TPC field cage resistors. High fields result in no detected photons.

Other: The origin of these events could not be ascertained. Many of these will be ordinary S1-S2 events, where the S1 is lost.

Events from the field cage resistors and cathode are difficult to discriminate, as both produce very wide S2 pulses: as a result, the two are categorised together in this analysis. Similarly, hotspots and e-trains are difficult to distinguish, due to their small pulse areas, and will also be combined. The distribution of iS2 pulse area for each population can be seen in Figure 4.2.

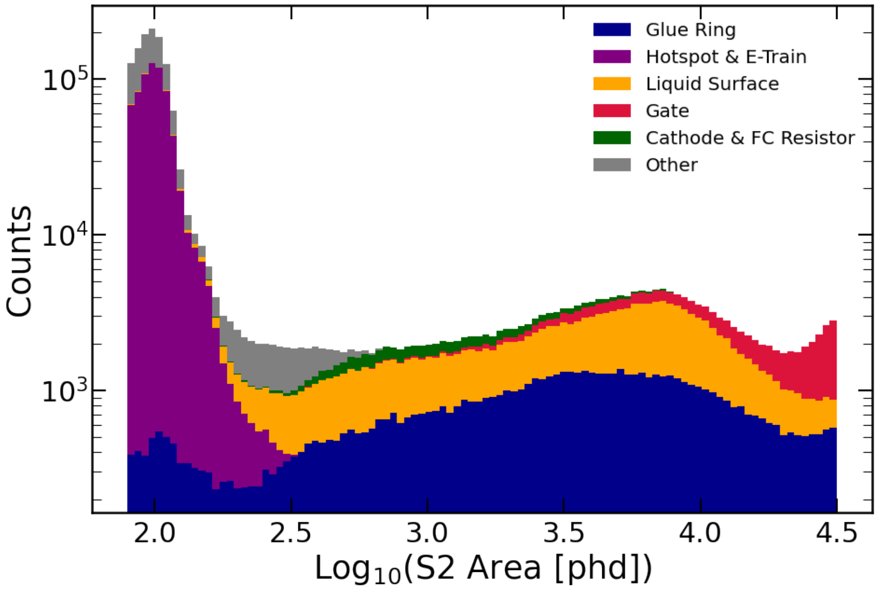


Figure 4.2: The pulse area of iS2 pulses in LZ, clustered using a simple GMM. Clusters are identified as originating from the glue ring (blue), field emission hotspots and E-trains (purple), near to the liquid surface (orange), the gate (red), and the cathode and field cage resistors (green). iS2s without a certain origin are included in grey.

Glue Ring

The TPC gate and anode wires are each captured between two support rings, which are held together by epoxy. This epoxy is a source of radiogenic backgrounds in LZ [125, 132]. The location of these rings, referred to as 'glue rings', is shown in a cross-section of the TPC in Figure 4.3.

Photons from interactions in this region are unlikely to reach PMTs, making this region a key source of iS2s. However, the field strength is also distorted in this region, such that electrons travel over a very large (x, y) distance during electroluminescence. This spread of S2 light results in a higher χ^2 on reconstructed (x, y) position of an S2, which can be used to discriminate glue ring events.

The χ^2 of a selection of events from each cluster can be seen plotted against S2 pulse area in Figure 4.4. At larger S2 pulse areas, using χ^2, events from the glue ring can be removed from data with minimal signal loss.

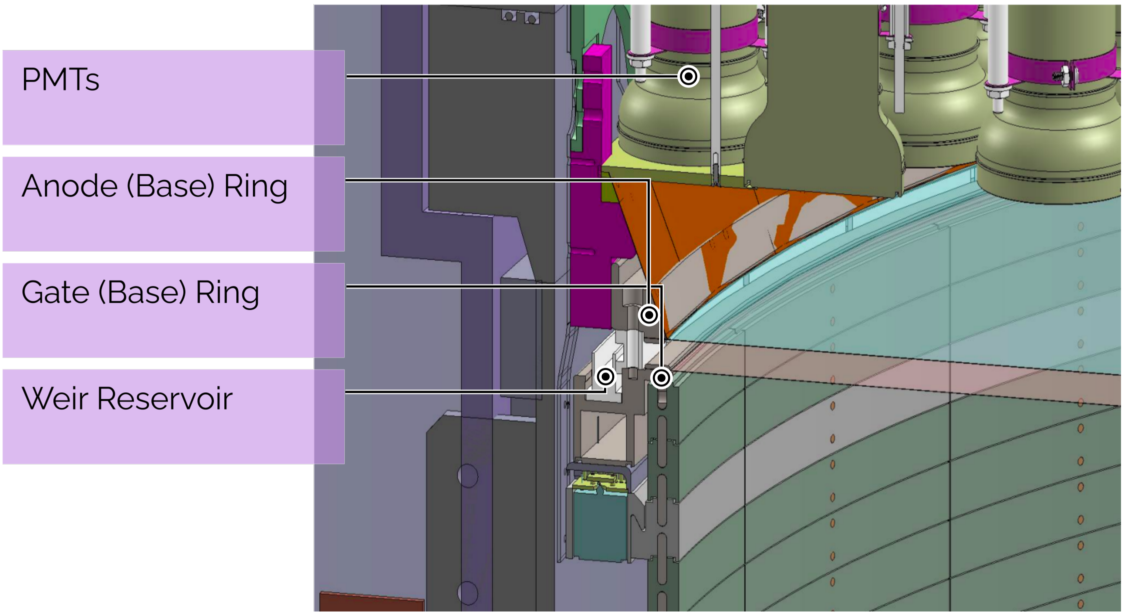


Figure 4.3: An enhanced schematic of the extraction region in LZ. The anode and gate rings, the sources of 'glue ring' events, are annotated, as are the weir reservoir and PMTs.

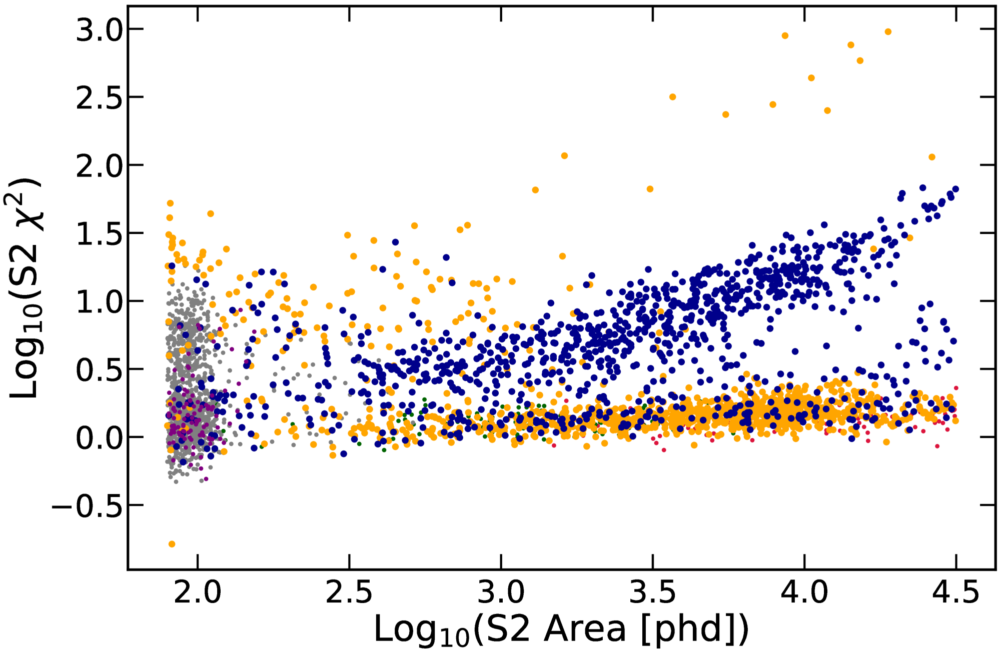


Figure 4.4: (x, y) reconstruction χ^2 of iS2 events, clustered by their origin. Events of poor reconstruction quality predominantly originate from the glue ring (blue). Clusters and their respective colours are annotated in Figure 4.2.

Hotspots

The gate, anode, and cathode in LZ each contain regions of small field perturbations, as a result of imperfections and unavoidable residue. This results in field emission of electrons, resulting in iS2 pulses of sizes $\mathcal{O}(1e)$. As these iS2s are localised in (x, y), they can be removed using a cut in (x, y) during periods of hotspot field emission [133].

Figure 4.5 shows the (x, y) position of events from each cluster of iS2 events. iS2s from hotspots are clustered in fixed positions. The location of glue ring events and the field cage resistor iS2s can also clearly be seen.

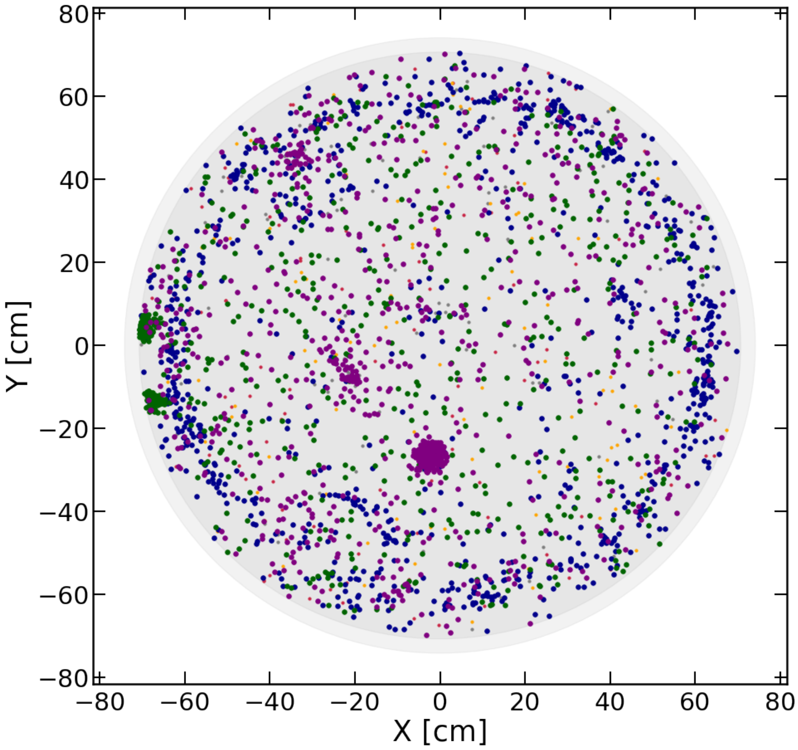


Figure 4.5: (x, y) position of iS2 events, clustered by their origin. A distinct population can be seen from a number of hotspots (purple), as well as the field cage resistors (green) and the glue ring (blue). Clusters and their respective colours are annotated in Figure 4.2.

E-Trains

Electrons from an S2 pulse can be delayed when drifting towards the extraction region. Delays from thermalisation at the liquid surface, $\mathcal{O}(100\,\mathrm{ns})$, are negligible relative to the TPC drift time, $\mathcal{O}(1\,\mathrm{ms})$ (940 µs in the first science run [85]).

However, delays from the absorption and re-emission of electrons by LXe oxygen impurities $\mathcal{O}(100\,\mathrm{ms})$ result in a region of low data quality across multiple event windows following a high-energy event. This delayed emission is referred to an an E-Train.

If several electrons from an E-Train are coincident, they can be classified as an iS2 pulse. E-Trains have been observed in previous dual-phase TPCs [117, 134–137], and an analysis cut using this understanding is applied in LZ, rejecting events within a time window following an S2 pulse [133].

Liquid Surface

Events produced near to the liquid surface in LZ produce an S1 and S2 pulse which cannot be resolved, producing an iS2. Events in this region do not diffuse through the TPC, and can therefore be rejected through pulse shape discrimination.

Figure 4.6 shows contours of the time for a waveform to reach 5% of its integrated pulse relative to 1% (AFT5-AFT1) and the full width at half maximum (FWHM) for iS2 populations. iS2s tend towards lower values of AFT5-AFT1 and FWHM compared to signal events. As the minimum drift time in data is set at 70 µs to reject above-anode events (discussed in Chapter 6), a cut on low rise times and FWHM can reject the majority of liquid surface iS2s with minimal signal loss.

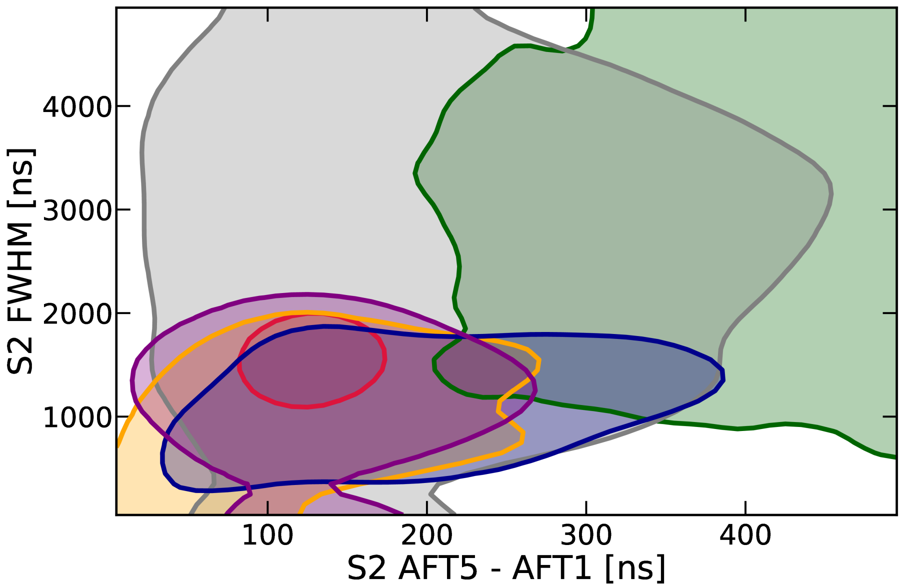


Figure 4.6: 3% Contours of AFT5-AFT1 vs FWHM of iS2 events, clustered by their origin. As the majority of iS2s are produced by the extraction region, iS2s tend to have low AFT5-AFT1 and FWHM compared to signal data. Clusters and their respective colours are annotated in Figure 4.2.

Gate and Cathode

The components of the grids were part of an intense radiopurity campaign before installation in LZ [126], in order to mitigate backgrounds. However, the gate and cathode arrays remain a source of radiogenic events in LZ, produced by ^{222}Rn plated out on the wires during production. The origin of grid radiogenic events, as well as how to reject them as part of an analysis of iS2 data, is the focus of Chapter 6.

Field Cage Resistors

The TPC field cage, surrounding the walls, ensures that the electric field in the drift region of LZ is constant and uniform. The field cage contains large chains of resistors [125]. One of these chains is observed to produce a large number of distorted S2s and iS2s in data, most likely the result of a field distortion. These events are easily removed with a cut in (x, y), as seen in Figure 4.5.

4.1.3 Isolated S1s in LZ

S1 pulses within the ROI are small, providing limited information as to their origin. However, some iS1 properties can still be identified in data, by clustering events as in Section 4.1.2. iS1 events cluster into 6 populations:

Dark Counts: The stochastic combination of PMT dark rate into a small S1 pulse.

High Single Channel: Radiation internal to a PMT.

Wide S1: Regions of high photon rate misclassified as an S1 pulse.

Low TBA: S1 pulses produced in the RFR, where an S2 is not detected.

High TBA: S1 pulses produced above the anode, where an S2 is not detected

Other: The origin of these events could not be ascertained.

The distribution of iS1 pulse area for each population can be seen in Figure 4.7.

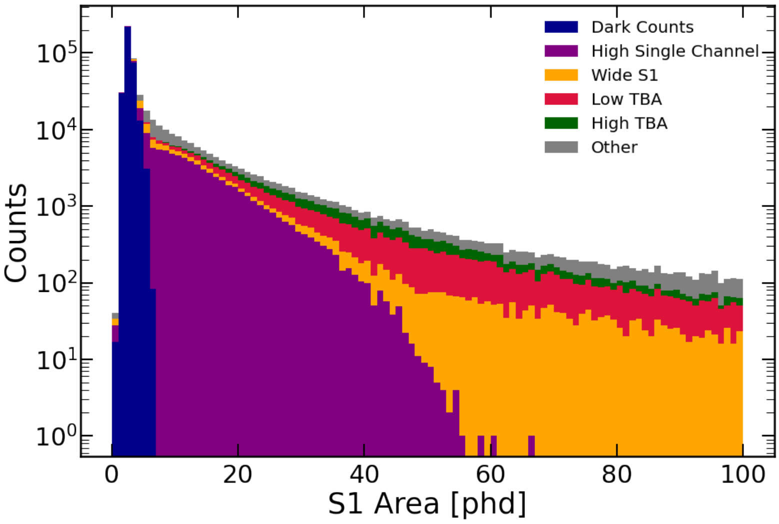


Figure 4.7: The area of iS1 pulses in WIMP search data, clustered using a simple GMM. iS1s are clustered as dark rate (blue), high single channel S1s (purple), broad pulses (orange), iS1s with low (red) and high (green) TBA, and other (grey).

Dark Counts

The most prominent contribution to iS1 rates is from the stochastic combination of PMT dark counts. When not exposed to external light, a PMT will still produce a small signal. Where sections of this signal are indistinguishable from signal produced by a single photoelectron, the resultant pulse is referred to as a dark count [138]. Sources of dark counts in PMTs can include internal radiogenics, field electrons emitted from high internal voltage, and the impact of avalanche photons with the walls of the PMT resulting in fluorescence. The latter does not contribute to dark counts in LZ, as PMTs are constructed with a Kovar body, which absorbs avalanche photons without fluorescence [125].

The average dark rate of the PMTs used in LZ was measured as 30 Hz [138]. As Cerenkov photons are not considered dark counts due to their ability to be discriminated in analysis, and fluorescence photons are not produced in PMTs, this dark rate is uncorrelated between PMTs. Dark counts can therefore be modelled as Poisson sources.

The probability of seeing a single Poisson event in a time interval can be expressed as

$$P(\lambda) = \lambda e^{-\lambda}, \tag{4.1}$$

where λ denotes the mean counts expected in a time window. When considering the pileup of dark counts into an S1 pulse, this is $2rT$, where r is the dark rate for a single PMT and $2T$ is the time before and after one dark count which another dark count must occur in for the waveform to be classified as an iS1.

As $rT \ll 1$, Equation 4.1 can be Taylor expanded to first order as $P = 2rT$. The process of n counts combining in coincidence from N PMTs is then a Binomial process, expressed as

$$P(n, N, r, T) = \frac{(N-1)!}{(n)!(N-1-n)!}(2rT)^n(1-2rT)^{N-1-n}. \tag{4.2}$$

An S1 is defined as three or more photons present within a 100 ns time window. Substituting the dark rate, $n = 3$, and $T = 100$ into Equation 4.2 gives an iS1 rate of approximately 60 mHz.

Dark count pileup produces small iS1 pulses, which are difficult to discriminate from signal S1s. As the average charge from a dark count is lower than that produced in a PMT by an incident photon, a fraction of iS1s from dark rate are rejected by pulse classification software.

High Single Channel

Many sources of radiation internal to a PMT will produce an iS1 that can be discriminated from dark counts. These can include PMT afterpulsing, and Cerenkov light produced through the β decay of ^{40}K contaminants in the PMT quartz window. If light from such an interaction escapes into the LXe bulk and reaches other PMTs, or the interaction coincides with other single photoelectrons (SPEs), this can be detected as an iS1. Such events can be easily discriminated, however, due to the large pulse area within a single PMT, and the large pulse width of the iS1 [139].

Wide S1

Many backgrounds in LZ can result in a temporarily high TPC photon rate. These include hotspots and E-Trains, discussed in Section 4.1.2, but also the fluorescence of PTFE surrounding PMTs [140]. As with dark counts, these photons can arrive in coincidence, producing an iS1 of high pulse width. The majority of these iS1s are rejected through the removal of periods of low data quality.

TBA

Events occurring at the very top of the TPC may produce an S2 which does not reach the extraction region. Similarly, events in the RFR produce an S1 pulse without an accompanying S2. As these iS1s occur near to the top and bottom of the detector, more light is produced in PMTs on the top or bottom array, resulting in a top-bottom asymmetry (TBA) in pulse area. A cut using the top-bottom asymmetry (TBA) of an event and its drift time to reject accidentals is discussed in Section 4.3.1.

4.2 Modelling Accidental Coincidence Events

4.2.1 Theoretical Estimation

The rate of isolated pulses can be used to calculate the expected rate of accidental coincidence events. As accidentals are assumed to be uncorrelated, the accidental rate can be calculated as the product of the iS1 rate and the probability of an iS2 arriving ahead of this within the maximum drift time T, expressed as

$$R(r_1, r_2, T) = r_1 P_2 = r_1(1 - e^{-r_2 T}) \approx r_1 r_2 T. \tag{4.3}$$

Before operation, the isolated pulse rates in LZ were derived from simulations to be $\mathcal{O}(1\,\mathrm{Hz})$ and $\mathcal{O}(0.1\,\mathrm{Hz})$ for iS1s and iS2s respectively. The maximum drift time during the first science run of LZ was 940 µs. Inserting these values into equation 4.3 gives a projected accidental rate of 0.094 mHz, or approximately 8 events/day within the NR band.

While accidental backgrounds were negligible in previous dual-phase TPCs, they are a primary background in an experiment such as LZ. As a result, a model of accidentals is required for a measurement of WIMP sensitivity, and analysis cuts are required to reduce the number of accidental events wherever possible.

4.2.2 Empirical Modelling of Accidentals Rates

For the impact of analysis cuts on accidentals to be measured, and for their impact on WIMP sensitivity to be evaluated, a model of accidental events is required.

However, waveforms are processed through a multi-layered pulse and event classification software, referred to as the LZ analysis package (LZap). The pulse and event classification algorithms in LZap are discussed in Chapter 5.

LZap is tuned using signal-like data. As a result, many accidental events which do not resemble signal are not classified as single scatters. For example, as mentioned in Section 4.1.3, LZap does not classify all instances of dark count pileup as iS1 pulses. Pulses which in theory would form iS1s and iS2s may not do so, and events which contain an iS1 and iS2 may still not be classified as a single scatter event.

This feature of LZap successfully lowers the rate of accidental events in data. However, it is too complex to be accounted for in simple calculations of accidental rate and distribution. As a result, a selection of accidental-like events are required which have been processed through LZap, for tuning of cuts and modelling of accidental backgrounds.

Unphysical Drift Time (UDT) Events

Two methods have been designed for modelling LZap-classified accidental events. The first of these exploits a feature of event classification: that there is no dependence on the S1-S2 drift time of an event. As a result, there exist a population of events in data with a S1-S2 time separation above the maximum TPC drift time. This concept is illustrated in Figure 4.8.

These unphysical drift time (UDT) events can be used to model accidental backgrounds in LZ. A distribution of event drift times in data from the first science run of LZ can be seen in Figure 4.9.

Events can be separated into four distinct regions. The first contains events

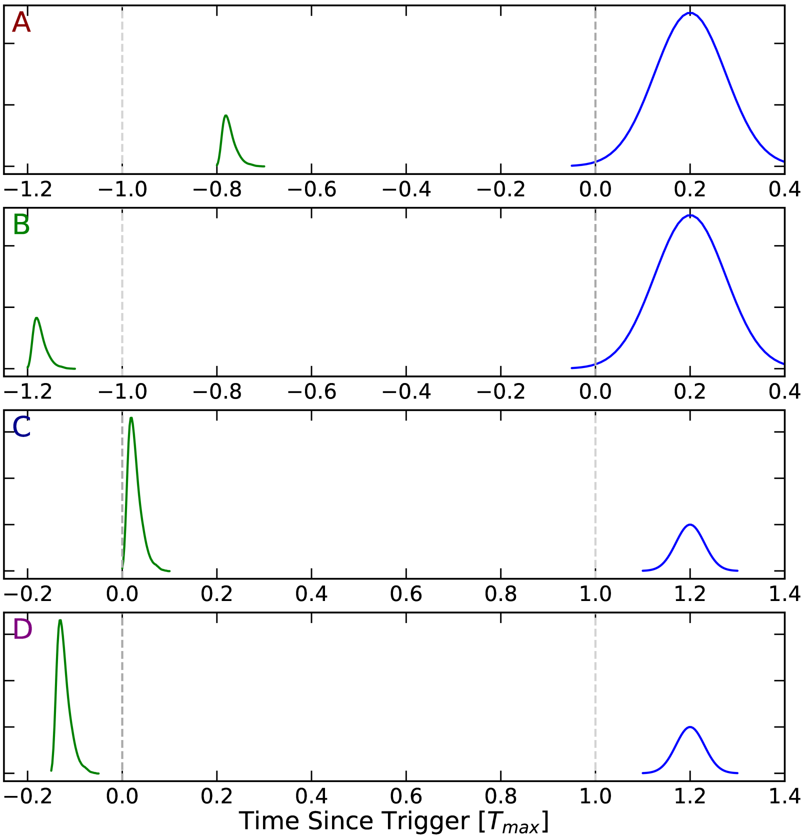


Figure 4.8: Illustration of the four primary trigger conditions in LZ: A) A physical event, B) An S2-triggered accidental, C) An S1-triggered accidental, and D) an accidental triggered by neither the event's S1 or S2.

of a physical drift time. The second is comprised of events where an iS2 has been triggered on, with the iS1 appearing between then and the beginning of the event window −2000 µs. At drift times above this, the iS1 has been triggered on, with an iS2 appearing between then and the end of the event window 2500 µs. Finally, if neither the iS1 or iS2 are triggered on, an accidental event can take a drift time of up to 4500 µs.

The details and origin of each UDT population in Figure 4.9 are outlined in Table 4.1.

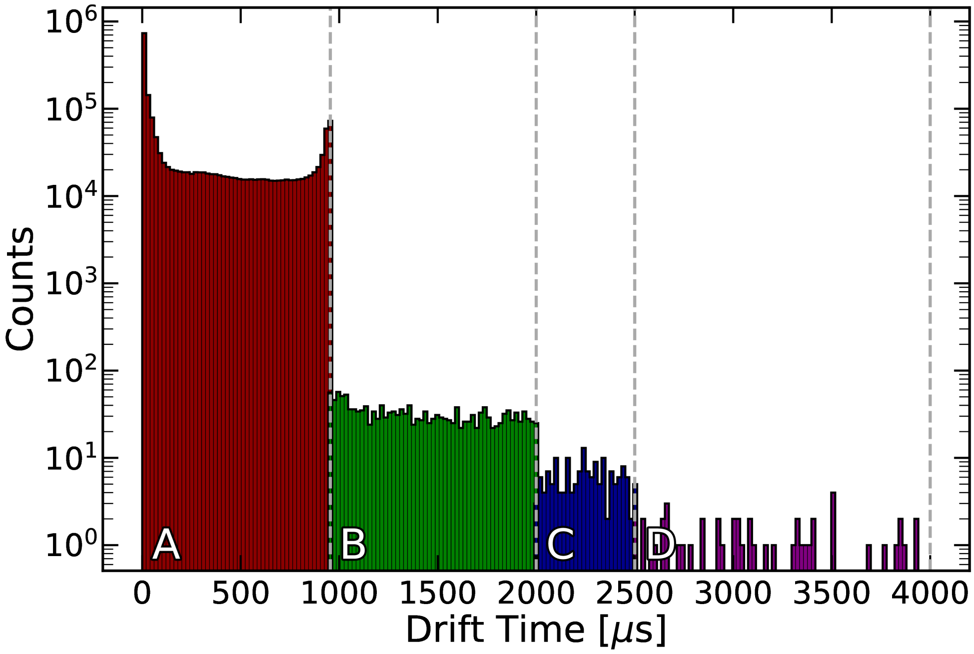


Figure 4.9: Distribution of the drift times of events in the first science run of LZ data. Dotted lines are included to highlight four distinct populations of data. Events outside of population A are UDT events.

Table 4.1: Drift time populations in WIMP search data. For each population, the triggered pulse as well as the main source of this population are given.

Region	Drift Times [μs]	S1	S2	Main Pathology
A	0 µs–940 µs	Triggered	Not triggered	Standard SS events
B	940 µs–2000 µs	Triggered	Not triggered	Typical accidentals
C	2000 µs–2500 µs	Not triggered	Triggered	Large-S1 accidentals
D	2500 µs–4000 µs	Not triggered	Not triggered	E-Train accidentals

Accidental Chop-Stitch (ACS) Events

UDT events are known to be accidentals, and are processed through LZap. However, they are limited in statistics, and can be correlated with the event trigger system. A method of modelling accidentals which avoids this issue involves artificially combining iS1 and iS2 waveforms, and then reprocessing this through LZap.

These accidental chop-stitch (ACS) events can be generated with high statistics and in a manner that is trigger-agnostic, but are not capable of measuring the accidental event rate. A final accidentals PDF for LZ is therefore generated using both of these populations. ACS events are used to generate the shape of the PDF in S1-S2 space, as well as to evaluate the rejection power of cuts, while UDT events are used to scale this PDF to the true accidental event rate.

Figure 4.10 shows the PDF of ACS events for the first science run of LZ within the WIMP ROI, cross-checked and scaled to reflect rates observed of UDT events. Integrating the entire PDF gives an accidental count of 113 events in 60 live days of data, with 9 of these within the NR band.

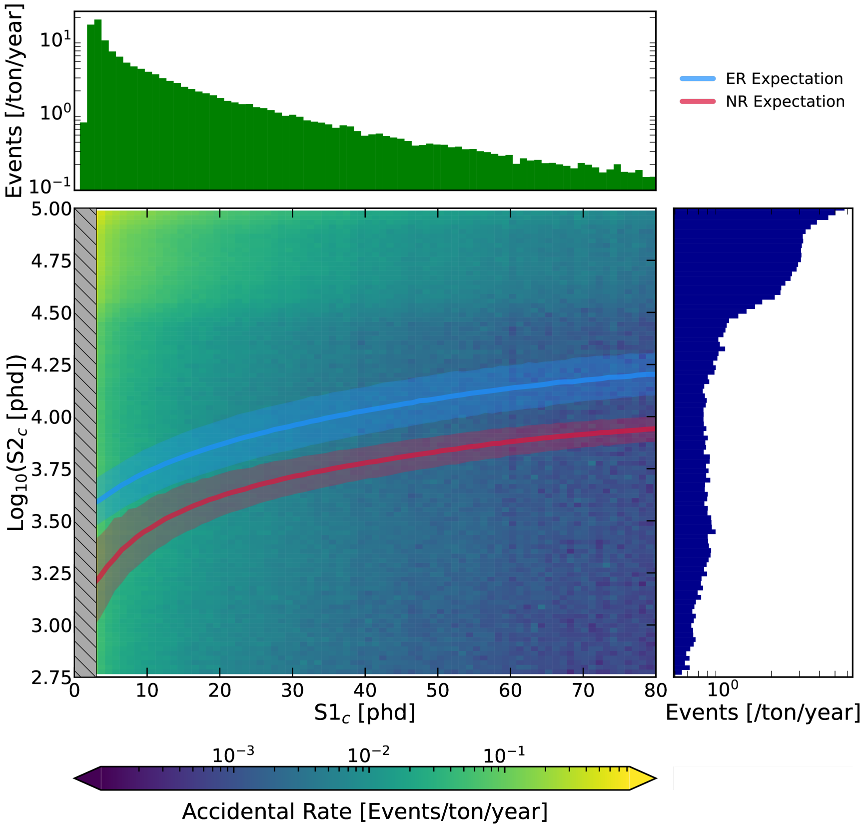


Figure 4.10: PDF of $\mathrm{S1}_c$ vs $\mathrm{S2}_c$ of accidental events for the first science run of LZ, measured in events/tonne/day. Histograms above and to the right show projections of this data in $\mathrm{S1}_c$ and $\mathrm{S2}_c$. A grey shaded region covers $\mathrm{S1}_c < 3\,\mathrm{phd}$, which was not considered in analysis. The ER and NR band are overlain in blue and red, with shaded regions indicating the 3σ contour of each.

4.3 Rejection of Accidental Events

Two approaches can be taken to rejecting accidental events. The first is background-focused, and looks at properties of iS1s and iS2s that are different from signal S1s and S2s. Cuts exploiting these differences are discussed in Sections 4.1.2 and 4.1.3.

The second approach is instead signal-focused, and exploits the fact that the drift time of an accidental event is not correlated to pulse parameters. In signal data, the width of an S2 pulse is correlated to its drift time, as electrons diffuse as they drift towards the extraction region. Likewise, the TBA of an S1 will tend towards higher values when an event occurs towards the top of the TPC, at low drift times.

4.3.1 S1 Top-Bottom Asymmetry

Theory

The top-bottom asymmetry (TBA) of a pulse can be defined as

$$\text{TBA} = \frac{A_{top} - A_{bottom}}{A_{top} + A_{bottom}}, \tag{4.4}$$

where A denotes the total pulse area collected by the top and bottom PMT array, typically given in units of photons detected (phd). As scintillation is produced at the interaction vertex, S1 TBA is correlated with drift time for signal events.

As photons reflect from the liquid surface towards the bottom PMT array, TBA in S1s tends towards negative values. The probability of reflection also varies with drift time, in a manner which cannot be derived. However, a fourth-order polynomial is found to best reflect the relationship of S1 TBA with changing drift time:

$$\text{TBA} = p_0 + p_1 t + p_2 t^2 + p_3 t^3 + p_4 t^4 \tag{4.5}$$

Fitting to Data

The S1 TBA relationship can be measured by fitting Equation 4.5 to data. As smaller S1s are more likely to be from accidentals, events with S1 areas in the range 600 phd to 1000 phd from CH_3T calibration data are used. Data is divided into slices of drift time. Each slice is then fitted to a normal distribution. The mean and standard deviation of these fitted slices can then themselves be fitted to Equation 4.5. This fit, and the distribution of events used for this fit, can be seen in Figure 4.11.

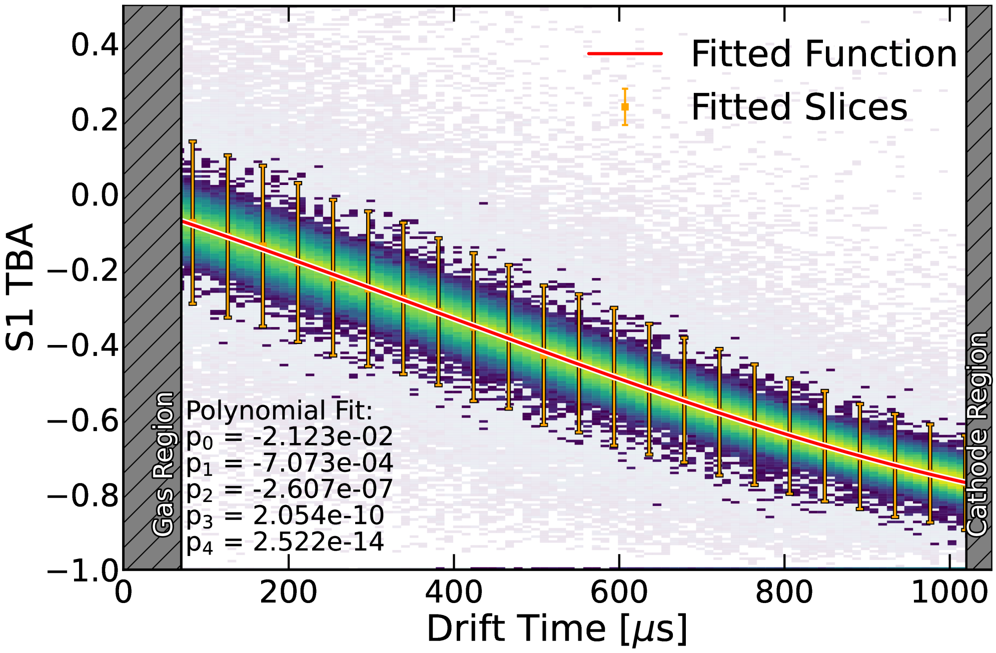


Figure 4.11: Fitting S1 TBA of CH_3T calibration data to Equation 4.5. Slices of data fitted to normal distributions are shown in orange (error bars denote standard deviation), and fitted Equation 4.5 is shown in red ($p_0 = -2.123 \times 10^{-2}$, $p_1 = -7.073 \times 10^{-4}\,\mu s^{-1}$, $p_2 = -2.607 \times 10^{-7}\,\mu s^{-2}$ $p_3 = 2.054 \times 10^{-10}\,\mu s^{-3}$, and $p_4 = -2.522 \times 10^{-14}\,\mu s^{-4}$, to 4 significant figures).

Resolution

The mean relationship between drift time and S1 TBA is expected to remain in (x, y) and S1 pulse area. However, the resolution of the derived S1 TBA improves for larger S1s.

The number of photons recorded by the top and bottom PMT array can be defined as

$$A_T = A_{\text{SPE}} n_T, \qquad A_B = A_{\text{SPE}} n_B, \tag{4.6}$$

where A_{SPE} is the area in phd of a single photon (≈ 1) and n_i is the total number of photons in the top and bottom arrays. If it is assumed the value n_i is binomially distributed, and that all photons impact either the top or bottom arrays, Equation 4.6 can instead be written in the form

$$A_T = A_{\text{SPE}} NP, \qquad A_B = A_{\text{SPE}}(N - NP), \tag{4.7}$$

where N is the number of detected photons and P is the probability of these photons hitting the top array. Substituting these values into Equation 4.4 gives us

$$\text{TBA} = \frac{A_T - A_B}{A_T + A_B} = \frac{A_{\text{SPE}}(NP + NP - N)}{A_{\text{SPE}} N} = 2P - 1. \tag{4.8}$$

From this, a value of the binomial probability P can be measured for a given value of TBA.

There are two dominant sources of error on the value of S1 TBA - binomial fluctuations in the number of photons incident on each PMT array, and fluctuations in the pulse area of each photon. The former can be assumed to take the form $\sigma_{n_T} = \sqrt{NP(1-P)}$, and the latter has been measured to be $\sigma_{A_{\text{SPE}}} \approx 0.4$ using the area distribution of single photoelectron pulses.

If it is assumed that the errors on S1 TBA are uncorrelated, the error on each of these parameters can be given as

$$\sigma_{A_T} = \sqrt{A_{\mathrm{SPE}}^2 NP(1-P) + (NP\sigma_{A_{\mathrm{SPE}}})^2} \tag{4.9}$$

and

$$\sigma_{A_B} = \sqrt{A_{\mathrm{SPE}}^2 NP(1-P) + (N(1-P)\sigma_{A_{\mathrm{SPE}}})^2}. \tag{4.10}$$

The error on TBA can then be expressed as

$$\sigma_{TBA} = \sqrt{\left(\frac{2A_B\sigma_{A_T}}{N^2}\right)^2 + \left(\frac{2A_T\sigma_{A_B}}{N^2}\right)^2}, \tag{4.11}$$

where σ_{TBA} is the error on a given measurement of S1 TBA.

The validity of Equation 4.11 can be demonstrated by subtracting the mean from each recorded value of S1 TBA and scaling by σ_{TBA} - in other words, by measuring the number of standard deviations of a given event from the mean. The resulting distribution, if Equations 4.5 and 4.11 perfectly represent real data, should have a mean of 0 and standard deviation of 1. Figure 4.12 shows this data fitted to a Gaussian of mean -0.01 and standard deviation 1.06. These equations can therefore be said to accurately represent the TBA of S1 pulses.

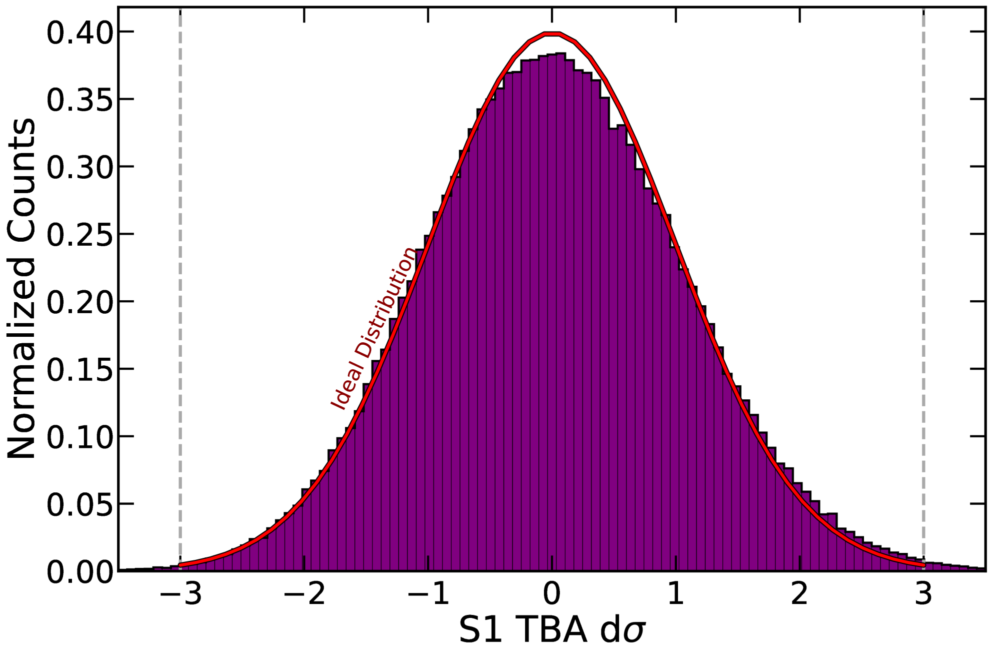


Figure 4.12: The difference between mean and recorded S1 TBA in CH_3T calibration data, scaled by σ_{TBA}. A normal distribution is included for reference. The mean and standard deviation of this data are close to 0 and 1, indicating that Equations 4.5 and 4.11 are accurate.

4.3.2 S2 Pulse Width

Theory

The width of an S2 pulse is determined by two parameters - the time for an electron to pass through the extraction region, which determines the pulse width of a single electron σ_0, and the variation in time to reach the extraction region σ_e.

Transverse diffusion is not expected to exceed 1 cm at any position in the TPC [141]. The standard deviation of electron time displacement can therefore be expressed as

$$\sigma_e = \frac{\sqrt{2D_L t}}{v_d}, \tag{4.12}$$

where D_L is the longitudinal diffusion coefficient, v_d is the electron drift velocity, and t is the drift time.

If $\sigma_0 \lessapprox 0.8\sigma_e$, the convolution of the single electron response with the electron time distribution gives the S2 time variance, in the form of Equation 4.13:

$$\sigma_{S2} = \sqrt{\sigma_e^2 + \sigma_0^2} = \sqrt{\frac{2D_L}{v_d^2}t + \sigma_0^2}. \tag{4.13}$$

From this, it can be concluded that the standard deviation of an S2 pulse and the drift time of an event are correlated for non-accidental single scatters [141].

The definition of pulse width typically used in LZ involves the time for a pulse to reach a fraction A% of its total integrated charge from the pulse start T_0, known as the Area Fraction Time (AFT). A width can then be defined as AFT(100-A%) - AFT(A%). This prevents fluctuations in pulse widths from pulse classification, and can be directly related to the standard deviation σ for a Gaussian-like pulse.

The AFT of a pulse can then be approximated as

$$\mathrm{R} = \int_{T_0}^{\mathrm{AFT_R}} N(x, \mu = 0, \sigma) = \frac{1}{2}\mathrm{erf}(\frac{\mathrm{AFT_R}}{\sqrt{2}\sigma}) + 1 \tag{4.14}$$

$$\mathrm{AFT_R} = \sqrt{2}\sigma\mathrm{erf}^{-1}(2\mathrm{R} - 1), \tag{4.15}$$

where R denotes the area fraction. As an S2 is assumed to be symmetric, the width of an S2, W_{R}, can be approximated as

$$W_R = 2\mathrm{AFT_R}. \tag{4.16}$$

Equation 4.16 can be used to scale the standard deviation σ_{S2} to any given pulse width metric. A pulse width defined using AFTs therefore follows the relationship expressed in Equation 4.13. This also indicates that, for a standard S2 pulse, definitions of pulse width can be expected to correlate with one another, a feature exploited to identify mis-classified pulses in Chapter 5 and to reject grid radiogenic events in Chapter 6.

A broader definition of S2 pulse width has a greater sensitivity to changes in pulse shape. As the majority of iS2 pulses are non-Gaussian, a wider metric is therefore capable of greater rejection of accidental events. However, this comes at the expense of signal acceptance, the fraction of signal-like events preserved by the cut. As this cut is designed to reject accidentals with minimal impact on signal acceptance, and other pulse width metrics were found to vary with iterations of pulse classification software and as a result of PMT saturation, the decision has been made to use AFT75-AFT25 in this context.

Fitting to Data

Equation 4.13 can be rewritten as

$$W_{S2} = \sqrt{M^2 t + C^2}, M = \frac{A\sqrt{2D_L}}{v_d}, C = A\sigma_0, \tag{4.17}$$

where A denotes the scale factor as described in Equation 4.16, and M and C can be fitted to data. As in Section 4.3.1, slices of calibration data in drift time can be fitted to normal distributions, and the mean of those normal distributions fitted to Equation 4.17, as seen in Figure 4.13.

Resolution

As the width of an S2 pulse is empirically measured, there is no clean analytic method with which to define its resolution. However, if the errors of parameters in Equation 4.17 are treated as uncorrelated, the error on a given S2 pulse width can be approximated as Equation 4.18.

$$\sigma_{W_{S2}} = \sqrt{\left(\frac{Mt}{W_{S2}}\sigma_M\right)^2 + \left(\frac{C}{W_{S2}}\sigma_C\right)^2 + \left(\frac{M^2}{2W_{S2}}\sigma_t\right)^2} \tag{4.18}$$

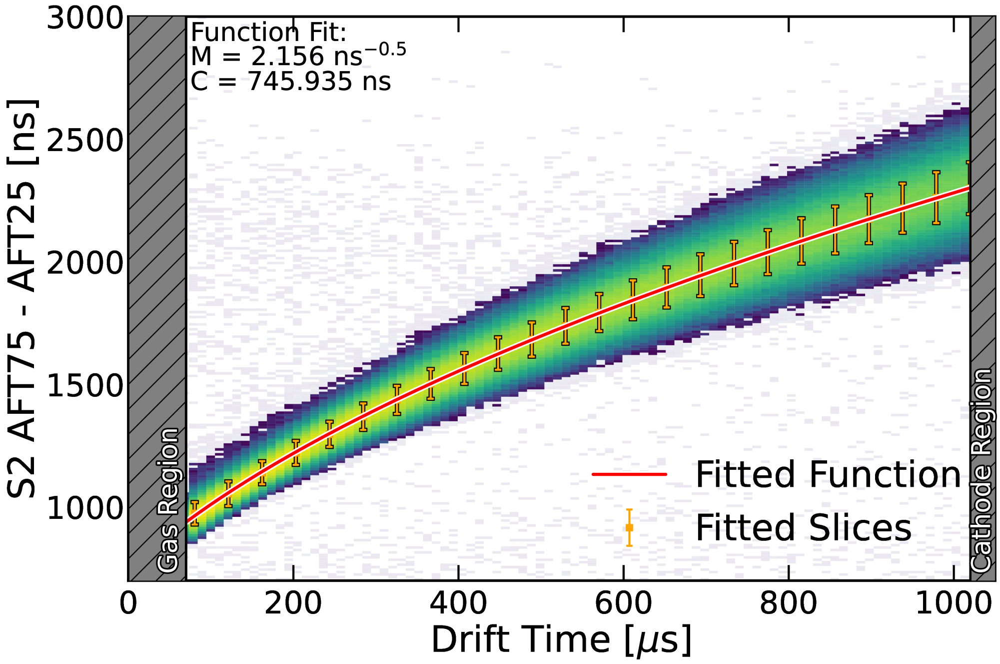


Figure 4.13: Fitting the S2 pulse widths of ^{83}Kr calibration data to Equation 4.17. Slices of data fitted to normal distributions are shown in orange (error bars denote standard deviation), and fitted Equation 4.17 is shown in red ($M = (2.156 \pm 0.001)\,\mathrm{ns}^{-0.5}$, $C = (745.9 \pm 1.5)\,\mathrm{ns}$, to 3 significant figures).

It has been demonstrated with toy model S2s that the values of σ_M and σ_C scale with $\frac{1}{\sqrt{N}}$, where N denotes the number of electrons in a pulse. The value σ_t should be constant, and would ideally match the PMT time resolution of 10 ns - however this value is allowed to float in a fit, to account for the shape of non-Gaussian S2 pulses.

Equation 4.18 can be fitted to data by subtracting the mean S2 pulse width from values of AFT75 - AFT25, scaling by $\sigma_{W_{S2}}$ to give the number of standard deviations, and tuning parameters in Equation 4.18 such that the standard deviation of the resulting distribution is equal to 1. Figure 4.14 displays such a distribution. The fitted values of σ_M, σ_C, and σ_t are $1.581\,\mathrm{ns}^{0.5}$, 150 ns, and 25 ns respectively. This distribution is not centred at zero, as higher-energy Kr events tend to higher pulse widths as the two constituent decays become resolvable. In signal data, a similar effect is observed as a result of PMT saturation. A cut using this fit therefore maintains signal acceptance for these events, with minimal background leakage (the fraction of accidental events rejected by a given cut).

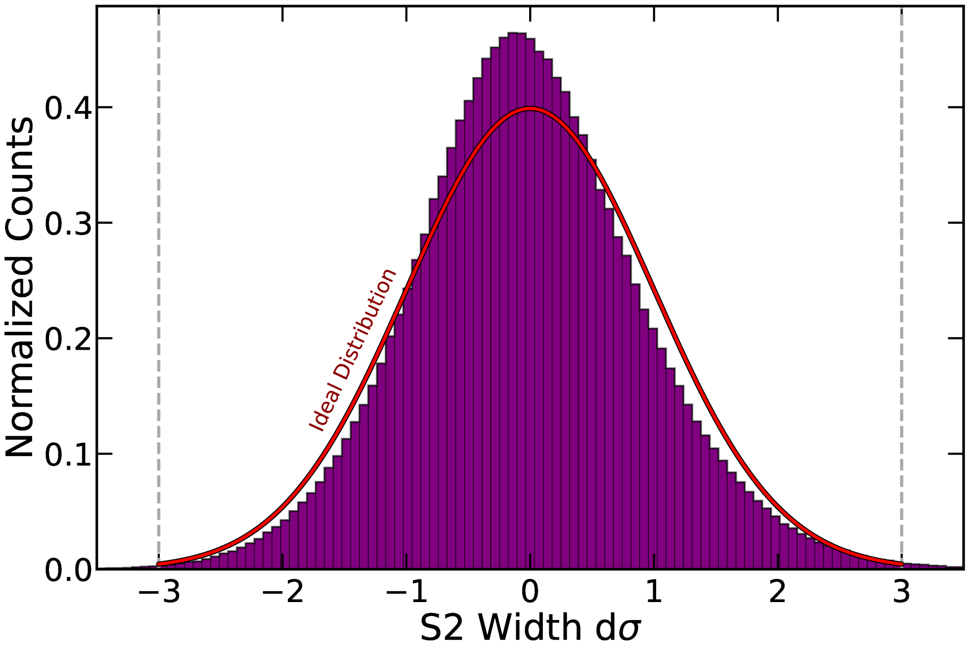


Figure 4.14: The difference between mean and recorded S2 AFT75-AFT25 in ^{83}Kr calibration data, scaled by $\sigma_{W_{S2}}$. A normal distribution is included for reference. The mean of this distribution differs from 0 due to deviations in pulse widths for high-energy Kr events. However, applying a cut at $\pm 3\sigma$ still preserves the vast majority of signal data, while enabling the removal of accidental coincidence events from analysis.

4.3.3 Signal Acceptance and Background Leakage

The relationships in an event's drift time identified in Sections 4.3.1 and 4.3.2 can be used as data quality cuts in LZ. The difference between an events S1 TBA and S2 pulse width from the mean expected relationship can be measured in units of σ. Events greater than $n\sigma$ from the mean can be rejected from analysis as accidental events.

As the distribution of S1 TBA in Figure 4.12 is symmetric, a cut is placed at $\pm3\sigma$. However, the distribution of S2 pulse width in the first science run is asymmetric due to PMT saturation. As a result, a cut is applied at $^{+3.5\sigma}_{-3\sigma}$.

The background leakage of data quality cuts is measured using ACS data. In the first science run, signal acceptance is measured using measurements of CH_3T, AmLi, ^{83}Kr and ^{220}Rn calibration data, discussed in Chapter 3. Preliminary cuts are applied in livetime, the outer detectors, and through the application of the fiducial volume cut.

The signal acceptance and background leakage of a cut in S1 TBA can be seen in Figure 4.15. Across all S1 pulse areas, signal acceptance is above 96%, with similar results for measurements using data from the first science run. Background leakage is 20% for large S1s and increases with decreasing pulse area.

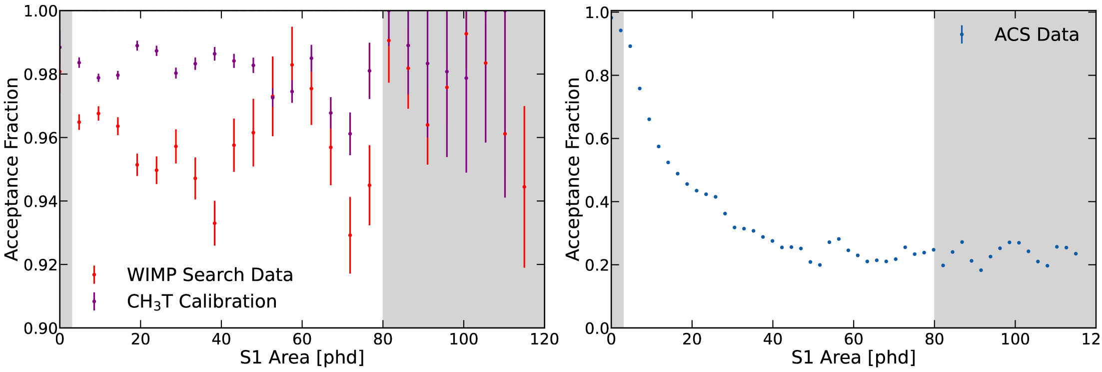


Figure 4.15: S1 TBA cut acceptance vs S1 pulse area for WIMP search data (left), CH_3T calibration (left), and ACS events (right). Shaded regions indicate S1 pulse areas outside of the WIMP ROI. Signal acceptance is above 98% in calibration data, while leakage of accidental events is as low as 20%.

The signal acceptance and background leakage of a cut in S2 pulse width can be seen in Figure 4.16. Signal acceptance in calibration data is greater than 99% up to an S2 pulse area of 40 000 phd, outside of the WIMP ROI. Above this, acceptance is observed to drop to 96% in Rn calibration data, due to PMT saturation. Measured signal acceptance is lower in WIMP search data than in calibrations, although this can be attributed to increased presence of accidental events. At high S2 pulse areas, background leakage is below 5%. At lower pulse areas, this leakage increases, reaching 100% at threshold.

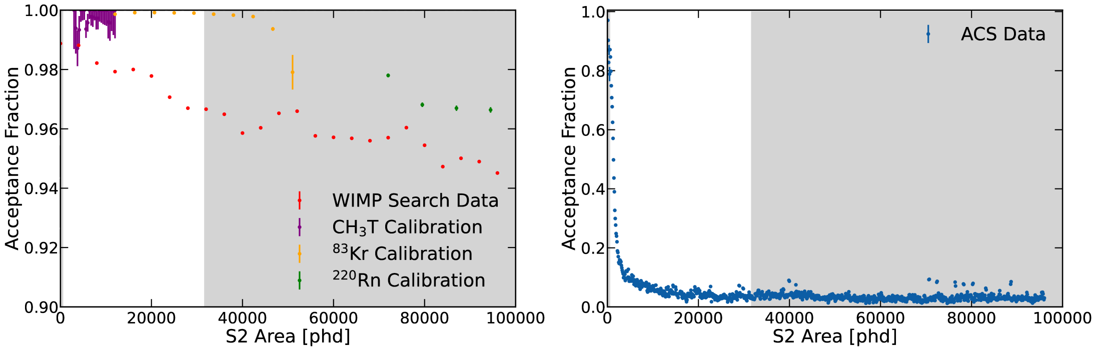


Figure 4.16: S2 width cut acceptance vs S2 pulse area for WIMP search data (left), CH_3T calibration (left), Rn calibration (left), Kr calibration (left), and ACS events (right). Shaded regions indicate data outside of the WIMP ROI. Signal acceptance decreases with increasing S2 pulse area, but remains above 95% in calibration data, while background leakage is as low as 5%.

The acceptance curves in Figures 4.15 and 4.16 can be combined to measure background leakage in S1-S2 space. The result of this can be seen in Figure 4.17. For the majority of the ER and NR bands, drift time cuts are capable of rejecting above 95% of accidental events.

4.3.4 Impact on WIMP Sensitivity

After the application of cuts targeting iS1 and iS2 pulses using the features outlined in Sections 4.1.2 and 4.1.3, above 80% of signal events in the WIMP ROI are preserved, with a background leakage below 0.5%. This results in the measured

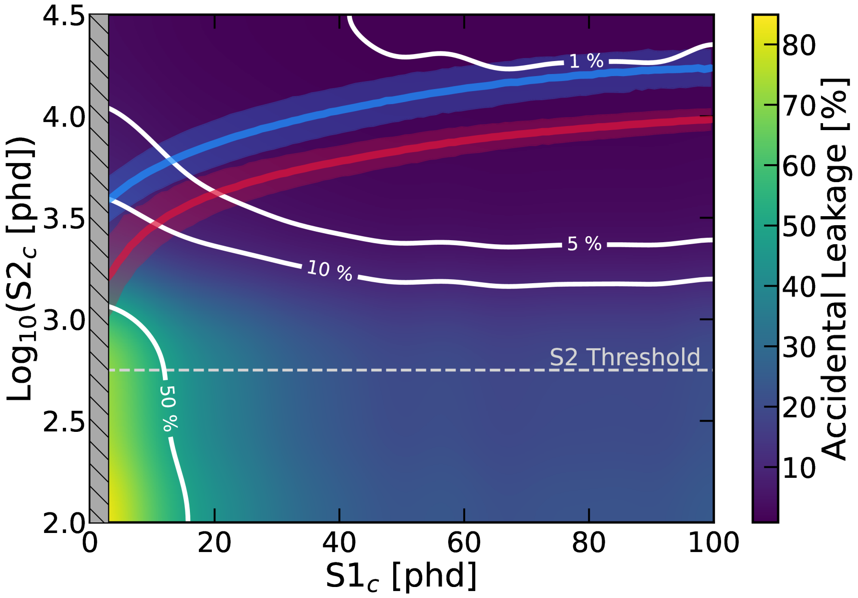


Figure 4.17: Leakage of accidental events after drift time cuts in the first science run of LZ, measured using ACS data. 50%, 10%, 5%, and 1% leakage are annotated as white contours. The S2 threshold is annotated in light grey, and the NR and ER bands are highlighted in red and blue, respectively.

number of accidentals in the first science run of LZ being just 1.2 ± 0.3 events [85].

To demonstrate the impact of the rejection of accidental events on WIMP sensitivity, the PLR analysis used to generate Figure 3.10 can be repeated using the distribution and projected number of accidental events before the application of data quality cuts.

The projected WIMP sensitivity without the rejection of accidental events can be seen in Figure 4.18. At low WIMP masses, accidental events are the dominant background, and impact sensitivity by up to two orders of magnitude. However, sensitivity at high WIMP masses is also limited by a high accidental rate, as the S1-S2 contour of high-mass WIMPs still predicts an excess of signal at small S1-S2. Without data quality cuts targeting accidentals, it would require up to an order of magnitude more live days of data to reach the achieved sensitivity of LZ to WIMPs.

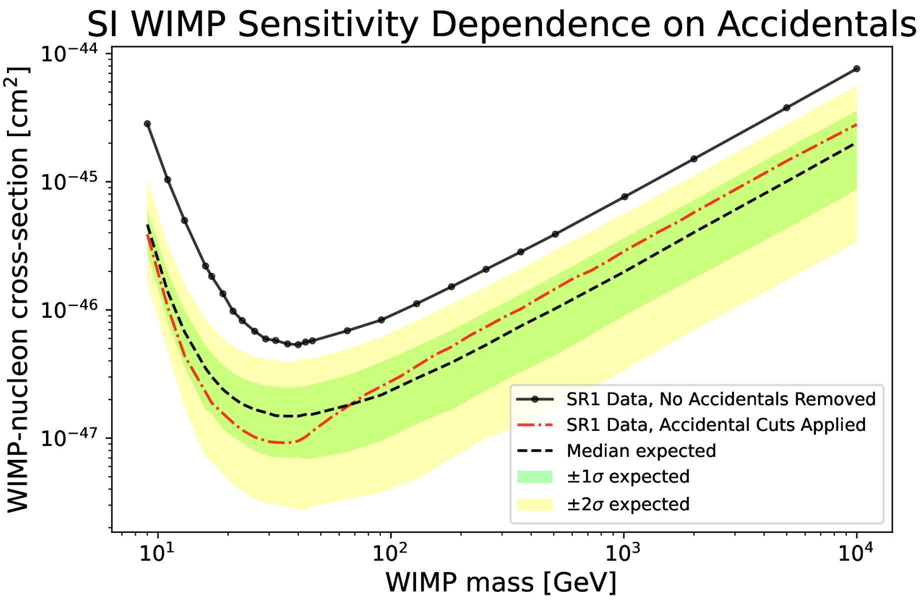


Figure 4.18: The 90% confidence limit for the spin-independent WIMP cross section vs WIMP mass, with accidental events present (black line) and removed (red dot-dashed line), from PLR analysis of data from the first science run of LZ [85]. The green and yellow bands are the 1σ and 2σ sensitivity bands. The black dotted line shows the median of the sensitivity projection. Without data quality cuts targeting accidentals, sensitivity decreases by up to an order of magnitude.

4.4 Accidental Backgrounds in Varying Drift Fields

4.4.1 Motivation

During the initial planning of LZ, it was proposed that a drift field of around 700 V/cm would be ideal for sensitivity to high-mass WIMPs [142]. During designs for LZ, the baseline drift field was chosen to be 310 V/cm [125]. During the first science run of LZ, an unanticipated surge of pulses in the skin region meant data was taken at a drift voltage of 193 V/cm [85]. This lower drift field has resulted in lower sensitivity to high-mass WIMPs than anticipated, although results in this region continue to surpass those of previous experiments. However, a lower drift field reduces the rate of recombination in events, increasing the probability of $S1_c > 3$ phd. This increases the sensitivity of LZ to low-mass WIMPs.

The initial decision of drift field was made due to the impact of ER leakage on sensitivity to high-mass WIMPs, which increases at lower drift fields. Figure 4.19 shows the variation of ER leakage, using projections from LUX Run 3 [143]. At high WIMP masses, ER leakage is high. As a result, ER leakage is a dominant background, and sensitivity is maximised with high drift field. However, ER leakage decreases exponentially at low WIMP masses. This is due to the movement of the NR band centroid in $S2_c$ at low WIMP mass [144, 145]. As a result, ER leakage has less impact on low-mass WIMP sensitivity.

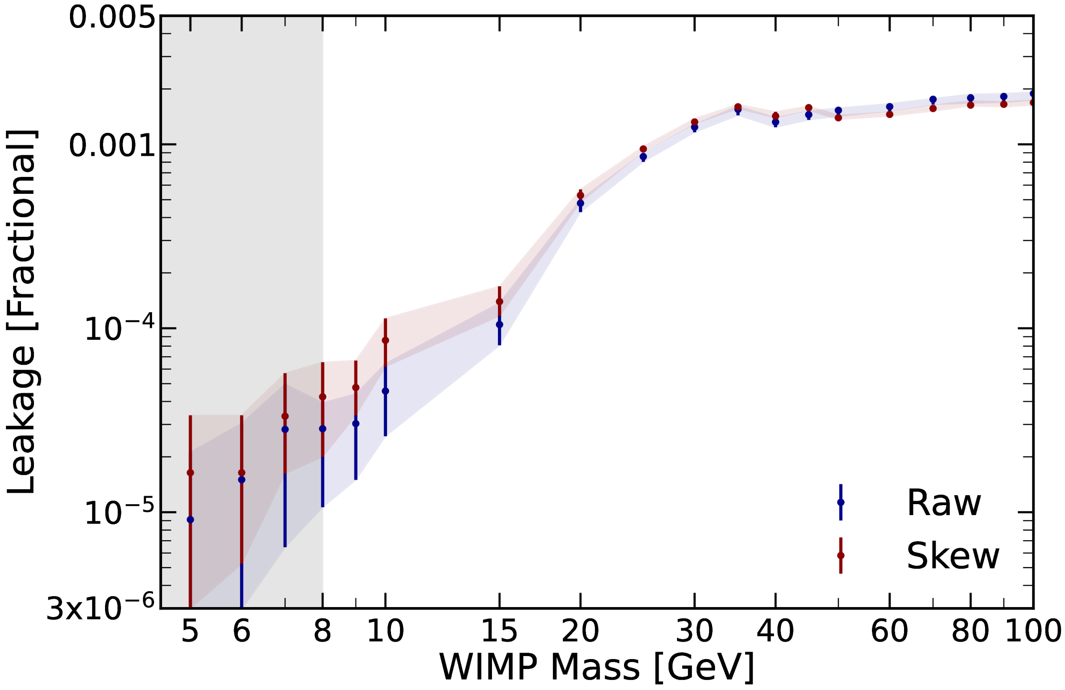


Figure 4.19: ER leakage, assuming a flat-background spectrum, as a function of WIMP mass from counting (red) and skew fits (blue) for LUX Run 3 [143], digitised from Ref. [144]. While leakage rises for a fixed NR band, the movement of the NR band centroid downwards in S2 pulse area for lower WIMP masses results in a lower ER leakage [145]. A grey shaded region indicates the range of WIMP masses not currently considered in LZ S1-S2 WIMP searches, below 8 GeV/c^2.

In low-mass WIMP searches, the dominant background is from accidental coincidence events. However, measurements of accidental backgrounds in dual-phase TPCs are minimal, preventing projections being made of the ideal operating conditions for a low-mass WIMP search.

This section outlines measurements of accidental event rates both before and after analysis cuts, for different drift fields in LZ. For various periods where LZ has operated at different drift fields, rates of iS1s and iS2s are extracted and used to derive an accidental event rate. For each field, the longitudinal diffusion coefficient, drift time, and pulse width resolution are interpolated. From this, the rejection power of a cut in S2 pulse width can be extrapolated. These measurements allow for the accidental pulse rate both before and after analysis cuts to be interpolated for all drift fields in LZ, and extrapolated to drift fields outside of existing data. These measurements can be used to identify the ideal drift field for a low-mass S1-S2 WIMP search in LZ.

4.4.2 Variation of Diffusion Coefficients

Measurements have been made in previous dual-phase LXe TPCs of the variation of drift velocity v_d and the longitudinal diffusion coefficient D_L with drift field [141, 146, 147]. These measurements have been integrated into NEST, such that it is capable of precisely modelling diffusion of all known signal and background sources in LZ [121]. However, fewer measurements have been made at low drift fields.

NEST calculates v_d using the empirical function

$$v_d = e^{A - \frac{B}{x} + C \log_{10}(x)}, \tag{4.19}$$

where x denotes the drift field strength in units of kV/cm, and A, B, and C are coefficients that vary based on LXe conditions. D_L is calculated using the functional form

$$D_L = A^{\lambda_A x} + B^{\frac{\lambda_B}{x}}, \tag{4.20}$$

where A, B, λ_A, and λ_B are empirically fitted variables.

The absence of measurements of D_L and v_d at low drift fields introduces a large uncertainty in calculated values below the operating conditions of LZ during its first science run. However, measurements can be made in LZ, using the relation between S2 pulse width and drift time.

Data was taken at a range of drift fields in LZ between science runs. The position of cathode radiogenic events in drift time can be used in each dataset to derive v_d. The S2 pulse width and drift time of data can be used to measure M from Equation 4.17, which can then be used to derive D_L.

Figure 4.20 shows the measured values of v_d and D_L in varying LZ drift fields. The variation of v_d with changing drift field is found to be in agreement with that measured in previous dual-phase TPCs [141, 146, 147].

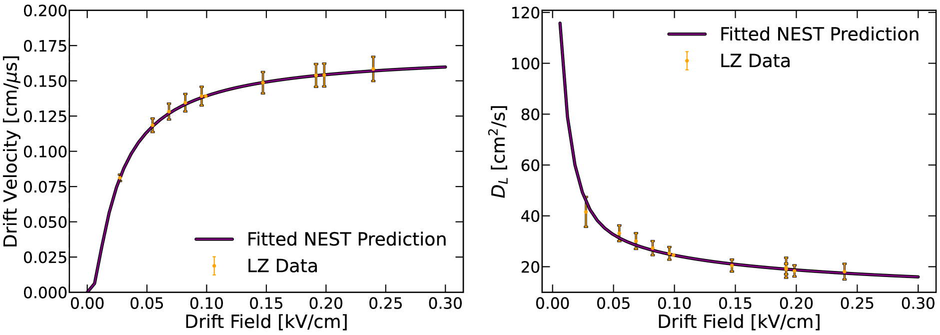


Figure 4.20: Left: Variation of v_d with changing drift field strength, with the functional form used by NEST fitted. The fitted relationship is in agreement with previous measurements. Right: Variation of D_L with changing drift field strength, with the function used in NEST fitted. Measurements of D_L have been used to improve models of diffusion in NEST. All errors are given to 3σ, for legibility.

Data shown in Figure 4.20 was implemented into NEST, to allow for more precise predictions of diffusion at low drift fields. Further measurements of diffusion using α radiation in LZ measure values of D_L in agreement with this data, within errors.

4.4.3 Variation of Accidental Event Rates

While taking data at different drift fields between science runs, iS1 and iS2 rates were monitored in LZ. These measurements concluded that the rate of isolated pulses remains consistent with changing drift field, and increases with increasing extraction field. These rates were measured in data as 0.8 Hz and 3 mHz for iS1s and iS2s, respectively. With this, the rate of accidental events can be interpolated for various drift fields, using Equation 4.3.

Figure 4.21 shows the variation of accidental event rate with changing drift field. At drift fields above 100 V/cm, the accidental event rate is relatively consistent. However, below this, the rate of accidentals increases asymptotically.

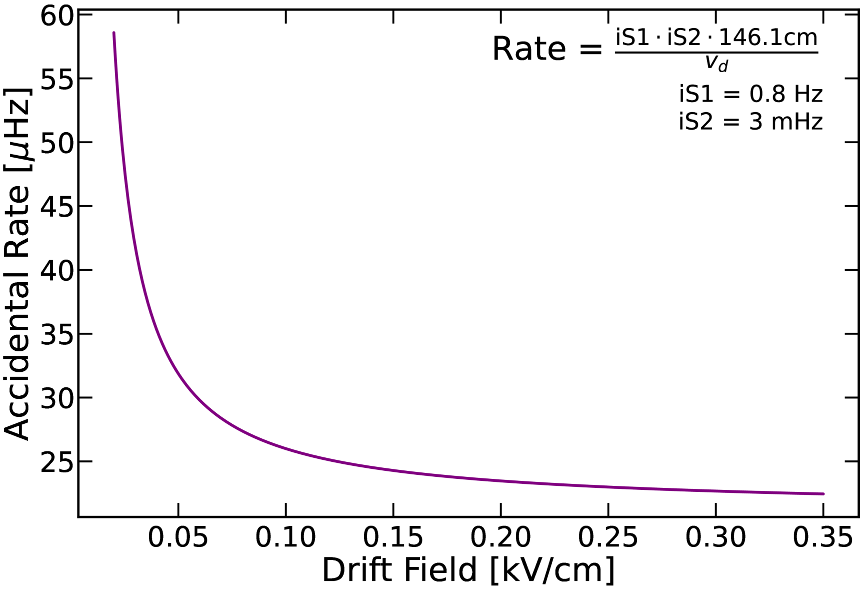


Figure 4.21: Using Equation 4.3 and data from Figure 4.20 to quantify accidental rates at varying drift fields. Above 100 V/cm, accidental event rate changes little with increasing drift field. Below this, the accidental event rate increases asymptotically.

A higher drift field reduces the rate of accidental events in data, as the drift time in which accidental events can take place is smaller. However, this relationship does not account for the impact of cuts on the rejection of accidental events. If accidental rates were high but rejection power approaches 100%, the number of accidental events present after cuts for statistical inference would be significantly lower than in current WIMP search data, improving sensitivity regardless of the initial accidental event rate.

4.4.4 Interpolation of Accidental Rejection Power

The shape and size of S1 pulses in LZ is found to change little with varying TPC drift voltage. As a result, the variation of cuts targeting iS1 events, outlined in Section 4.1.3 is negligible with changing drift field. Similarly, as S1 TBA is not expected to vary, the rejection power of the S1 TBA vs drift time cut will not change with varying drift fields.

Cuts targeting iS2 events target regions of livetime and the shape of iS2s very near to the extraction region, as outlined in Section 4.1.2. As these are events where electron diffusion is negligible, the rejection power of these cuts has not been observed to vary with changing drift field. However, the distribution of pulse width against drift time changes with drift field, affecting accidental rejection power.

S2 Pulse Width vs Drift Time in Changing Drift Fields

Equations 4.17 and 4.18 are found to be valid where $\sigma_e > \sigma_0$ [141]. A cut in pulse width is therefore valid up to drift fields $\mathcal{O}$(kV/cm). These equations more closely resemble data at lower drift fields, where diffusion dominates. However, below 50 V/cm, electron lifetime losses are projected to increase considerably, resulting in smaller S2s with increased pulse mis-classification (discussed in Chapter 5), and statistical effects changing the shape of the pulse width relationship in data.

The S2 pulse width of signal data is dependent on the variation of the parameters in Equations 4.17 and 4.18 with drift field. M is dependent on D_L and v_d, both of which are measured in Figure 4.20. The width of single electron pulses does not vary with changing drift field. As a result, the parameter C and its error σ_C will not vary with changing drift field. σ_t, an error of order 10 ns, is negligible relative to other errors on S2 pulse width within the WIMP ROI. This leaves a measurement of the variation of σ_M with changing drift fields.

The limited data taken at varying drift fields results in large errors when measuring σ_M in data. However, a toy model of S2 pulse widths can be used to quantify how σ_M varies with M. The details of this toy model are outlined in Appendix B.

Toy model S2s are generated at a drift time of 940 µs and of size $n_e = 100$ for constant C and varying values of M. The standard deviation of pulse widths for each generated dataset was then derived. This is plotted against M in Figure 4.22. The standard deviation of pulse widths is directly proportional to M. As a result, it can be concluded that σ_M is also proportional to M.

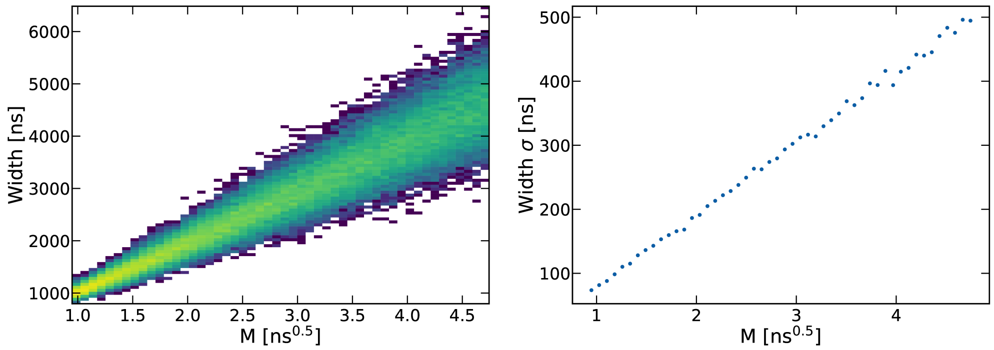


Figure 4.22: Left: Data generated using Equation 2 at T of 940 µs and N of 100 for constant C and varying values of M. Right: Standard deviation of this data plotted against M.

S2 Pulse Width Rejection Power

With this understanding of pulse width variation, the number of accidental events remaining after a cut in S2 width can be calculated using ACS data from the first science run of LZ, which have been shown to represent true accidental events through thorough cross-checks with UDT events. The distribution of ACS events in S2 pulse width and drift time is shown in Figure 4.23, alongside an illustration of the S2 width cut.

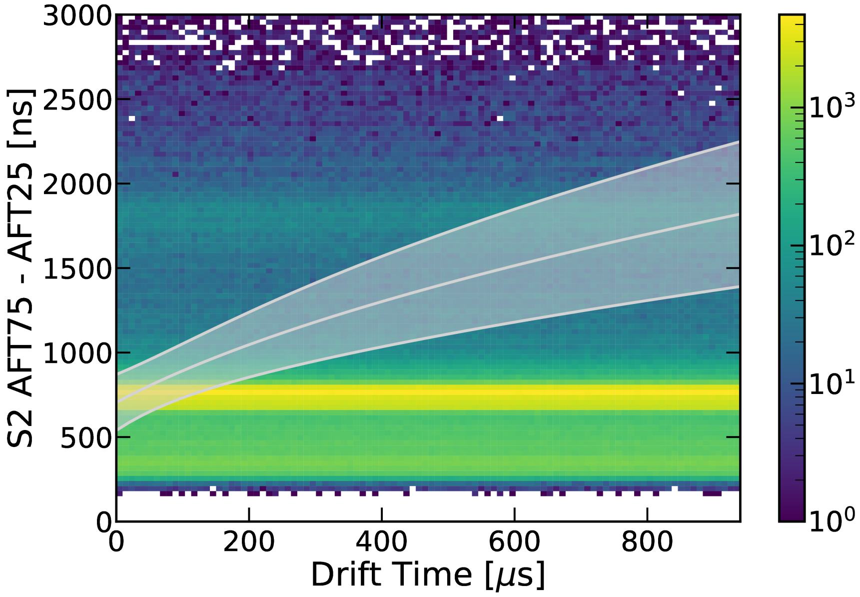


Figure 4.23: The width vs drift time of ACS events generated for analysis of the first science run of LZ. The mean S2 pulse width relationship from data, as well as the 3σ bounds for an S2 pulse, $n_e = 100\,\mathrm{e}$, are overlain.

The following methodology can be used to measure the rejection power of the S2 pulse width cut on accidentals using ACS data:

1. ACS events failing any data quality cuts excluding the S2 pulse width cut are removed from data.

2. For a given drift field, v_d and D_L are interpolated from Figure 4.20. These are used to derive M, calculate a maximum drift time T, and scale the value

of σ_M.

3. ACS data is assigned a random drift time between 0 and T.

4. The width of ACS events is scaled from W to W_{new} through the expression

$$W_{new} = \sqrt{\frac{M_{new}}{M}(W^2 - C^2) + C^2} \tag{4.21}$$

5. ACS events are passed through the S2 pulse width cut, as defined in Equations 4.13 and 4.18, with the adjusted values of M and σ_M.

6. The fraction of events passing the S2 width cut with these parameters is returned as an acceptance fraction of the cut at this drift field.

Following this process for varying field strengths produces the distribution in Figure 4.24. The rejection power of a cut in S2 pulse width increases with decreasing drift field, asymptotically approaching 100%.

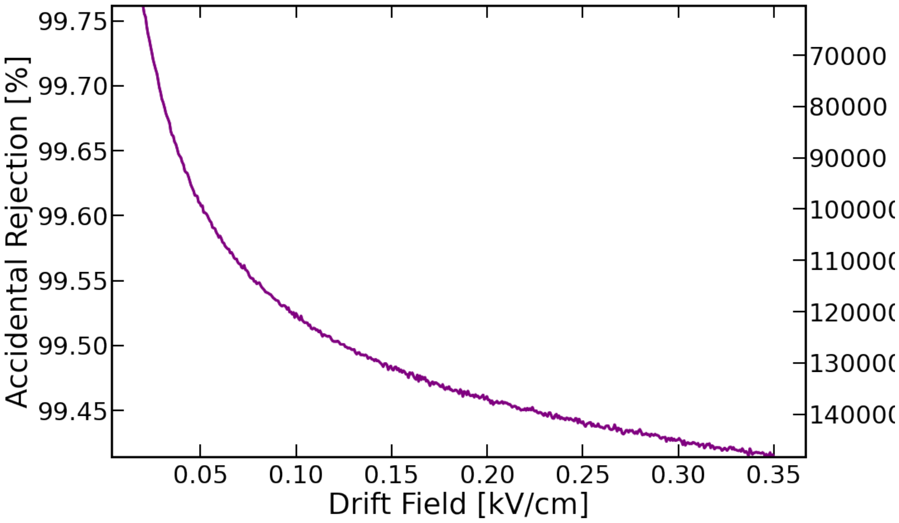


Figure 4.24: S2 pulse width cut rejection power vs drift field, measured using ACS data. Rejection power increases with decreasing drift field.

Accidental events take on S2 pulse widths within the boundaries defined by Equation 4.13, with uncorrelated drift time. The rejection power of a cut is therefore inversely proportional to fractional S2 pulse width resolution, $\sigma_{W_{S2}}/W_{S2}$. $\sigma_{W_{S2}}$, defined in Equation 4.18, contains three components. Only one of these, σ_M, changes with drift field. As a result, $\sigma_{W_{S2}}/W_{S2}$ decreases with decreasing drift field, and rejection power increases.

This is illustrated in Figure 4.25. At a drift field of 0.01 kV/cm, the shaded volume in pulse width and drift time covers 72% of the plot, indicative of the fraction of accidental events rejected by a cut in pulse width. At a drift field of 3 kV/cm, this rejection power decreases to 64%.

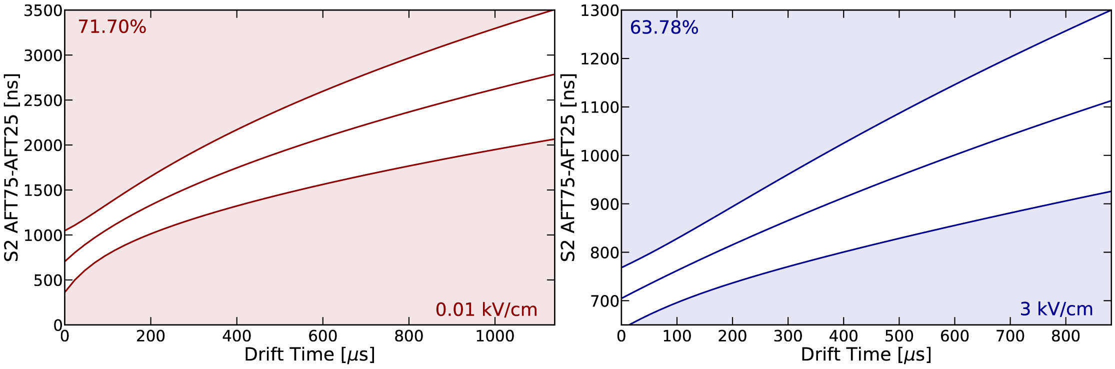


Figure 4.25: Illustration of the rejection power of the S2 width cut in varying fields, $n_e = 100$. Left: Projected bounds of a cut in S2 pulse width at a drift field of 0.01 kV/cm. The shaded region covers 71.7% of the plot, indicative of the fraction of accidental events rejected by the cut. Right: The same bounds at a drift field of 3 kV/cm. The shaded region covers just 63.78% of the plot, indicating that rejection power is lower at higher drift fields.

4.4.5 Projecting Accidental Rates Post-Cuts

Figure 4.21 demonstrates that the rate of accidental events increases with lower drift fields, due to higher drift times. Figure 4.24 demonstrates that the rejection power of data quality cuts will also increase with decreasing drift field, due to improved pulse shape discrimination. The ideal drift field for a low-mass WIMP search in LZ is therefore a compromise, where accidental events are minimised while maintaining high rejection power.

The accidental event rate after analysis cuts can be calculated by multiplying the accidentals rate in Figure 4.21 with the accidentals acceptance in Figure 4.24. Figure 4.26 shows the two distributions together for reference, and their product, the accidental rate expected to survive cuts. Accidental rate after analysis cuts decreases linearly with decreasing drift field, towards a minimum of between 0.05 kV/cm and 0.1 kV/cm. Below this point, the rate increases asymptotically with decreasing field. At this point, ER leakage increases considerably for both high-mass and low-mass WIMP searches [148].

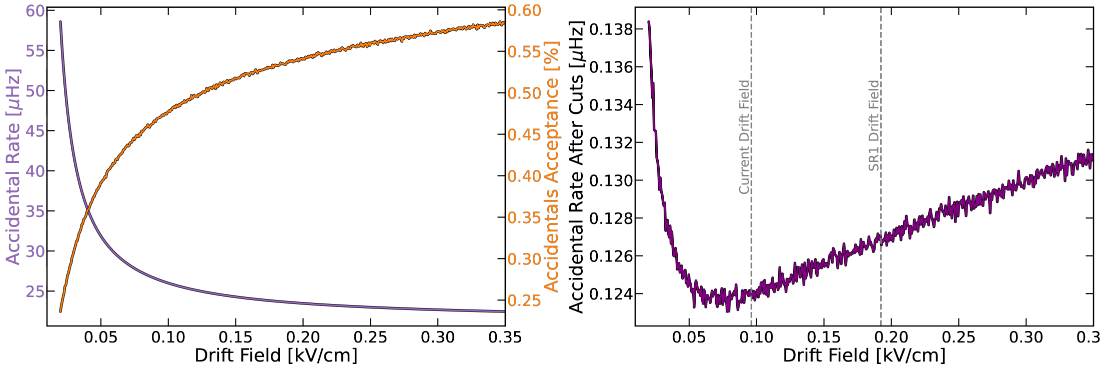


Figure 4.26: Left: Interpolated Accidentals Rate and S2 Width Rejection Efficiency Overlain against Drift Field Strength. Right: The product of these two distributions to give Accidentals Rate Post-Cuts. Vertical grey lines denote the Drift Field for the first science run of LZ and in 2024.

The operating field of LZ in 2024 is 97 V/cm, near to the drift field with a minimum projected accidentals rate. While this lowers sensitivity to high-mass WIMPs, it places LZ in an opportune position for low-mass WIMP searches.

5

Measurements of S2 Mis-classification

During analysis, S2 pulses have a non-zero probability of being mis-classified. This lowers the sensitivity of LZ, particularly to low-mass WIMPs. For example, an S2 of interest may be mis-classified as multiple S2s and be rejected from analysis, or may have a portion of S2 pulse mis-classified as separate, resulting in a lower reconstructed event energy.

This chapter presents efficiency measurements of the LUX-ZEPLIN pulse and event classification software LZap. Calibration data is used to identify candidate mis-classified events, which are used to measure for the first time S2 mis-classification and its dependence on event energy and position. This is substantiated using simulation data, extrapolating these probabilities to energies below that possible using calibrations alone.

Candidate events from simulations and low-energy calibration sources can be used to tune LZap for low-mass WIMP searches, minimising the probability of mis-classification of small S2s. Where these effects are irreducible, mis-classification measurements can be inserted as nuisance parameters into PLR analysis, to minimise sensitivity uncertainty.

5.1 Pulse Classification in LZ

5.1.1 Overview

The LZ DAQ samples data at a rate of 100 MHz. When a first-level trigger is activated, this data is recorded as a Pulse-Only Digitisation (POD) 2 ms before and 2.5 ms after the trigger timestamp [149].

The waveform is then passed through the LZ analysis package (LZap). LZap first applies a suite of shape-based filters to data, in order to classify sections of waveform as individual pulses. These pulses are classified using the HADES decision tree as an S1, S2, single photo-electron (SPE), multi photo-electron (MPE), single electron (SE), or other pulse [150]. LZap then derives properties of interest in a pulse (reduced quantities, or RQs) such as amplitude and pulse area. An event classifier determines whether an event contains a single scatter (SS), a multiple scatter (MS), or an other interaction. This process is based on classification software used in LUX [151].

This process is illustrated in Figure 5.1, where the LZ DAQ has triggered on an S2. The recorded waveform has been divided by LZap into a series of pulses. The presence of a single S1 and S2 results in the event being classified as a single scatter.

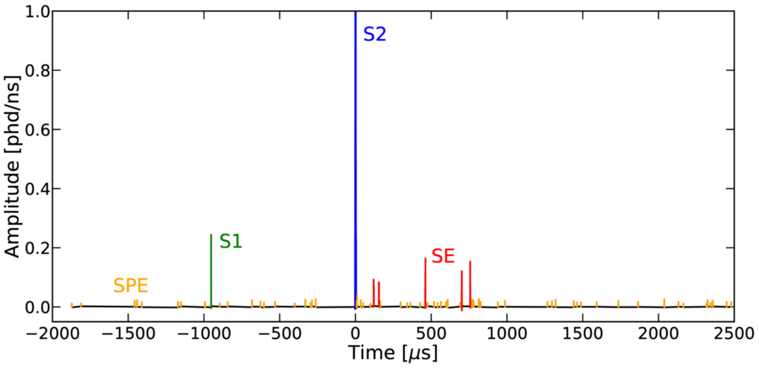


Figure 5.1: A recorded LZ waveform. Sections of waveform have been classified as S1 (green), S2 (blue), SPE (orange) and SE (red) pulses by the pulse and event classification software LZap.

LZap has been designed for the classification of single scatter S2s within the WIMP ROI, which take on a Gaussian pulse shape. However, non-Gaussian waveforms can result in the true beginning and end of an S2 pulse being incorrectly defined. At high energies, this leads to a source of MS inefficiency. At low energies, this leads to a source of SS inefficiency, as well as a dominant source of S2 pulse area uncertainty.

This chapter establishes with precision for the first time the probability of these mis-classifications in LZ calibration data. These measurements are then substantiated with simulation data for small S2s. Candidate mis-classified events can be used to tune LZap, such that mis-classification has minimal impact on a given analysis. Where mis-classification is irreducible, probabilities are integrated into the LZ PLR as nuisance parameters, to minimise sensitivity uncertainty.

5.1.2 Pulse Identification

Pulse Positions

In order to identify the nature of the PODs recorded within an event window, LZap first passes these through a series of filters. These filters identify sections of waveform which are likely the product of scintillation and electroluminescence.

The probability of scintillation is quantified through an exponential filter, containing the known time constant of LXe recombination. A pulse is classified as an S1, as opposed to an SPE or MPE, if a signal is observed in three or more PMTs within a 100 ns time window. As the shape of an S1 pulse is consistent across the TPC, mis-classification of S1s is minimal.

An S2 pulse, however, changes in shape depending on TPC position, as discussed in Section 4.3.2. This variation requires a more complex filter, which increases the probability of mis-classification.

The probability of electroluminescence is inferred using a Difference of Gaussians

(DoG) filter. A waveform $W(t)$ is convolved with Gaussians of standard deviations σ_1 and σ_2 as

$$\Gamma_{\sigma_1,\sigma_2}(t) = W(t) * \frac{1}{2\pi}(\frac{1}{\sigma_1}e^{\frac{-t^2}{2\sigma_1^2}} - \frac{1}{\sigma_2}e^{\frac{-t^2}{2\sigma_2^2}}). \tag{5.1}$$

The value of $\Gamma(t)$ is highest for a Gaussian-like pulse of standard deviation $\sigma_1 < \sigma < \sigma_2$, preventing the classification of waveforms too narrow or broad as S2 pulses in analysis. The choice of σ_1 and σ_2 is motivated using the maximum and minimum widths of S2 pulses observed in data, derived during the fitting process outlined in Section 4.3.2.

The DoG filter is illustrated in Figure 5.2 for a toy model S2 waveform. The filtered output contains a distinct peak at the centre of the event. This section of waveform can therefore be classified as an S2 pulse.

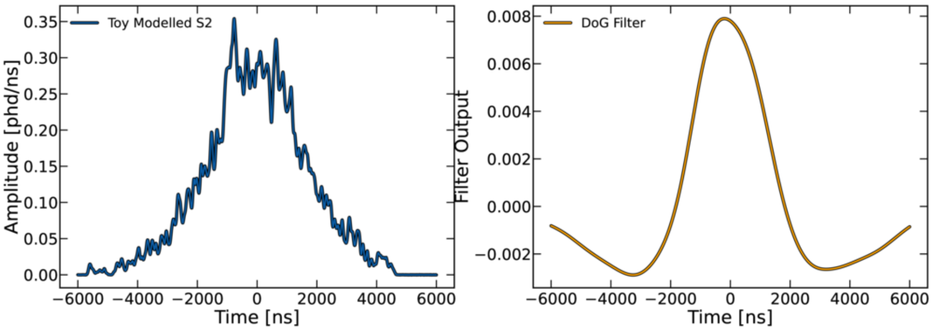


Figure 5.2: (Left) A toy-model S2-like waveform. (Right) The waveform passed through Equation 5.1. The peak of this output indicates that this section of waveform is an S2 pulse.

Pulse Boundaries

Once the position of expected pulses is identified, the boundaries of these pulses can be ascertained. A bounding box is placed around an S2-like waveform through the application of a skewed moving average filter:

$$S(i) = \frac{1}{S_l + S_r + 1}\Sigma_{j=i-S_l}^{i+S_r} W(j). \tag{5.2}$$

The tangent of this waveform $Mi + C$ can then be derived for each value of $S(i)$. As sections of waveform not containing a pulse have a gradient approaching zero, a condition can be placed on the boundary of a pulse that

$$M \times (S_l + S_r + 1) > T, \tag{5.3}$$

where T is a tuneable threshold.

S2 pulses with long tails will reach a gradient below threshold before the end of an S2. To resolve this, a variable $l(M, C; i)$ is defined as

$$l(M, C; i) = \Sigma_{j=i-S_l}^{i+S_r}(M \times j + C - S(j))^2. \tag{5.4}$$

The boundary of a pulse is then defined as the final sample satisfying either the condition expressed in Equation 5.3 or

$$W(i) - min(l(M, C; i)) > T. \tag{5.5}$$

This definition of S2 pulse boundaries is illustrated in Figure 5.3.

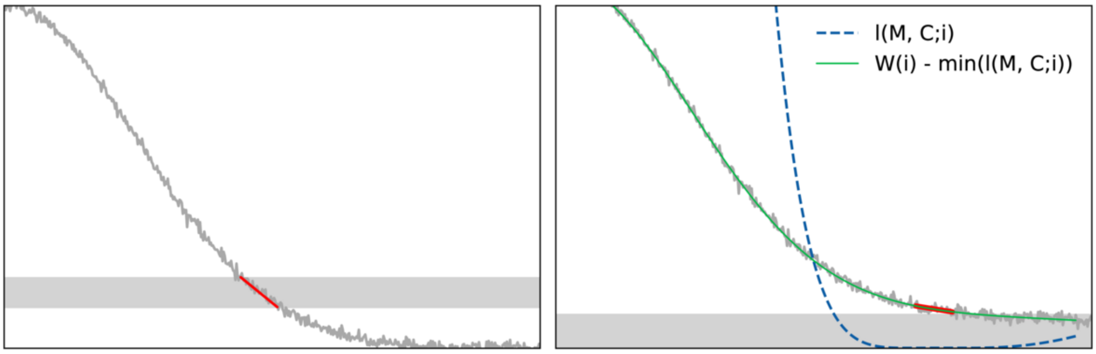


Figure 5.3: An illustration of S2 pulse boundaries in LZap. Left: A simple Gaussian waveform. The annotated tangent of the waveform (red) is above a threshold, defining the pulse boundary. Right: A waveform with an extended tail. The tangent (red) is below threshold. However, the value $W(i) - min(l(M, C; i))$ reaches threshold at this position. $l(M, C; i)$ is zero at this point, but is greater than zero for other samples.

Pulse Merging

Occasionally, the pulse boundaries of an S2 are defined such that some charge is classified as a separate pulse. These are combined with the main S2 using a separate filter, demonstrated to be effective during analysis of data from XENON10 [152].

A narrow moving average filter is first passed over the original waveform:

$$A_i^1 = \frac{1}{N_1 + 1} \Sigma_{m=i-N_1/2}^{i+N_1/2} W(t). \tag{5.6}$$

This filter is designed to identify sections of waveform which are more likely to originate from a single electron (SE) than scintillation (S1). As such, the width of this filter, N_1, is tuned such that it resembles the average S1 pulse width, $\mathcal{O}(10)$ samples. A second filter is then applied, taking in the output of the first filter:

$$A_i^2 = \frac{1}{N_2 + 1} \Sigma_{m=i-N_2/2}^{i+N_2/2} W(t) - max(A_j^1), j\epsilon[i - \frac{N_2}{2}, i + \frac{N_2}{2}]. \tag{5.7}$$

Here, N_2 is tuned to resemble the average SE pulse width, $\mathcal{O}(60)$ samples.

This filter is close to $W(i)$ when charge is continuous, and returns lower values when charge is clustered similarly to S1 pulses. A ratio of $W(i)$ and A_i^2 can therefore be used to identify SE-like waveform and merge these pulses into a primary S2.

Pulse merging is illustrated in Figure 5.4. The boundaries for this S2 do not encompass the entire pulse. However, the filter described in Equations 5.6 and 5.7 is continuous, and $W(i) \approx A_i^2$ at the boundary between pulses. The S2 and its tail would therefore be merged into a single pulse.

5.1.3 S2 Prominence

Many pulses are classified in an event window as S2s. However, some of these pulses should not be considered when evaluating the nature of an event. Examples of this can be seen in Figure 5.5. While some of these S2s indicate that an event is a single

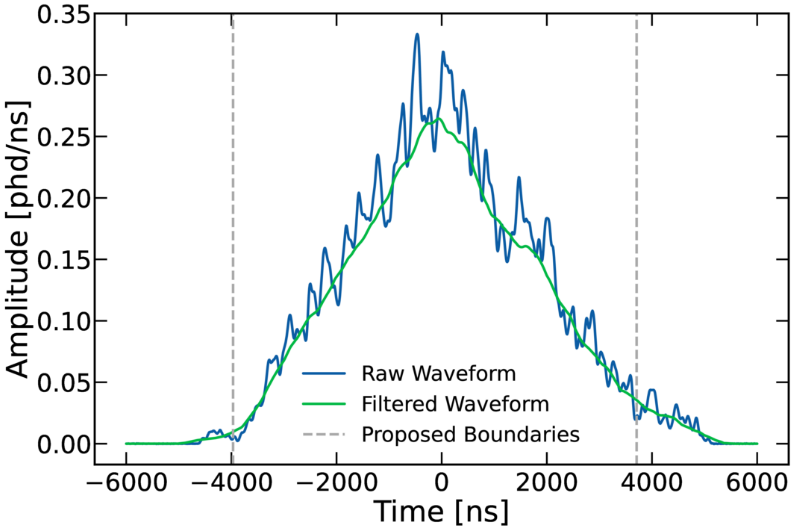


Figure 5.4: An illustration of pulse merging in LZap. A toy model S2 (blue) has boundaries as defined by Section 5.1.2 (grey dashed). The filter described in Section 5.1.2 is overlain in green. The lower bound, at −4000 ns, is appropriately defined. The upper bound, at 3900 ns is not, but the main S2 and its tail would be selected to be recombined in later stages of analysis using the filtered waveform.

scatter or multiple scatter, other S2 pulses should not be taken into consideration.

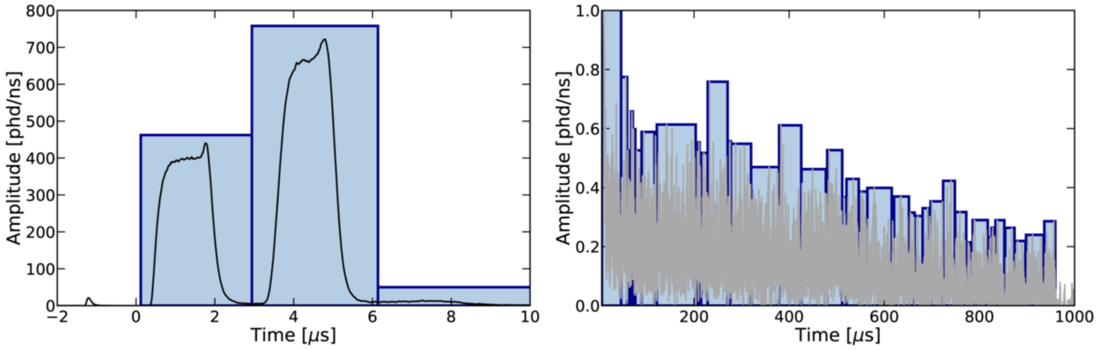


Figure 5.5: Left: A multiple scatter event. S2 pulses are highlighted in blue. Two pulses should be considered when evaluating the event. The third, from photoionisation of the gate array, should not. Right: An E-Train following a single scatter S2 pulse. Delayed emission is classified as many separate S2 pulses, which should not be considered when evaluating the event.

LZap classifies the largest S1 and S2 in an event window as ’prominent’. Other S1 and S2 pulses are then compared to these, to evaluate if they are also prominent. An event is then classified as a single scatter or multiple scatter depending on the number of prominent S1s and S2s.

The prominence of smaller S2s in an event is evaluated using a decision tree. S2s are generally considered prominent if they are of similar size to the main S2 and have a high amplitude in comparison to their pulse width. These elements of the decision tree are illustrated in Figure 5.6.

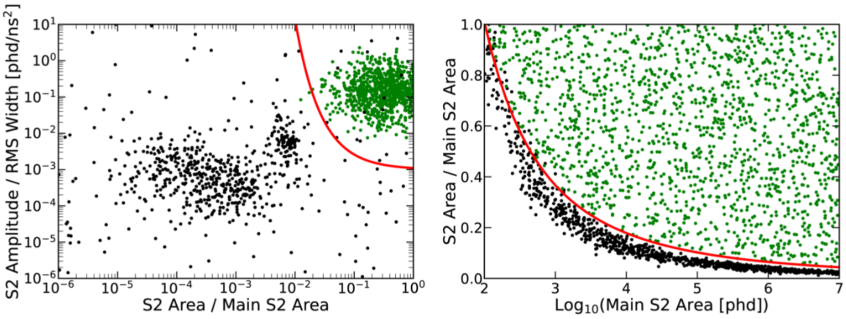


Figure 5.6: Illustration of elements of the S2 prominence decision tree implemented in LZap. Prominent S2s are highlighted in green. Left: Relative S2 pulse area against S2 pulse amplitude/width. S2s pass this check if they are of similar size to the main S2 and are of Gaussian pulse shape. Right: S2 pulse area against relative S2 pulse area. The range of S2s considered prominent increases for events containing a larger S2 pulse.

5.1.4 Mis-Classification

The boundaries of an S2 pulse can be mis-identified in three distinct ways, illustrated in Figure 5.7. In the first, the boundaries of an S2 are set such that a portion of charge has been classified as a separate pulse, which is not merged in. This 'fragment' of charge is not identified as prominent: however, it still results in a loss of S2 pulse area, resulting in a mis-reconstruction of energy. The second, where a pulse is 'split' into two evenly-sized pulses, is less common. However, these similar pulses are both identified as prominent, resulting in a fraction of SS events being incorrectly rejected from analysis. When LZap is tuned to mitigate these mis-classifications, the third begins to arise: S2s in multiple scatter events can be merged into a single S2, resulting in an efficiency loss of MS events and a high-energy SS background.

There is a trade-off between tuning LZap to mitigate pulse splitting and merging

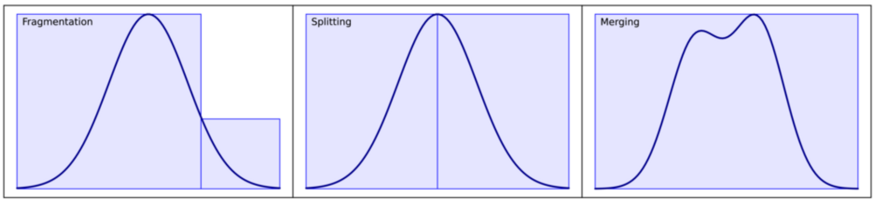


Figure 5.7: Illustrations of the three anomalies present in LZap data: the loss of charge in an S2 (fragmentation), the mis-classification of a SS as a MS (splitting), and the mis-classification of a MS as a SS (merging).

in analysis: reducing the number of S2s split or fragmented will increase the likelihood of large MS S2s merging. LZap can therefore be tuned for a given analysis, balancing the probability of mis-classifications such that sensitivity is maximised.

Splitting and Fragmentation

The DoG filter is effective at classifying S2s, with pulse merging filters after this proving capable of combining fragments of charge into the pulse. However, there remains a non-zero probability of a single scatter S2 being identified as multiple pulses. As these S2s appear almost identical to two multiple scatter S2s, it is impossible to correctly identify these pulses without drastically increasing the rate of S2 pulse merging.

This concept can be illustrated using toy S2s. Toy waveforms are passed through a Kolmogorov-Smirnov (KS) test [153] to evaluate their similarity to a Gaussian. A p-value of 0.05 is indicative that hypothesis has been rejected. As such, waveforms with a $p < 0.05$ can be said to be non-Gaussian, and are therefore likely to be classified as multiple pulses in LZap. The method used to generate toy model S2s in this section is discussed in Appendix B.

A set of toy model waveforms with a p-value greater than 0.95 can be seen in Figure 5.8. As waveforms closely resemble a Gaussian, these waveforms are very likely to be classified as a single S2 in LZap. Waveforms continue to exhibit

Gaussian pulse shapes until $p = 0.05$.

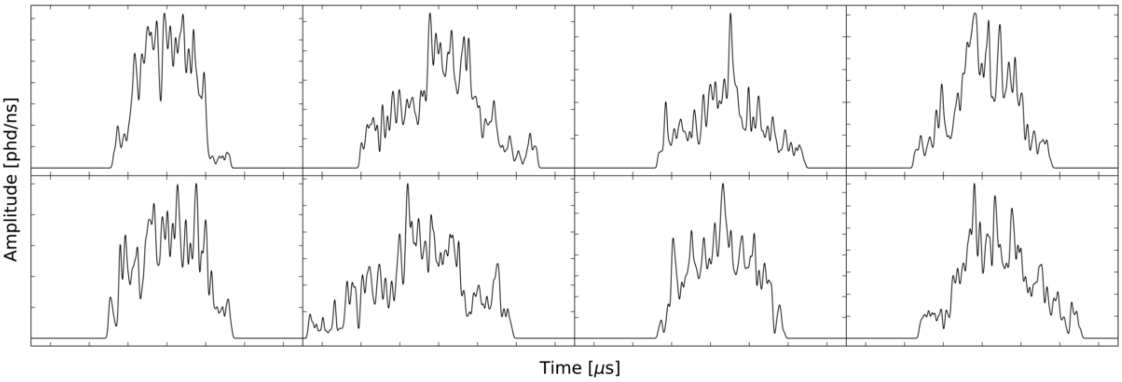


Figure 5.8: Toy model S2s produced with a KS test p-value > 0.95 compared to a normal distribution. Waveforms would be classified as a single S2 pulse in LZap.

A set of toy model waveforms with a p-value below 0.05 can be seen in Figure 5.9. As waveforms do not resemble Gaussians, many would be classified as multiple S2 pulses in LZap, with the probability dependent on the tuning of pulse merging filters and the ratio of splits to fragments depending on the S2 prominence decision tree.

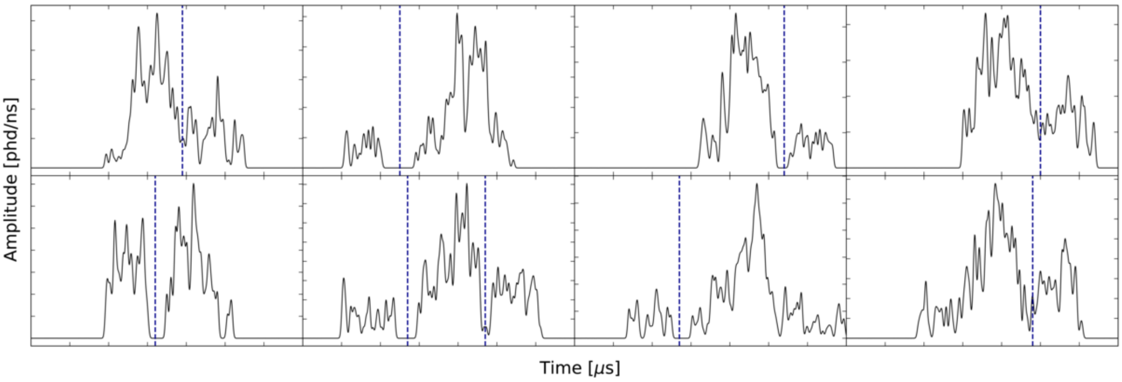


Figure 5.9: Toy model S2s produced with a KS test p-value < 0.05 compared to a normal distribution. Waveforms are very likely to be classified as multiple pulses in LZap. The boundaries of these pulses are illustrated as vertical lines (blue).

Only a small fraction of non-Gaussian S2 waveforms will fragment and split. This fraction can only be quantified using calibration and simulation data. However, the number of non-Gaussian waveforms provides insight into the relationship between mis-classification and an event's energy and position. If these relationships are observed in data, it can be concluded that candidate events are most likely mis-classified, and that LZap has been correctly tuned.

The toy model used to produce S2 waveforms seen in Figures 5.8 and 5.9 was used to generate events of varying pulse area, drift time, and SE width. The fraction of events of a KS test p-value < 0.05 can be seen in Figure 5.10. Non-Gaussian waveforms are observed to be most common in smaller S2 pulses at higher drift times, as expected. A peak of 80% is observed at $n_e = 5$, $\approx 200\,\mathrm{phd}$. The probability of non-Gaussian pulses, and therefore split and fragmented S2s, is low within the S1-S2 WIMP ROI, but some mis-classification will be observed at threshold. For a fixed pulse area and drift time, non-Gaussian pulses become more probable with decreasing SE pulse width, as the probability of two electron waveforms not overlapping increases. As grid deflection causes SE pulses towards the TPC centre to become narrower, pulse splitting and fragmentation is expected to be greatest at the TPC centre, decreasing with increasing radius.

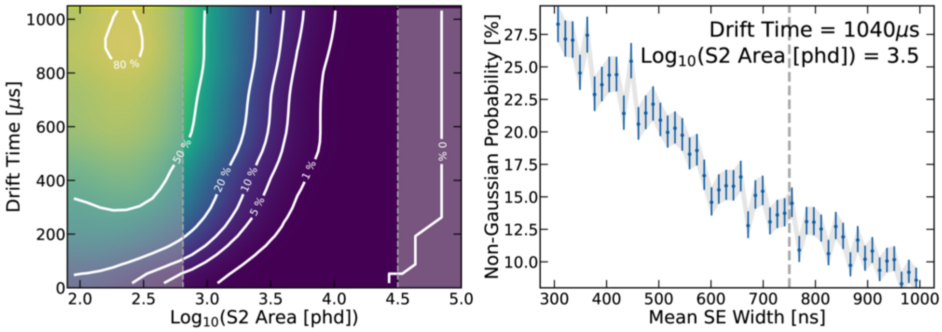


Figure 5.10: Probability of toy model S2s exhibiting non-Gaussian shapes ($p < 0.05$) against pulse area and TPC position. Left: Probability against S2 pulse area and drift time. Grey shaded regions are outside of the S1-S2 WIMP ROI. Smaller pulses at higher drift times are more likely to exhibit non-Gaussian waveforms, although probabilities in the WIMP ROI are low. Right: Probability against SE width, for fixed pulse area and drift time. A dashed line indicates the SE width, 750 ns, used in the left plot. Probability increases linearly with decreasing SE width, indicating that non-Gaussian pulses are more likely at the TPC centre.

S2 Merging

Multiple scatter events in LZ produce two or more S2 pulses, separated in (x, y, z). S2s with sufficient separation can be resolved. However, LZap classifies pulses using waveform time, which is only sensitive to (z). If separation in (z) is too low, regardless of (x, y) separation, two S2 pulses are classified as a single S2 pulse.

If (z) approaches zero, it is impossible for LZap to separate the S2s, regardless of tuning. At high (z), S2s can be easily resolved. Between these ranges, the probability of two S2s merging is dependent on the tuning of LZap merging filters.

The minimum separation in (z) at which two S2s can be separated can be derived using KS tests. Figure 5.11 shows the KS-test output for two Gaussian distributions of varying separation σ. At a separation 3σ, waveforms continue to look Gaussian-like, with $p = 1$. At this separation, two symmetric S2 pulses would be classified as a single S2 pulse using a DoG filter. Test values then drop exponentially with increasing separation, reaching $p < 0.05$ at 4.2σ. At a separation of $\sim 4\sigma$, two S2s would be successfully split by LZap in most instances using a DoG filter.

The value of σ depends on pulse area and drift time, as discussed in Section 4.3.2. The mean S2 pulse width variation in drift time, seen in Figure 4.13, can be scaled to σ using Equation 4.16. When scaled by the average drift velocity, this can be used to give the physical separation in (z) at which two S2s can be discriminated with confidence.

Figure 5.12 shows this calculation for varying drift times. Multiple scatter events with a physical separation greater than 5 mm can be resolved throughout the TPC. Below 1.5 mm, two S2s pulses are unlikely to be resolved by LZap regardless of TPC position. At high drift times, events of (z) separation below those listed in Figure 5.12 will still be resolvable, as the S2 pulse width will exceed σ_2 in the DoG filter.

For isotropic and homogeneous backgrounds and signals in the TPC bulk,

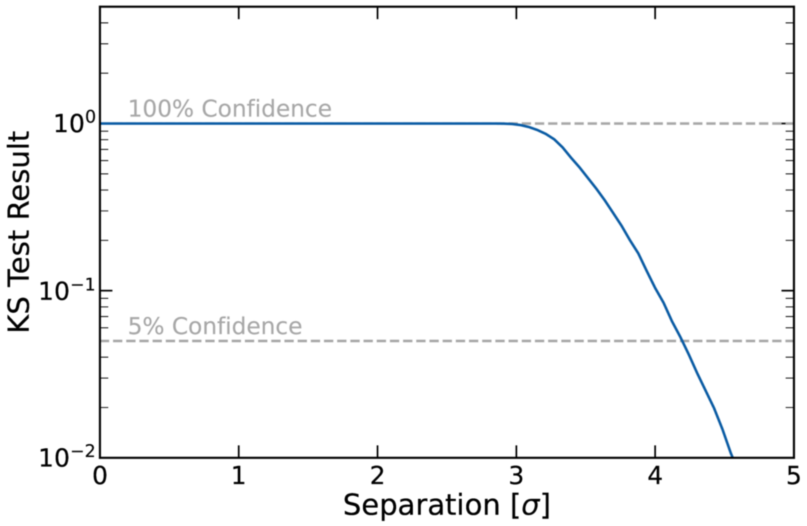


Figure 5.11: Gaussian KS test results of two Gaussians of separation σ. Two Gaussians are easily discriminable from one at separation 4σ, but cannot be easily discriminated at separation 3σ.

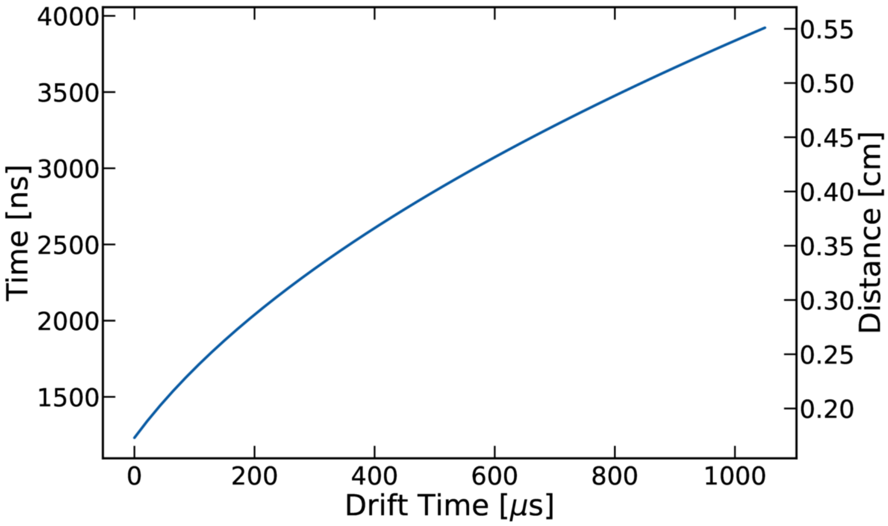


Figure 5.12: The MS separation threshold vs drift time. Below this separation in event time/physical distance, it is no longer possible to confidently resolve two S2 pulses.

merging is more probable at high drift times towards the TPC walls, where SE pulse width is greatest. However, the majority of MS backgrounds are produced by the TPC walls. An event from the walls which travels to the TPC bulk has generally lower separation in (z). As such, the majority of merged S2s in data are anticipated in the TPC bulk.

A measurement of S2 pulse merging is of interest for Multiply Interacting Massive Particle (MIMP) searches in LZ [154], as well as searches for $0\nu\beta\beta$ [155]. However, it is also of importance when tuning LZap for the rejection of thermal neutrons, a key NR background, and when searching for decays of ^{127}Xe, a low-energy calibration source of interest [156].

5.2 S2 Mis-classification From Calibration Data

5.2.1 Motivation

The probability of signal mis-classification can be quantified using LZ calibration data. A suite of cuts are applied to reject events which do not resemble signal. After this, additional cuts are applied to identify which remaining S2s appear to be split, fragmented, or merged. The fraction of these to signal-like events allows the mis-classification probability against event energy and position to be determined.

Multiple scatter events introduce a systematic error to measurements of splitting and fragmentation. Likewise, distorted single scatter S2s introduce uncertainty into measurements of S2 merging. While the latter is an irreducible error, the former can be mitigated by corroborating established probabilities with analysis of simulated data, known to originate from single scatters.

5.2.2 Selecting Candidate Events

In order to select signal-like events in data from which to identify mis-classified events, preliminary cuts must first be applied, to reject events which do not resemble NRs in the TPC bulk. However, the mis-classification of events occurs before the application of data quality cuts. The cuts applied to data must therefore be minimal, removing only populations which do not reflect signals or backgrounds in the TPC bulk. These criteria also allow for mis-classification to be understood outside of the fiducial volume and WIMP ROI, which has applications in alternative physics searches such as S2-Only analysis.

The following cuts are applied to data, before selecting mis-classified S2s.

Livetime Cuts: Several cuts in standard analysis remove periods of low data quality, such as around E-Trains, during hotspot emission, and whenever the outer detectors have been triggered.

Above-Anode Cut: Events from above the LZ anode dominate data at drift times below 70 µs. The nature of the cut used to remove these events in this analysis is outlined in Chapter 6.

S1 Tail Cut: A cut on broad S2 pulses near to large S1s rejects events produced from the photoionisation of the gate by high-energy scintillation.

Many-S2 Cut: Events with a high number of S2 pulses relative to the main S2 pulse area are rejected, in order to remove remaining periods of low data quality.

Below-Band Cut: In the context of of S2 pulse merging, where wall events dominate the list of candidates, events below the NR band are rejected.

Results

Approximately 40% of events are rejected using the outlined cuts, leaving a population of non-anomalous MS and SS events in LZ.

The distribution of these events for CH_3T data can be seen in Figure 5.13. At pulse areas below 1×10^4 phd, events are predominantly removed by the above-anode ('gas') cut and S2 TBA. Below 1×10^3 phd, events originate from mis-reconstructed α decay and iS1 pulses, and are rejected by the S1 tail cut. Above 1×10^4 phd, events are predominantly rejected by the skin and OD. Remaining events at low

S2 pulse area originate from the TPC grids, with a small population of low-energy accidental events. Within the WIMP ROI, data is predominantly standard ER and NR interactions, with data above the ROI comprised largely of MS wall events.

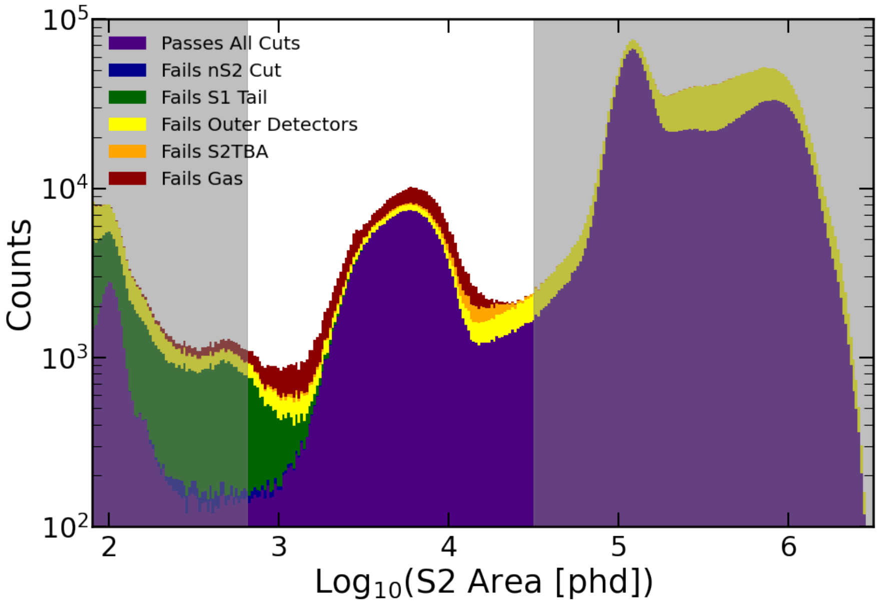


Figure 5.13: S2 pulse area of events in CH_3T calibration data remaining after analysis cuts, selected for measurements of pulse mis-classification. Shaded regions indicate the S1-S2 ROI, containing accidental events and high-energy wall backgrounds at low and high S2 respectively.

5.2.3 S2 Merging in Calibration Data

The first anomaly to measure is the merging of MS S2 pulses into one SS S2, as illustrated in Figure 5.7. For this work, cuts are applied to single scatter events to identify S2 pulses which are distorted, and appear more similar to two Gaussians than one. Remaining candidates can then be compared to MS events, to determine merging probability.

Measurements of MS S2 merging require a reliable MS calibration source. For this work, DD calibration data, shown in Table 3.1, are used to select merged candidates within the WIMP ROI. As events are produced by a generator at the top of the detector, the majority of events are distributed at low drift times. However, with $\mathcal{O}(1e6)$ events, the statistics are sufficient to demonstrate that S2 pulse merging is negligible in this region. At higher energies, the majority of MS events are produced by the TPC walls.

Merged S2 pulses take on an above-average pulse width due to the separation between S2s, (z). A list of merging candidates can be refined accepting events of only above-avergage pulse width, where the mean SS S2 width is defined in Chapter 4. Events of S2 pulse width $> 2\sigma$ for a given drift time and S2 pulse area, as defined in Equations 4.13 and 4.18, are selected as potential merged events. Candidate events are shown in Figure 5.14, in pulse width and drift time.

This selection of events contains many SS S2s. Many of these are standard S2s, due to the cut in pulse width being applied at 2σ. However, other pulses are subject to PMT saturation and wall effects, which also increase pulse width. Identifying merged S2s will therefore require cuts in pulse shape.

Previous studies searching for anomalous waveforms in LZ have found that cuts in AFT RQs disproportionately target merged S2 pulses [157, 158]. This is due to the fact that AFTs are correlated in a Gaussian waveform, and are not correlated in an asymmetric waveform produced by two S2s in close proximity. AFTs and

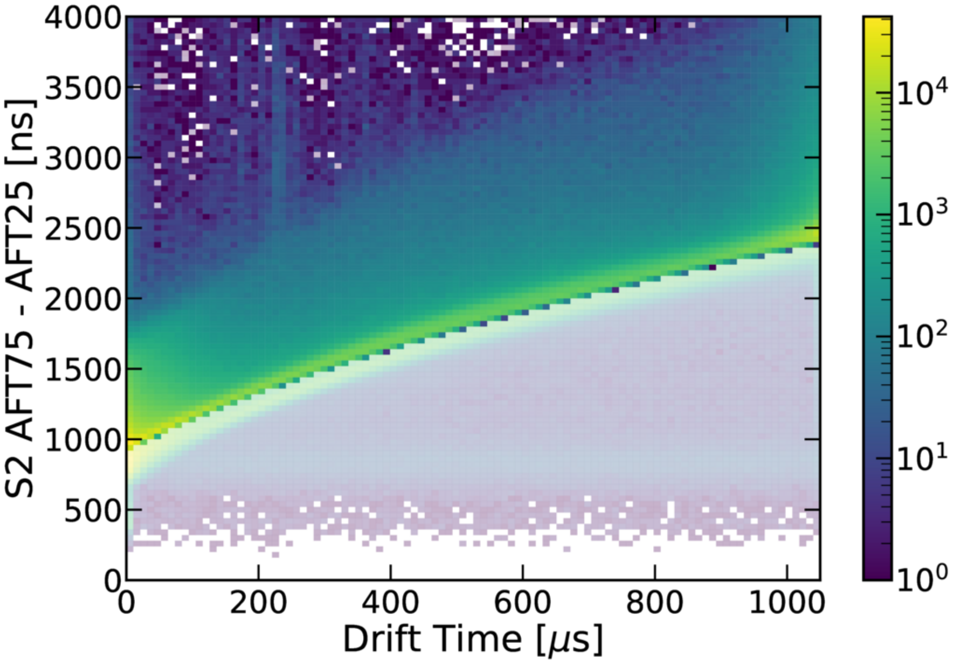


Figure 5.14: Events selected as potential merged S2s, shown in S2 pulse width vs drift time. Events not selected for this analysis are displayed at a lower opacity.

pulse width metrics can therefore be used in combination to reject SS S2s, leaving a dataset of merged S2 pulses.

This concept is illustrated in Figure 5.15. Events are plotted in two width metrics: AFT25-5 and AFT95-75. Most events follow a distinct linear relationship, indicating that an S2 is symmetric. Events that deviate from this relationship are asymmetric, and are more likely to be the product of pulse merging.

A cut in a single width metric will result in significant loss of merged S2s, as well as leakage from distorted SS S2s. As the width metrics of SS S2s each correlate with one another, a cut in multiple width metrics improves the efficiency of data selection. As merged S2s have not yet been identified in data, unsupervised machine learning is well-suited to this analysis. A Gaussian Mixture Model has been chosen to separate merged and un-merged populations in data. The nature of a GMM, as well as an illustrative example of GMM clustering, can be found in Appendix A.

A GMM was provided all candidate events, and sorted events into clusters, such

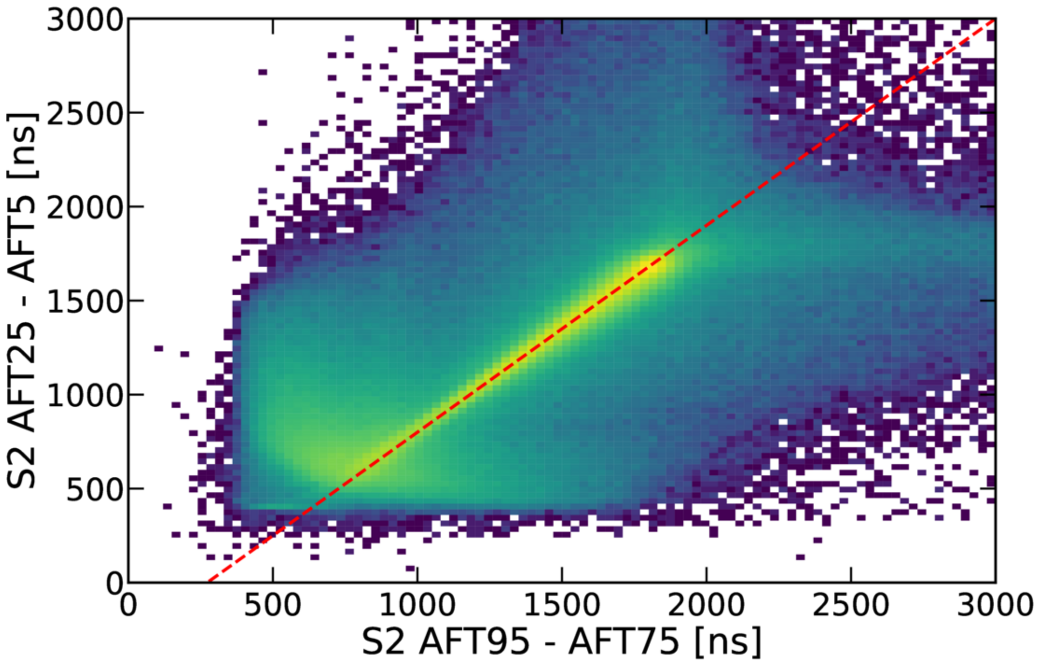


Figure 5.15: The S2 AFT95-AFT75 vs AFT25-AFT5 of S2s considered for a measurement of S2 pulse merging. A red dotted line indicates the relationship shown by a symmetric S2 pulse. Events deviating from this relationship are more likely to be merged S2 pulses.

that covariance was minimised. In order to fully separate known populations in data with minimal computation time, the GMM was trained to identify 25 clusters in data. Clusters containing asymmetric S2s were identified as merged candidates, while clusters exhibiting standard correlations in width metrics were identified as SS S2s. Several width metrics were chosen as input, based on Figure 5.15. These inputs, as well as the general region of an S2 pulse they are sensitive to, are outlined in Table 5.1.

Candidates identified as standard or distorted SS S2s are distributed in multiple RQs in Figure 5.16. The majority of events follow the expected relationship in SS S2 pulse width against drift time, and are closely correlated in multiple width metrics, as expected. The fraction of candidates to SS S2s is constant in the WIMP ROI, where the majority of events are standard S2s selected by a 2σ cut on pulse width. Above the ROI, this fraction increases due to PMT saturation. A small

Table 5.1: Inputs for a GMM designed to identify merged S2 waveforms, and the S2 property each input is sensitive to.

Input	Property
AFT75 - AFT25	S2 Peak
AFT95 - AFT5	S2 'Bulk'
AFT95 - AFT75	S2 Tail
AFT25 - AFT5	S2 Rise Time
RMS Width	'True' S2 Width
FWHM	S2 Peak

fraction of events in this data have an above-average width at low drift times, and show asymmetric pulse shapes. These S2s are produced from SS events near to the field cage resistors, around which S2s become distorted and iS2 pulse rate increases due to a higher field strength.

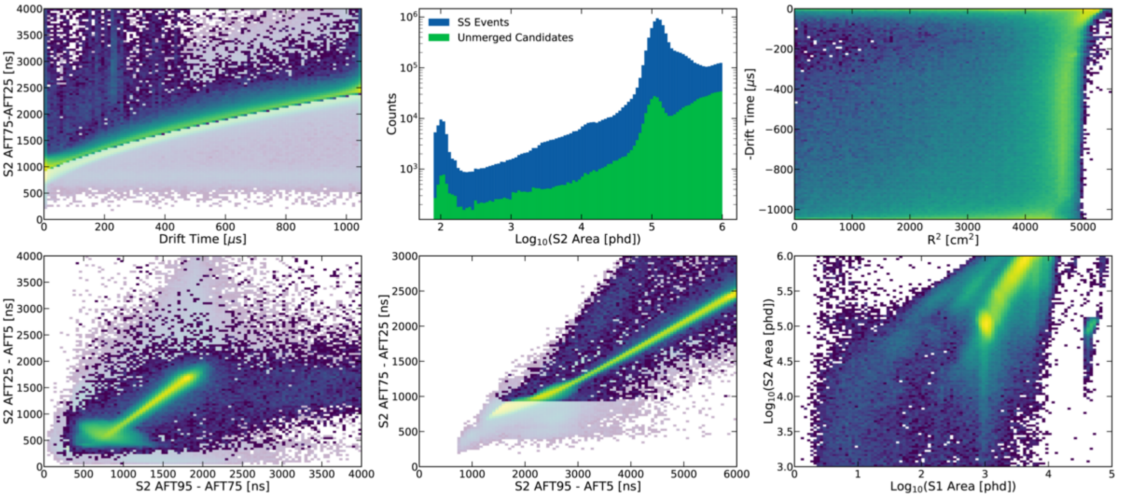


Figure 5.16: Candidate events identified as likely originating from distorted SS S2s, shown in metrics of width, size, skewness, and width correlation. Events generally display standard pulse widths for a given drift time, and are correlated in multiple width metrics. Events displaying an above-average width at low drift times originate from the field cage resistors.

The RQ relationships of candidates identified as merged MS S2s can be seen in Figure 5.17. Events are evenly distributed in pulse width, due to the variation (z) in MS S2 separation, and do not follow the same relationships as in Figure 5.16.

Example waveforms of S2s selected as merged pulses can be seen in Figure 5.18. Each waveform can be seen to closely resemble the combination of two Gaussian distributions, as expected. Events in the TPC are relatively symmetric, and are therefore clearly the combination of MS S2s. Towards the TPC wall, the majority of candidate waveforms become considerably more asymmetric. This suggests that while the waveforms of wall events appear to be the product of multiple clouds of ionisation, this is likely the result of the distortion of electrons from a SS S2, and not from the merging of MS S2s.

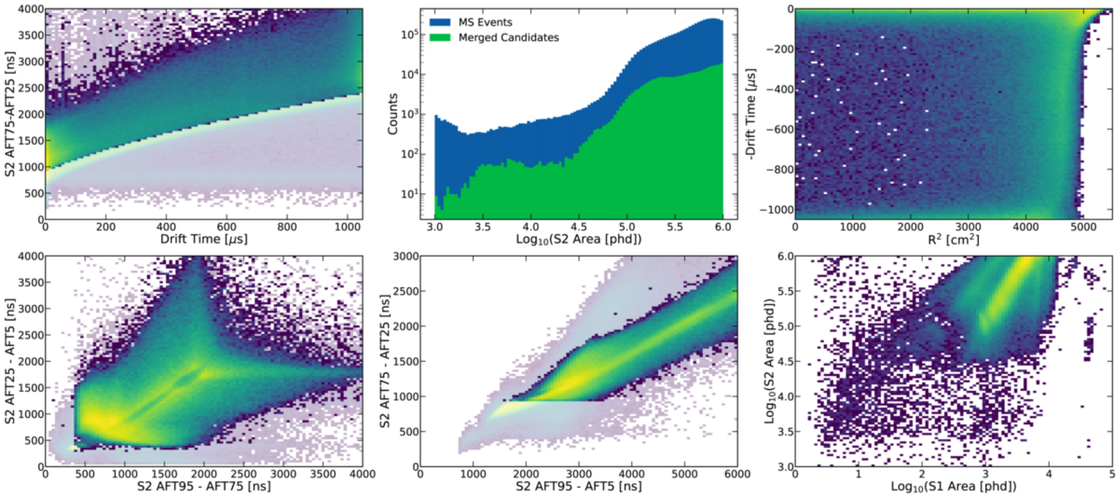


Figure 5.17: Candidate events identified as likely originating from MS S2 merging, shown in metrics of width, size, skewness, and width correlation. Events are more evenly distributed in pulse widths, and do not show the same correlations between pulse widths as observed in SS S2s.

A conservative measurement of S2 pulse merging can be made by assuming that no candidate events are the product of distorted SS S2s. The fraction of merged candidates can then be found for varying TPC drift time and radius. Pulse merging is not observed to change in probability with changing S2 pulse areas - however, events are divided into S2 pulse areas inside and outside of the WIMP ROI,

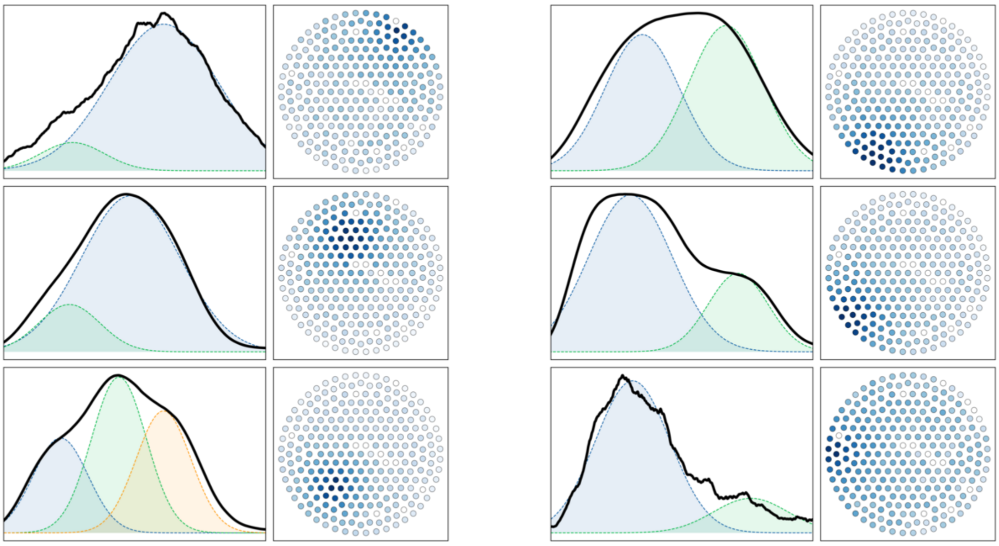


Figure 5.18: Examples of waveforms identified as potentially merged, alongside the PMT hitmap. For each waveform, several Gaussians are overlain, in order to illustrate that each is likely the sum of multiple electron clouds. Events from the TPC bulk are clearly produced by the pileup of MS S2s. Events nearer to the TPC wall are less symmetric, and are likely the product of distorted SS S2s.

as ROI events are predominantly from DD calibrations, and high-energy events predominantly from backgrounds in the TPC wall.

Figure 5.19 shows the merging probability of events inside and outside of the WIMP ROI against TPC position. 5% of candidate MS events in the ROI are identified as merged S2s. However, these events are disproportionately located towards the field cage resistors, where candidates are considerably more likely to be the product of distorted SS S2s. Outside of this volume, merging is measured to be consistently below 2%. Repeating this analysis using WIMP search data with an applied fiducial volume cut finds no candidate events in the WIMP ROI. Above the ROI, candidates are most probable towards the TPC wall, where they are most likely distorted SS S2s, and in the TPC bulk, where they are most likely a result of MS merging. In both volumes, the fraction of merged candidates is on the order of 10%.

Merged S2 pulses are negligible in calibration data of energies within the WIMP ROI, and no merged candidates have been found in WIMP search data. Above the

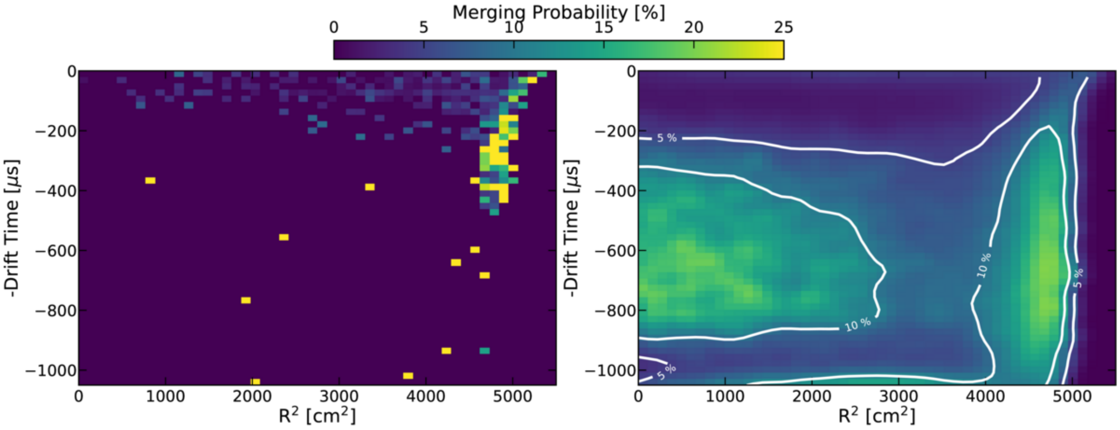


Figure 5.19: PDFs of S2 merging against TPC radius and drift time. Left: Events within the S1-S2 ROI. Merging probabilities are consistently below 2%, excluding a volume around the field cage reisistors, where candidates are produced by distorted SS S2s. Right: Events above the S1-S2 ROI. Events towards the TPC wall are the product of distorted SS S2s. However, events in the TPC bulk are produced by S2 merging. Merging probability does not exceed 12%.

ROI, around 12% of MS events are mis-classified as SS events in the TPC bulk. Tuning of LZap for high energy searches can minimise this fraction, using the results of this work as input.

5.2.4 S2 Fragmentation in Calibration Data

The second form of mis-classification to measure is the ’fragmentation’ of an S2, where a SS S2 is incorrectly identified as a prominent S2 and a non-prominent pulse - a ’fragment’ of charge, as illustrated in Figure 5.7. Two key metrics are required for SS analysis: the probability of fragmentation taking place, and the amount of charge lost in the process.

As CH_3T calibration data has been found to be reliable across most pulse areas and TPC positions, it has been selected for this analysis.

When an S2 fragments, a smaller pulse is observed immediately before or after the main S2 in an event. A series of cuts on the properties of this smaller pulse are detailed below, in order to select fragmented S2s while rejecting random coincidence of S2 pulses with TPC backgrounds.

Time Separation

As fragments of a primary S2 are from the same cloud of ionisation, the majority of fragmented pulses are immediately before or after the main S2 in an event, with zero separation in time. However, it is possible for small S2s to produce fragmented charge with some separation in time, as seen in Figure 5.9. A cut in the separation between the main S2 and neighbouring pulses can therefore be used to reject coincidence events.

The time separation of pulses after the main S2 in an event can be seen in Figure 5.20. SPE pulses immediately following an S2 form a peak at 0 ns. Above this, rate decreases, as a result of the pulse boundary algorithms used to define the beginning and end of an S2. This increases to a peak, the most frequent separation of pulse boundaries in LZap, before decreasing in a power law relationship. A small trough is observed in the range 300 ns to 400 ns, due to the pulse merging algorithms in LZap. The same general trend is observed for pulses before the main S2 in an event, with a lower overall rate, due to the spurious emission of photons and electrons following an S2 pulse.

For the analysis of fragmentation of S2s, events are selected if a pulse before or after the main S2 is separated by less than 2000 ns. This maximises acceptance of fragmented pulses from small S2s, while background leakage is mitigated through cuts in pulse size and shape.

Area Fraction

Fragmentation of an S2 results in the formation of an S2 pulse and a second smaller pulse. Pulses which are significantly smaller than the main S2 are more likely to originate from random coincidence, with actual small fragments having negligible impact on the reconstruction of an event. A cut can therefore be placed on the size of a fragment relative to the size of the main S2 in an event.

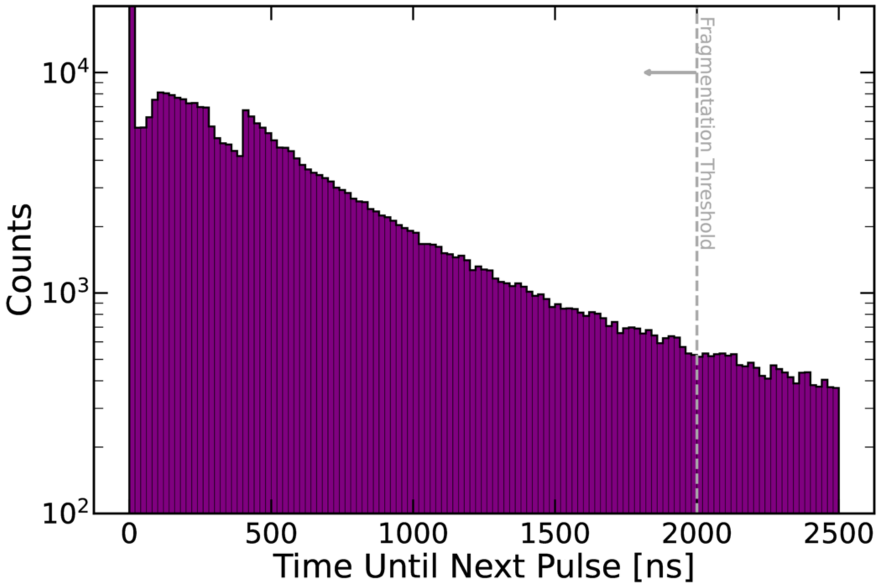


Figure 5.20: The time separation between an S2 and the following pulse. The threshold selected for fragmentation candidates is annotated.

The area fraction of fragments following the main S2 pulse in an event can be seen in Figure 5.21. Event rate is highest for small fractions, due to the delayed emission of photons and electrons following large S2s. Event rate quickly drops with increasing fractional area. However, two peaks are observed at 2% and 9%, produced by photoionisation of the gate array and glue ring.

When electrons begin to electroluminesce, some of the electroluminescence light strikes components of the TPC. This produces electrons, which then drift to the extraction region and produce a second 'echo' S2, as seen in Figure 5.5. An S2 echo is proportional to the size of the main S2, and the separation between the main S2 and its echo is dependent on the separation between the extraction region and the TPC component struck by electroluminescence light.

Figure 5.22 shows fragment pulses plotted in area fraction and separation from the main S2 in an event. The separation between the peak of the main S2 and a fragment is clustered to two values, 1500 ns and 2500 ns. These are the known drift times of S1-S2 events from the glue ring and gate, respectively. Photons

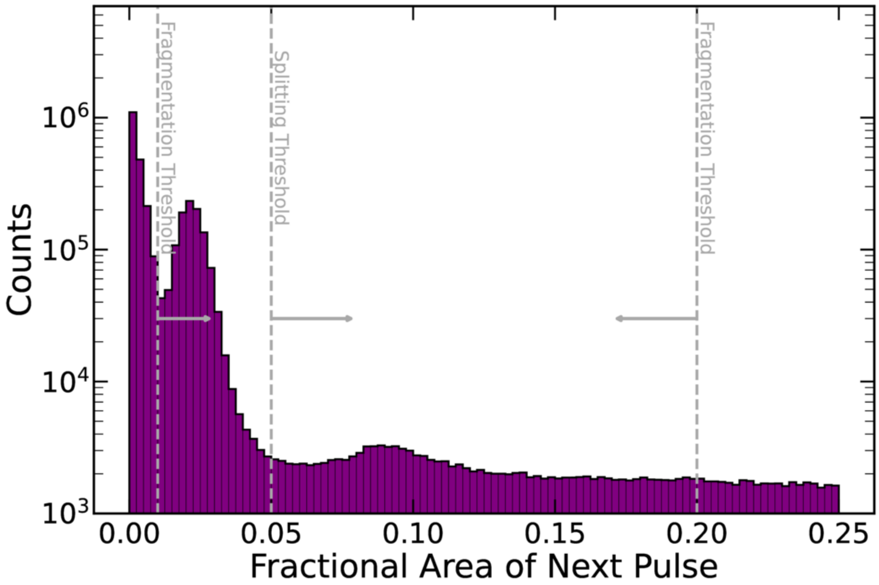


Figure 5.21: Fractional area of pulses following the main S2 in an event. The thresholds for splitting and fragmentation candidates are annotated in grey. Two peaks are observed in data at 2% and 9%, from photoionisation of the gate and glue ring.

are more likely to strike the glue ring, as it is nearer to the extraction region, resulting in a larger S2 echo. Figure 5.22 also shows a selection of events in (x, y) space. Events producing gate photoionsiation are evenly distributed, whereas events only photoionise the glue ring when produced directly above it, resulting in events forming a ring in (x, y).

Photoionisation is produced from electroluminescence light, rather than from ionisation at the event vertex. These pulses should therefore not be included as part of an S2, as this would result in the energy of an event being overestimated.

Candidate fragmented events are rejected if the area fraction of a fragment is below 1% or above 20%. This does not reject photoionisation of the gate and glue ring. However, these and multiple scatter events are rejected through a cut in pulse width.

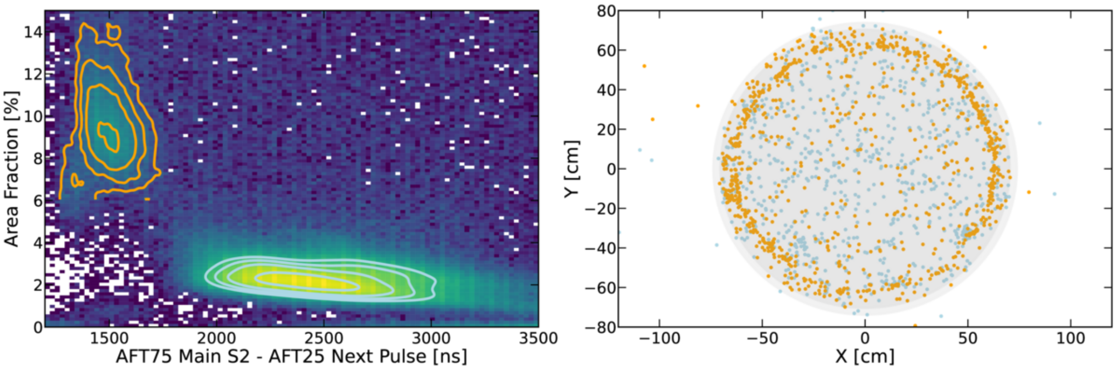


Figure 5.22: Distributions of fragments following an S2. Left: Peak-peak separation vs area fraction. Events are clustered at 1500 ns (orange) and 2500 ns (blue), indicative of photoionisation of the glue ring and gate. Right: XY position of a selection of events. Events photoionising the gate are evenly distributed, while events only photoionise the glue ring when directly above it.

Width Ratio

The waveform of a larger S2 that has been fragmented is continuous, such that the gradient of the waveform between the S2 and its fragment is constant. This can be derived using AFTs for the two pulses.

An illustration of this can be seen in Figure 5.23. Two S2 pulses overlap one another, but the total waveform has been correctly split. Each S2 is represented by a Gaussian distribution. The gap between the two Gaussians, derived as the distance from AFT75 of the first pulse to AFT25 of the second, is larger than the width of each individual Gaussian, derived as the AFT75-AFT25 of each pulse. The waveform can be rejected from analysis, as it is correctly classified. Where the gap between pulses is smaller than pulse widths, it is more likely that a waveform is an incorrectly split single scatter S2.

The distribution of the gap:width ratio of fragments following the main S2 in an event can be seen in Figure 5.24. The rate of events increases with increasing ratio, up to a peak at ~ 3 produced by MS S2s with low separation in (z). Below this peak, an event is less likely to be correctly classified.

In order to reject MS S2s, events of a gap:width ratio greater than 2 are rejected

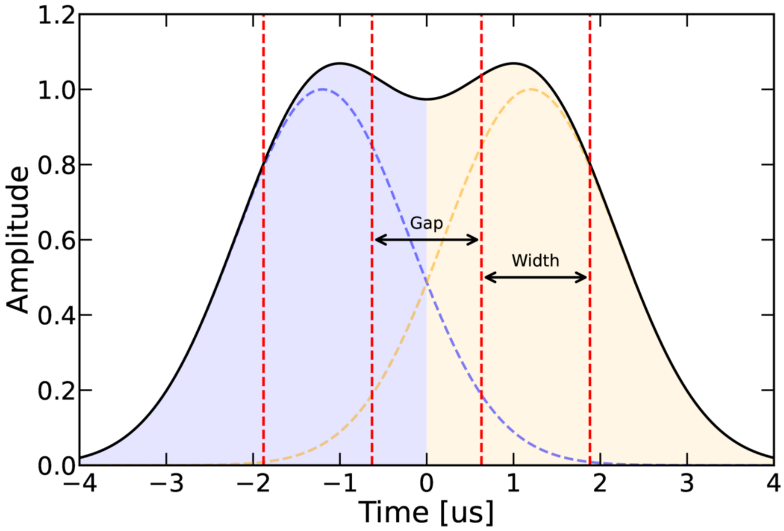


Figure 5.23: Illustration of two S2s with low separation in z. The AFT25 and AFT75 of each pulse is annotated. The gap between the first and second pulse is greater than the width of each individual pulse. The event can therefore be said to be correctly classified.

when selecting candidate fragmented S2s.

XY Separation

An S2 and its fragment both originate from the same electron cloud, are located nearby to one another in (x, y). In contrast, a MS event with low separation in (z) is more likely to have greater separation in (x, y). This can be used to reject MS events.

When split, an S2 should show a positional variation of less than 5 cm. However, position resolution decreases for small S2 pulses and fragments. Candidates are therefore selected with a positional variation below 20 cm.

Results

After the application of these cuts to data, 4668 candidate events are selected from an initial dataset of 3,412,983 events. Figure 5.25 shows the distribution of these candidates in S2 pulse area, relative to the number of events in the original dataset.

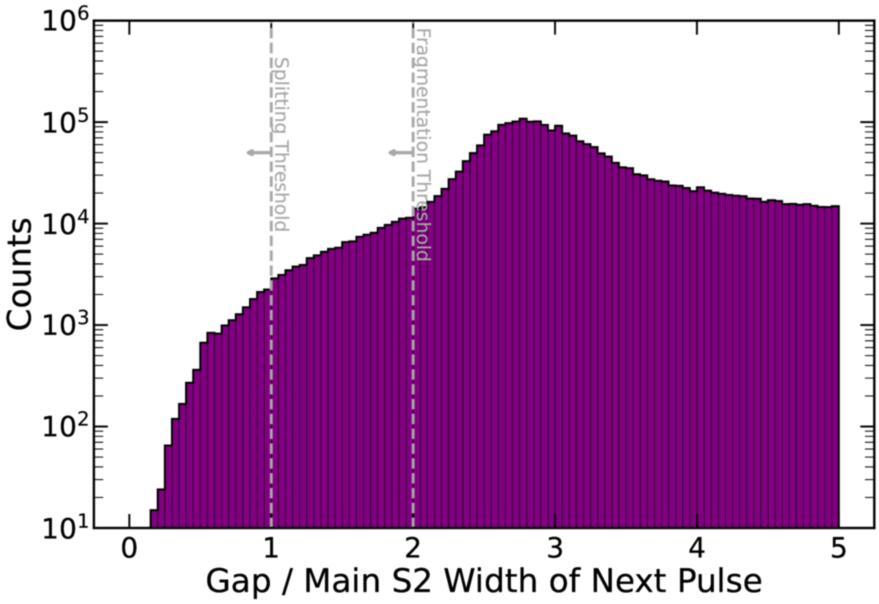


Figure 5.24: The AFT width fraction between the main S2 of an event and the pulse immediately following it. Thresholds for selecting fragmentation and splitting candidates are annotated in grey. A peak is observed at ~ 3, produced by MS S2s. Below this, events are more likely to be the product of pulse mis-classification.

Above the WIMP ROI, Fragmentation is $< 0.1\%$. Remaining events in this region contain S2s indicative of gate photoionisation, suggesting that fragmentation is less probable than this. Below the ROI, fragmentation becomes more probable, reaching a peak of $(6.5 \pm 1.3)\,\%$. Below this, fragmentation is observed to decrease: however, in this region, calibration data is subdominant to backgrounds. Candidate events identified as fragmented in this region are from backgrounds near to the field cage resistors.

The impact of this fragmentation on S2 pulse area resolution, a limiting factor on sensitivity in low-mass WIMP searches, can be evaluated by comparing the charge loss from fragmentation to other sources of S2 uncertainty. The area resolution of a single electron in LZ is approximately 20%. This introduces a baseline resolution to an S2 that scales with the number of electrons N as $\frac{1}{\sqrt{N}}$. Similarly, the electron lifetime is observed to be between 5 ms and 8 ms, resulting in the loss of up to 12-20% of S2 electrons during diffusion. While this is corrected for, the binomial

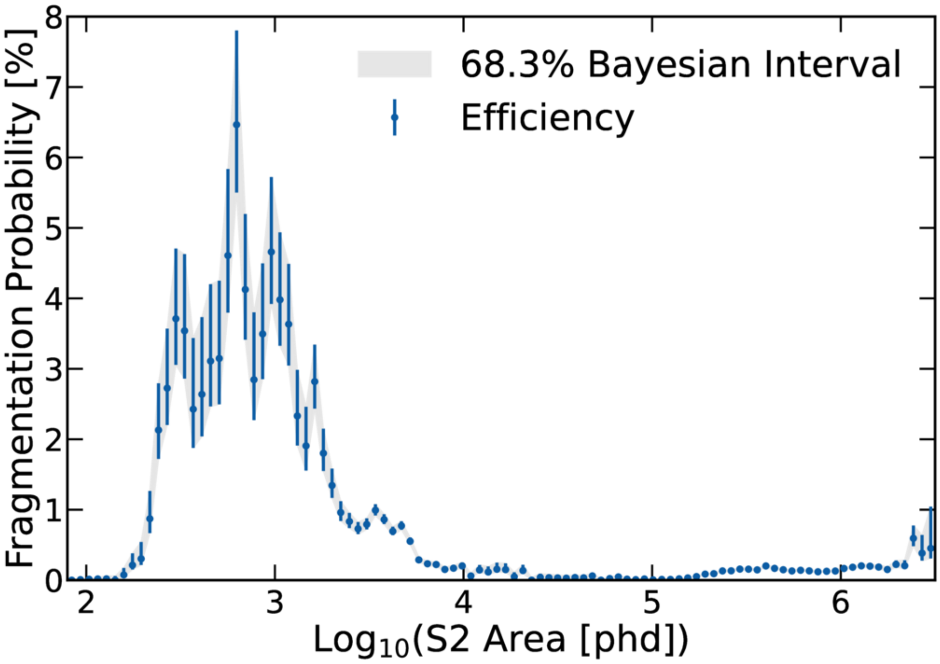


Figure 5.25: The probability of an S2 pulse fragmenting vs S2 area, derived using CH_3T calibration data. Fragmentation above the WIMP ROI is negligible. Towards the bottom of the ROI, fragmentation occurs in $\mathcal{O}(5\%)$ of calibration events.

distribution of electrons detected for a given S2 introduces another systematic uncertainty to corrected S2 pulse areas. This uncertainty on S2 pulse areas is $\sqrt{P(1-P)} = 40\%$ for small pulses at maximum drift time for the lowest observed electron lifetime, or 32% for the highest lifetime, and also scales with $\frac{1}{\sqrt{N}}$. Together, these constitute the baseline S2 pulse area resolution in LZ.

Figure 5.26 shows the fraction of charge lost to fragmentation in candidate events. The projected S2 pulse area resolution from SE area and electron lifetime are included for reference. At the top of the WIMP ROI, charge loss from fragmentation is negligible relative to other sources of S2 uncertainty. At the bottom of the ROI, around the S1-S2 threshold, the loss of charge in candidate events is comparable to the S2 pulse area uncertainty from electron lifetime.

These measurements demonstrate that pulse fragmentation is not a primary concern in standard S1-S2 analysis. In physics searches at lower energies, where

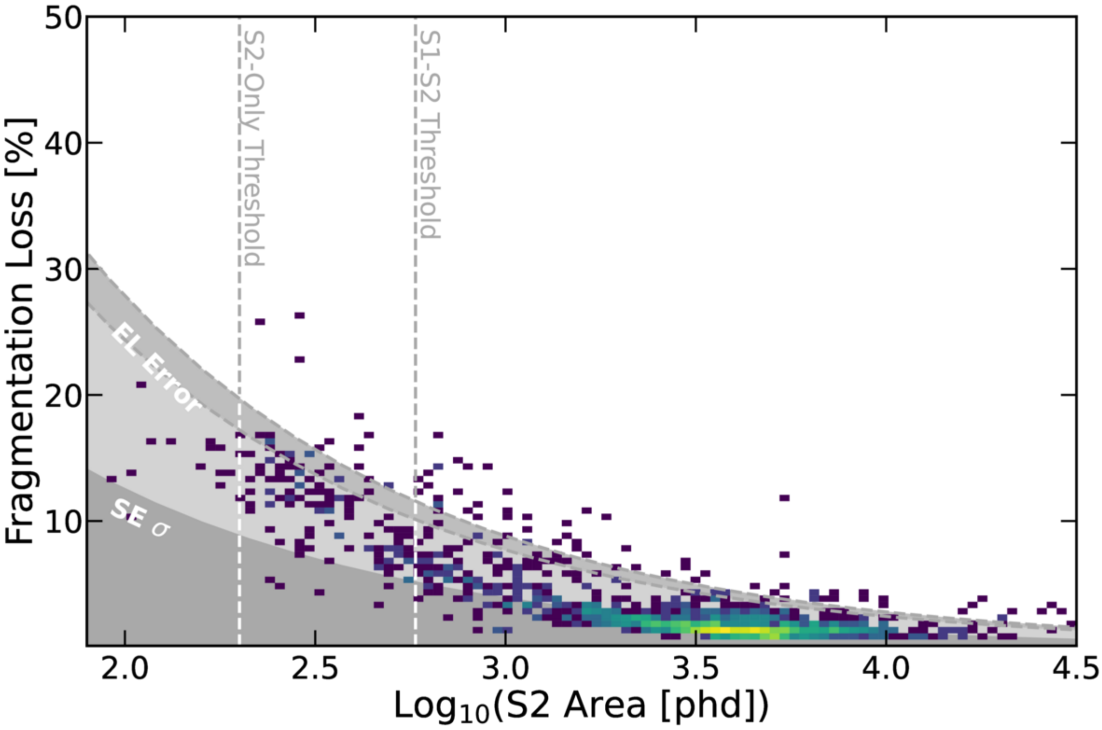


Figure 5.26: Fractional charge loss vs S2 pulse area for fragmented candidates in CH_3T calibration data. The projected S2 pulse area resolution from two other systematic uncertainties, electron pulse area resolution and the electron lifetime, are annotated. Charge loss from fragmentation is subdominant to other sources of systematic S2 pulse area uncertainty, down to the 14.5 e threshold considered in S1-S2 analysis.

fragmentation is comparable in impact to other systematic uncertainties, fragmentation probability can be introduced to PLR analysis as a nuisance parameter. This process is demonstrated using simulation data in Section 5.3.3.

In S2-Only analysis, discussed in Chapter 6, events are considered for analysis if an S2 contains no S1 or S2 pulse in a time window 1050 µs before it, regardless of prominence. This reduces the rate of background events considered in this analysis, which tend to be disproportionately split compared to signal events. However, as some fragmented S2s would be rejected by this cut, this results in pulse fragmentation becoming a source of signal inefficiency in S2-Only analysis.

Figure 5.27 shows the fraction of candidate events against S2 pulse area which contain a fragment pulse classified as an S2 before the main S2 in the event. While some signal events would be rejected by this cut, this fraction is found to be $< 1\%$ up to pulse areas of 6000 phd. At low pulse areas, CH_3T data does not fully represent signal, due to the presence of field cage resistor events. However, this measurement has been repeated using AmLi, with similar results. A cut on non-prominent S2 pulses is therefore found to have minimal impact on signal acceptance in an S2-Only analysis, and is capable of rejecting low-energy backgrounds.

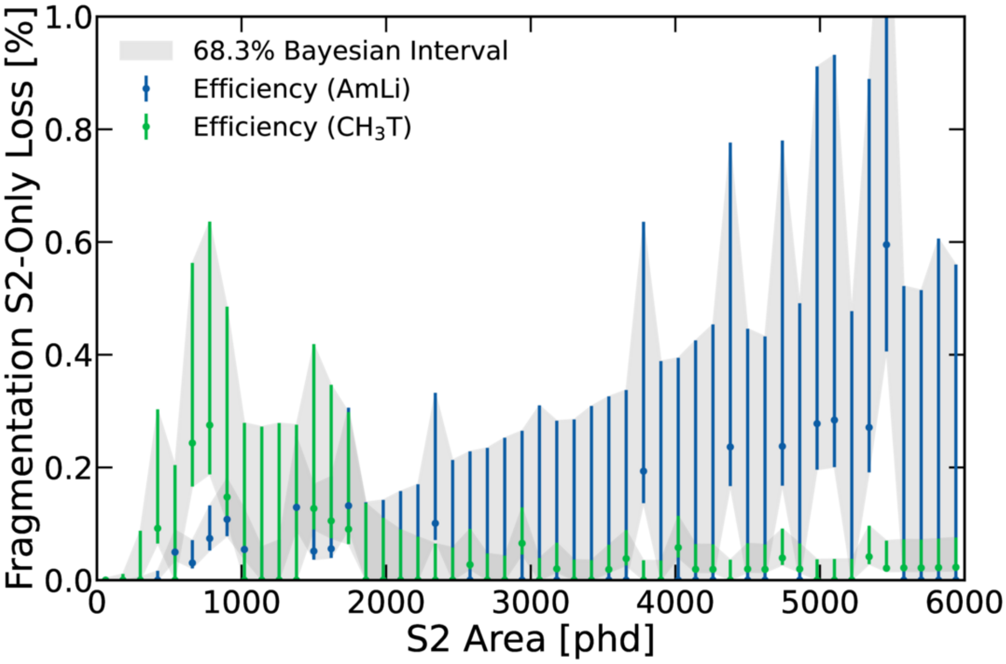


Figure 5.27: Efficiency loss to non-prominent S2s in an S2-Only analysis, measured using CH_3T and AmLi calibration data. Probabilities are consistently below 1%, indicating that a cut on non-prominent S2 pulses in S2-Only analysis has negligible impact on signal acceptance.

5.2.5 S2 Splitting in Calibration Data

The final form of pulse mis-classification to measure in calibration data is the splitting of a single S2 into two prominent S2 pulses, illustrated in Figure 5.7. As the detection efficiency of S2 pulses approaches 100% at S2 pulse areas of 500 phd [85], splitting is the dominant contribution to signal inefficiency in most physics searches in LZ, with greatest impact in sensitivity to low-mass WIMPs.

RQ-Based Cuts

S2 splitting and fragmentation both involve the separation of one pulse into two separate pulses. As such, the cuts outlined in Section 5.2.4 are also effective at selecting split candidates. However, the boundaries of these cuts are made more restrictive:

Time Separation: The probability of the two pulses having zero separation is much higher. As such, timing separation is reduced from 2000 ns to 20 ns.

Area Fraction: The secondary pulse is much closer in size to the main S2. The cut in area fraction has been changed from 1%-20% to 5% upwards, as seen in Figure 5.21.

Width Ratio: A prominent S2 nearer to the primary S2 is more likely to have a continuous waveform. The cut in width fraction has been changed from <2 to <1, as seen in Figure 5.24.

XY Separation: As split S2s contain larger pulses, XY reconstruction is more precise, allowing a cut to be placed at 5 cm.

These cuts reject >99% of MS events from the list of split candidates. With the application of RQ cuts alone, the probability of splitting can be said to be below 2% for any energy and TPC position in calibration data, using 3666 candidate events. However, the majority of candidate events are still correctly classified MS events. This would result in measurements of pulse splitting providing a conservative estimate by up to an order of magnitude.

In order to measure the probability of S2 splitting in calibration data, as opposed to an upper limit, it is essential to discriminate MS S2s from split candidates with a rejection power beyond that possible using RQs alone.

This remainder of this section outlines the analysis of S2 waveforms to collect information on an event beyond that available in LZap. This information is used to reject MS events through two approaches: the use of a simple cut in waveform width, and a multi-parameter cut trained through unsupervised machine learning. These approaches are demonstrated to produce results in agreement with one another within errors, demonstrating that remaining events are indeed split S2s.

Waveform Pulse Width

When an S2 pulse is split in half, the width of each resultant pulse is approximately half of the width of the original S2. When the waveform of the two pulses are combined, and the width of this pulse is derived, a split S2 will exhibit a physical pulse width. In contrast, two S2s which have been correctly split will exhibit a pulse width above the expected relationship when merged, similarly to merged S2s in Figure 5.17. A cut can therefore be applied to reject MS S2s from a list of candidates by calculating the pulse width were two pulses to be merged, and rejecting events where this pulse width is unphysically high.

Figure 5.28 shows the pulse width vs drift time relationship for candidate events, using both the LZap pulse width and a pulse width derived when merging waveforms of neighbouring pulses. A simple $\sqrt{Mt + C}$ cut is applied to data, with events passing this cut highlighted in red. Tagged events generally exhibit a physical pulse width when merged and an unphysically low pulse width in isolation, indicating that these events are likely split. Events rejected by the cut exhibit physical pulse widths in isolation and unphysically broad pulse widths when merged, indicating that events contain MS S2s which have been correctly separated.

A cut in merged pulse widths preserves just 347 events, out of 3666 passing all RQ cuts. However, $\mathcal{O}(50)$ of these events follow a physical relationship in both merged and unmerged S2 pulse widths, making it ambiguous in this metric as to whether or not these events are split. On top of this, it is anticipated that signal loss and background leakage will increase considerably in this metric for smaller S2s. This is as a result of the uncertainty on pulse width increasing with decreasing

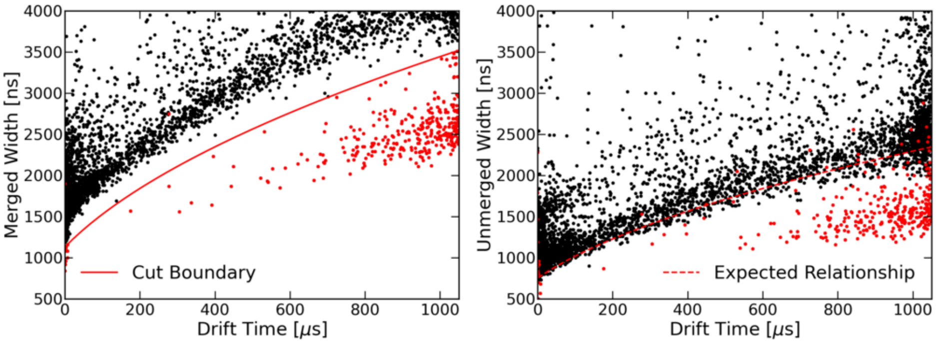


Figure 5.28: Identifying split candidates using the width of unmerged and merged waveforms. Width here is defined as the AFT75-25 of a waveform. Left: Merged pulse widths against drift time. A $\sqrt{Mt+C}$ cut is annotated in red. Events passing this cut have a physical S2 pulse width when merged, and are annotated in red. Right: Unmerged pulse widths against drift time. Events passing the cut in merged pulse widths exhibit unmerged pulse widths below the mean relationship in SS S2s, indicating that the majority of these events are examples of pulse splitting.

pulse area, as detailed in Section 4.3.2. For these reasons, a second cut rejecting events on the basis of their pulse shape has been developed, as a cross check.

Single-Gaussian KS Tests

If an S2 is split, the waveform of the two split pulses should resemble a Gaussian distribution, and the waveform of each individual pulse should resemble half of a Gaussian. In contrast, if an event is correctly classified, each individual pulse should resemble a Gaussian. A set of KS tests comparing the waveform of pulses to these distributions can therefore provide information on the likelihood of an event containing a split S2.

Figure 5.29 illustrates the three distributions used to test waveforms. The first, a normal Gaussian, is used to test if the merged waveform resembles an S2. The second, an asymmetric normal distribution, tests an individual pulse waveform to test if it resembles a MS S2 with low separation (z). The third, a half-Gaussian, tests if a pulse has been mis-classified.

An event containing pulses which produce low KS test results for a skew Gaussian (a 'good' split) and high KS test results for a half-Gaussian (a 'bad' split) is likely to

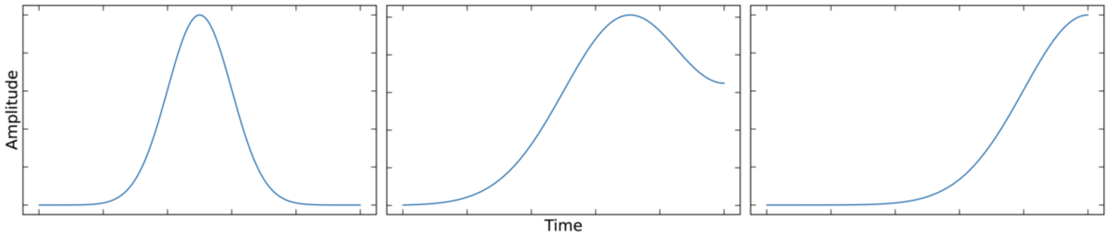


Figure 5.29: The example distributions used to test the shape of pulse waveforms. Left: A normal distribution, used to test if a merged waveform resembles an S2. Centre: A skew Gaussian distribution, used to test if a single pulse resembles a MS S2. Right: A half-Gaussian distribution, used to test if an S2 has been mis-classified.

contain a mis-classified S2. This concept is illustrated in Figure 5.30. The waveform of a mis-classified S2 is shown, with a vertical line annotating the position at which the waveform was split into two separate pulses. The right-hand pulse is then shown as a cumulative distribution function (CDF), alongside the CDF of the skew Gaussian and half-Gaussian distributions seen in Figure 5.29. The right pulse tests well against a half-Gaussian (a 'bad split') and badly against a skew Gaussian (a 'good split'), indicating that the waveform contains a mis-classified S2.

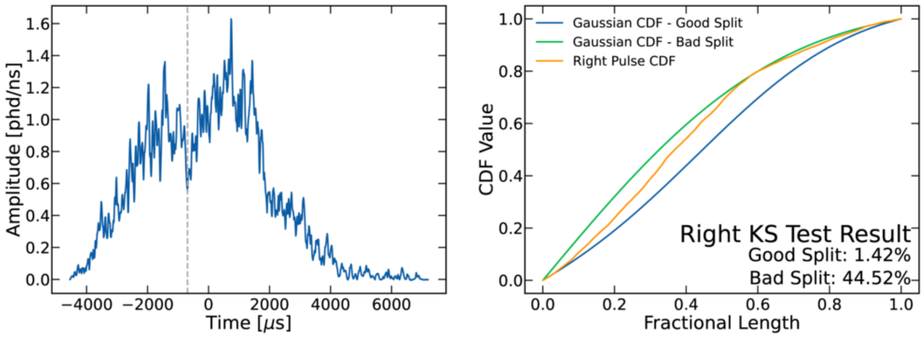


Figure 5.30: Left: Example waveform of a split S2 from data. The total waveform resembles a normal distribution. A vertical line indicates the point at which one pulse ends and another begins. Right: Cumulative distribution function (CDF) of the right-hand pulse. CDFs of a skew Gaussian distribution (a 'good' split) and of a half-Gaussian distribution (a 'bad' split) are overlain. The right pulse tests better with the latter than the former, indicating that the waveform contains a mis-classified S2.

Double-Gaussian KS Tests

Instead of probing individual pulses, a KS test can instead be used to quantify the likelihood of a waveform originating from two Gaussian distributions. A CDF is produced of two normal distributions separated by a distance D. The KS test result is then derived comparing this distribution to the full waveform of two pulses. This is repeated for different values of D. A correctly classified MS event will produce low test results at low D, which eventually increase at a fixed point - the separation (z) between MS S2s. In contrast, a mis-classified S2 will test well at low D, and will not tend towards a peak.

This concept is illustrated in Figure 5.31, using the waveform seen in Figure 5.30. The KS test result is shown for varying separation D. Test results remain at around 60% until $D = 0.8\sigma$, at which point test results rapidly decrease. As this test result is high at low D and does not form a peak at high D, it can be concluded that this waveform contains a single S2 pulse.

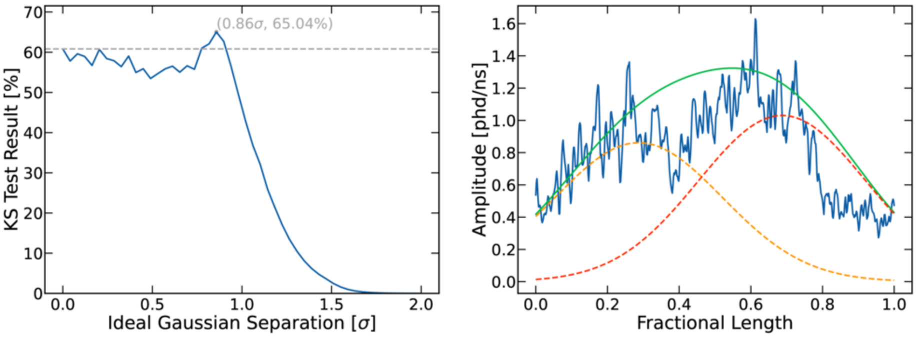


Figure 5.31: Left: A double-S2 KS test performed on an example waveform, with varying separation. Test results are high at low separation, and low at high separation, indicating the waveform is of a single mis-classified S2. Left: The example waveform, with the best test result overlain. The test result more closely resembles a single Gaussian, rather than two Gaussians.

Rejection of MS Events using a Gaussian Mixture Model

KS tests are, in isolation, capable of rejecting MS S2s. However, as with all cuts used to identify pulse mis-classification, a simple cut results in considerable background leakage or signal loss. A GMM is therefore used to identify split S2s, using the KS test results outlined in Section 5.2.5. Test results are inputted into a GMM and sorted into 15 clusters. Two clusters have been identified to contain test results indicative of split S2s. These 348 events have been selected as split candidates using this approach.

Events passing this cut can be used as a cross-check for the cut in pulse widths described earlier in Section 5.2.5. The merged pulse width of events identified as split using their pulse shape can be seen in Figure 5.32, alongside the same distribution for events identified as MS S2s. The cut boundary used in Figure 5.28 is shown for reference.

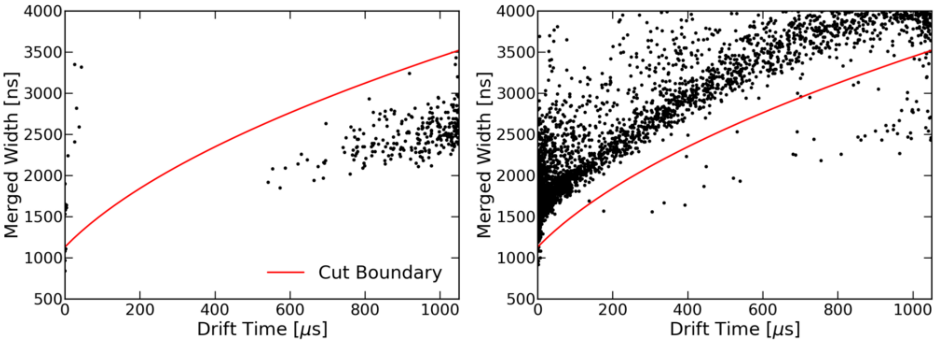


Figure 5.32: The merged width vs drift time distribution of candidate events. Left: The distribution of events identified as split using pulse shape. The cut used to identify split S2s using their pulse width in Section 5.2.5 is annotated. 13 events at low drift times pass a cut in pulse shape but fail a cut in pulse width. Right: The distribution of events identified as MS S2s using pulse shape. 85 events at high drift times fail a cut in pulse shape, but pass a cut in pulse width.

At high drift times, all events identified as split using their waveform shape exhibit a physical pulse width, as expected. 13 events are identified near to the gate array at low drift times, which do not follow the expected relationship. The majority of events identified as MS S2s using their waveform shape exhibit a pulse width as expected. However, 85 events pass cuts in pulse width, but fail a cut using pulse shape.

Figure 5.33 shows the waveforms of several events passing a cut in pulse width but failing a cut in pulse shape. Of the 85 events satisfying this condition, 55 are outside of the WIMP ROI. Upon closer inspection, all 55 originate from high-energy MS events where a waveform has split asymmetrically, as seen in the top half of Figure 5.33. This asymmetry allowed events to pass a cut in pulse width. Within the ROI, the remaining 30 events exhibit pulse shapes which make them difficult to distinguish from low-energy MS events. These events tend towards the field cage resistors, where MS events and distorted S2s are more probable.

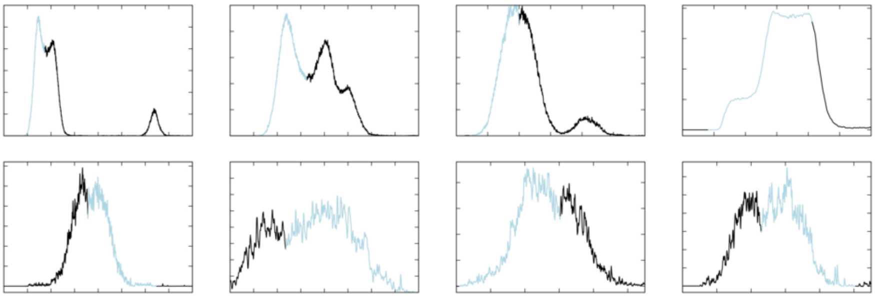


Figure 5.33: Example waveforms of CH_3T calibration events identified as split by pulse widths, but not identified as split by KS tests. S2s on the top row are visibly multiple scatters. It is difficult to identify by eye if S2s on the bottom row are single or multiple scatters.

Results

Candidate events selected by cuts on merged pulse width and pulse shape can be compared to the distribution of all SS events, in order to quantify the probability of splitting. The distribution of events in S2 pulse area before and after cuts can be seen in Figure 5.34. After cuts using RQs, the 3666 remaining events are distributed in two peaks above and within the S2 ROI. With the application of a cut in waveform width, the majority of events above the ROI are rejected, as are smaller S2 candidates from the field cage resistors. With a cut in pulse shape, all events above the ROI are rejected.

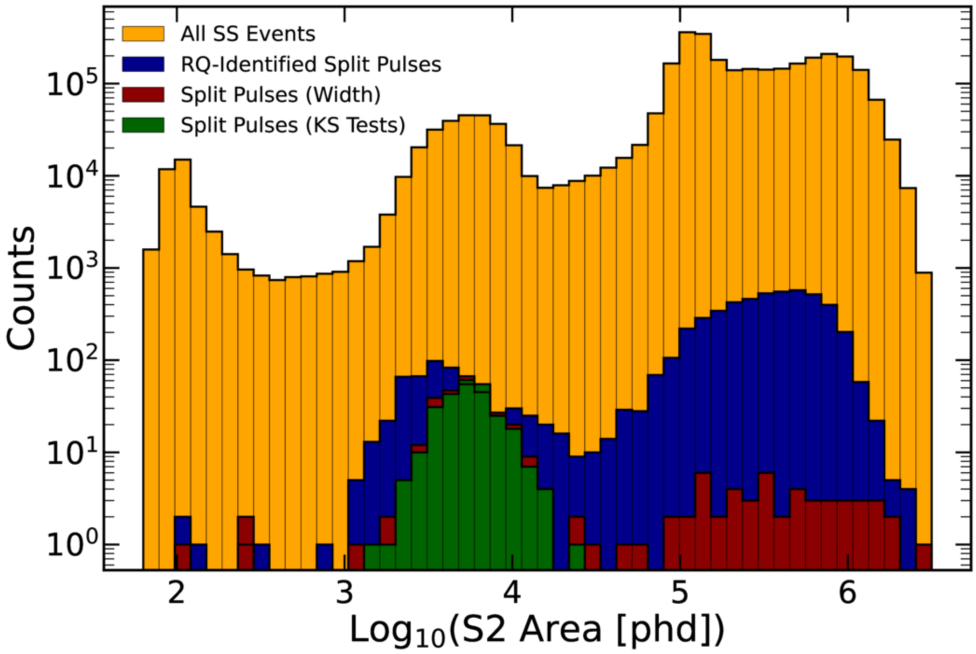


Figure 5.34: The S2 pulse area distribution of events in CH_3T calibration data which are classified as SS events (orange), tagged as split using RQ cuts (blue), tagged as split using the width of the merged waveform (red) and tagged using KS-tests of waveform shape (green). Pulse splitting is found to occur predominantly in the ROI.

The probability of pulse splitting for a given S2 pulse area can be seen in Figure 5.35. The mean probability is defined using the total number of SS events and candidates identified as split using pulse shapes, as seen in Figure 5.34. As some candidates may be incorrectly rejected through cuts in pulse shape, error bars are derived through a 1σ Bayesian interval using candidates passing a cut in merged pulse width.

Above the ROI, the probability of pulse splitting is found to be <0.01% in data. This probability increases within the ROI, to a maxima of $(0.15 \pm 0.02)\,\%$. This probability is observed to decrease in data below the ROI: however, this region contains fewer calibration events, instead being dominated by accidentals.

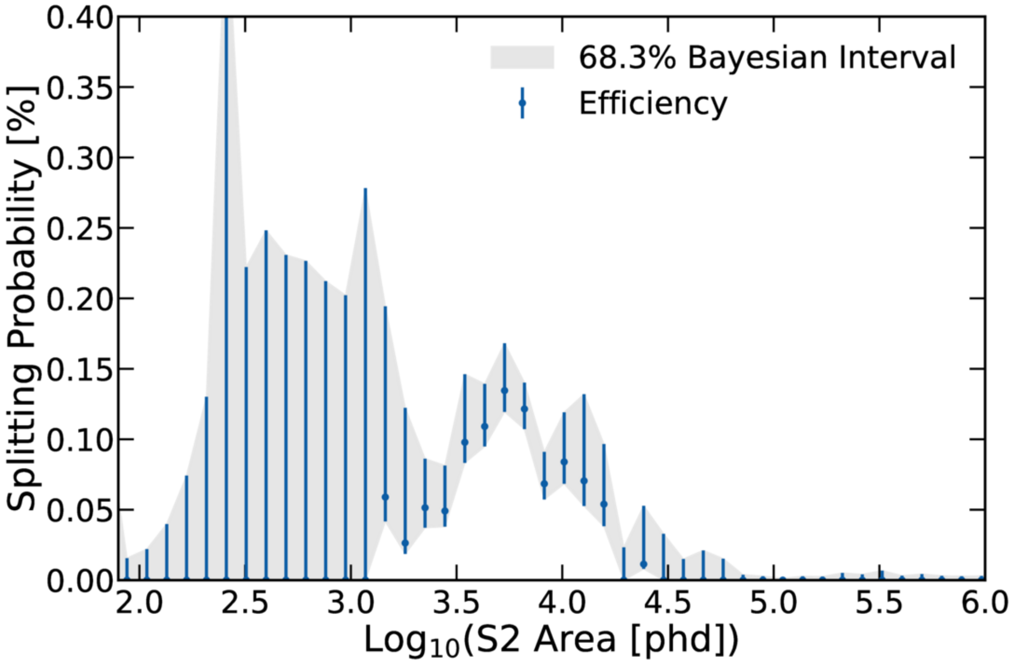


Figure 5.35: Splitting probability against S2 pulse area, for CH_3T calibration data. Above the ROI, the probability of pulse splitting is measured to be below 0.01%. Within the ROI, splitting does not exceed 0.15%. Splitting below the ROI is projected to be improbable, although CH_3T calibration data is not observed at these S2 pulse areas.

This probability can also be measured across all S2 pulse areas in R^2 and drift time, in order to find the variation of pulse splitting with TPC position. This distribution can be seen in Figure 5.36. Splitting probability is greatest in the TPC centre at high drift times. The probability of splitting decreases linearly with increasing event radius. This distribution matches predictions made using toy model S2s in Figure 5.10.

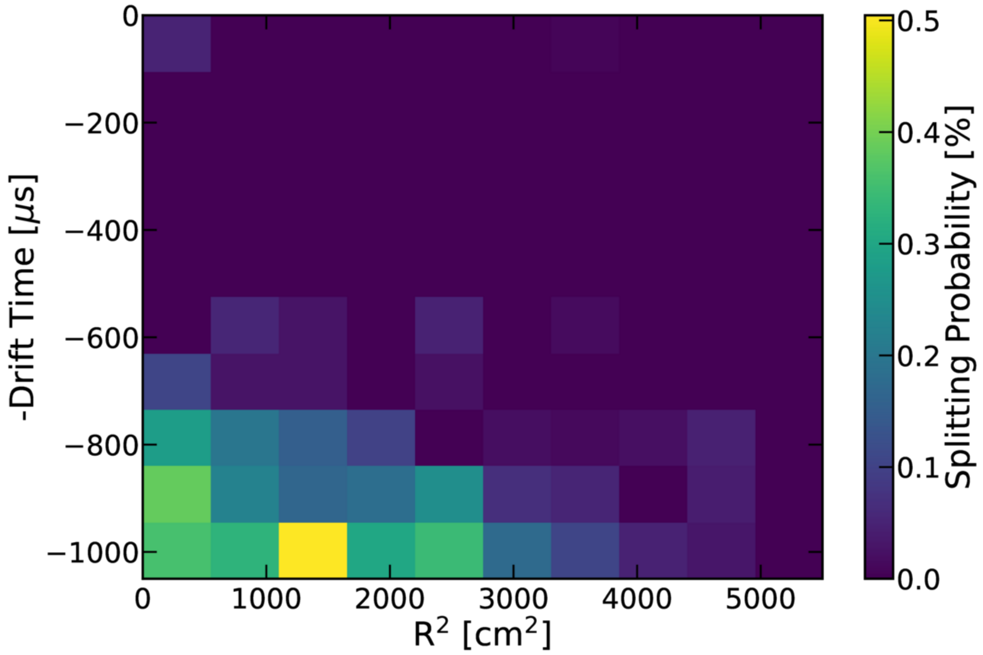


Figure 5.36: Splitting probability against radial position and drift time, for CH_3T calibration data. Probability is highest in the TPC centre at high drift times. Probability decreases linearly with radius, as expected.

5.3 Validation using Simulation Data

5.3.1 Motivation

LZap mis-classification has been measured in calibration data, and has been found to be of minimal impact in S1-S2 analysis. However, the presence of MS events introduces systematic errors. As the true origin of simulated events are known, they can be used to project mis-classification probabilities without background leakage.

The statistics available for measurements of mis-classification in calibration data for small S2s are also low. Mis-classification probabilities using calibration data decrease for small S2s below the S1-S2 ROI, as data is dominated by accidentals, wall events, and events near to the field cage resistors, while mis-classification is predicted by toy models to increase in this region. As simulated events can be produced of any energy, they allow for reliable projections of mis-classification probability below the ROI, assuming that simulations correctly model small electroluminescence waveforms within errors.

The GEANT4-based simulation framework BACCARAT [159, 160] has been developed for LZ to the point where full waveforms can be generated for simulated interactions. Clusters of N electrons are placed randomly inside the simulated LZ TPC and the resultant waveform processed through LZap. This can then be used to project the probability of splitting and fragmentation for a given size of S2 below those seen in calibrations.

BACCARAT has been tuned such that RQs from simulated data reflect those observed in real events. However, differences between simulations and real TPC physics may result in discrepancies between measured probabilities in data and simulations. For example, simulated events do not photoionise the gate or glue ring. To resolve this, simulated events are generated up to $n_e = 100$ (~4000 phd). If simulations and data are in agreement as to the splitting and fragmentation probability at this S2 pulse area, simulation data can be relied upon for projections of mis-classification below this.

5.3.2 Methodology

In BACCARAT, electrons are deposited at fixed TPC positions for a given event. 5×10^4 events have been generated using clusters of 10 and 20 electrons, with 2×10^4 events generated for multiples of 10 above this, up to 100 e.

As other detector backgrounds are absent, all pulses in an event originate from an S2. The true pulse area can therefore be defined as the sum of all S2 and SE pulses in an event. An S2 is considered lost if no S2 pulse is contained in an event, and detected if at least one S2 pulse is present. An event is fragmented if more than 1% of the total pulse area in an event window is not part of the main S2 pulse. An event is split if the event contains more than one S2 pulse. In order to ensure a conservative estimate of splitting, an S2 does not need to be identified as prominent in simulations for an event to be split.

In S2-Only analysis, fragmented events are a source of efficiency loss when a non-prominent S2 pulse is seen in an event before the main S2. When considering efficiency loss from fragmentation in an S2-Only analysis, events are only considered which contain more than one S2 pulse, where the first S2 is not the largest in the event. These classifications, and the conditions for these classifications, are outlined in Table 5.2.

Table 5.2: Classifications of events in simulation data, alongside the number of S2s and prerequisites for each classification.

Status	S2s	Other Conditions
Lost	0	None
Detected	1+	None
Fragmented	1 Prominent	Other S2 Area > 1%
Fragmented (S2-Only)	2+	$S2_a < S2_b$
Split	2+	None

5.3.3 Results

S2 Fragmentation

The percentage of events classified as fragmented in simulation data can be seen in Figure 5.37 against S2 pulse area. Fragmentation probability increases with decreasing S2 pulse area, to a maximum of (60±5) %. Below 100 phd, this probability is observed to decrease: below this point, events containing a mis-classifed S2 do not contain an S2 pulse, and therefore contribute to overall S2 detection efficiency. Within the S1-S2 ROI, fragmentation is found to be $\mathcal{O}(5\%)$, in agreement with measurements made using calibration data.

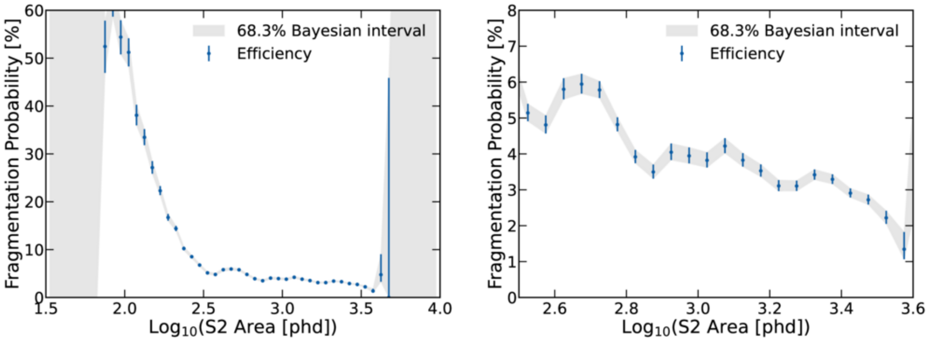


Figure 5.37: S2 fragmentation probability against S2 pulse area, derived using simulation data. Left: Fragmentation for all pulse areas considered. Fragmentation probability increases to a maximum of (60 ± 5) % at 100 phd, below which mis-classified events instead contribute to detection efficiency. Right: The same distribution, within the S1-S2 ROI. Fragmentation is found to be $\mathcal{O}(5\%)$, in agreement with calibration data.

The efficiency loss from fragmentation in an S2-Only analysis can be seen in Figure 5.38. Efficiency loss is consistently below 0.5% across all S2 pulse areas considered, in agreement with calibration data, down to S2 pulse areas of 300 phd. Below this, signal loss increases considerably. However, this is below the current established threshold for S2-Only analysis of 200 phd (see Chapter 6).

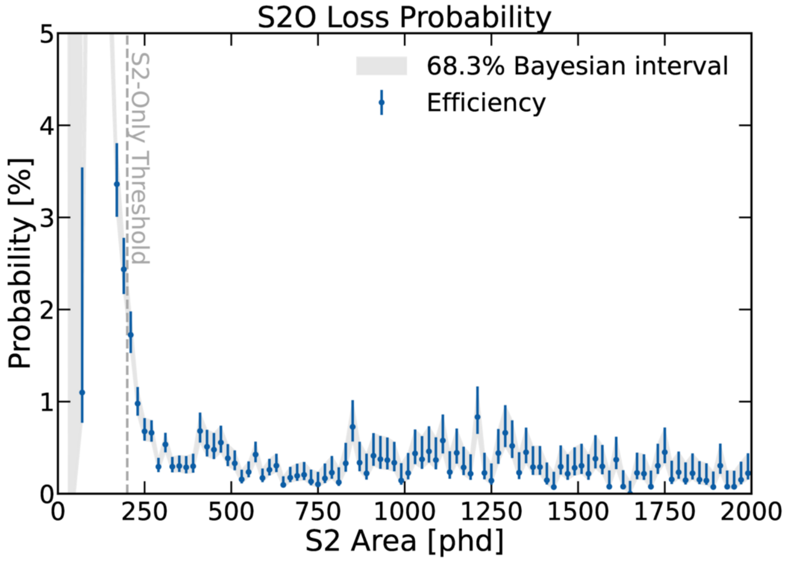


Figure 5.38: S2-Only efficiency loss from non-prominent S2s for varying primary S2 area, generated using simulation data. Signal loss is low across all S2 pulse areas, up to the threshold of S2-Only analysis.

The charge loss from S2 fragmentation can be seen in Figure 5.39, compared to other sources of S2 pulse area uncertainty. Within the S1-S2 ROI, charge loss is subdominant to other sources of uncertainty, as observed in calibration data in Figure 5.26. However, below this, fragmentation becomes more probable, and charge loss from fragmentation is greater than that from other sources of uncertainty. As a result, pulse mis-classification is the dominant source of S2 pulse area resolution in low-mass WIMP searches such as S2-Only analysis.

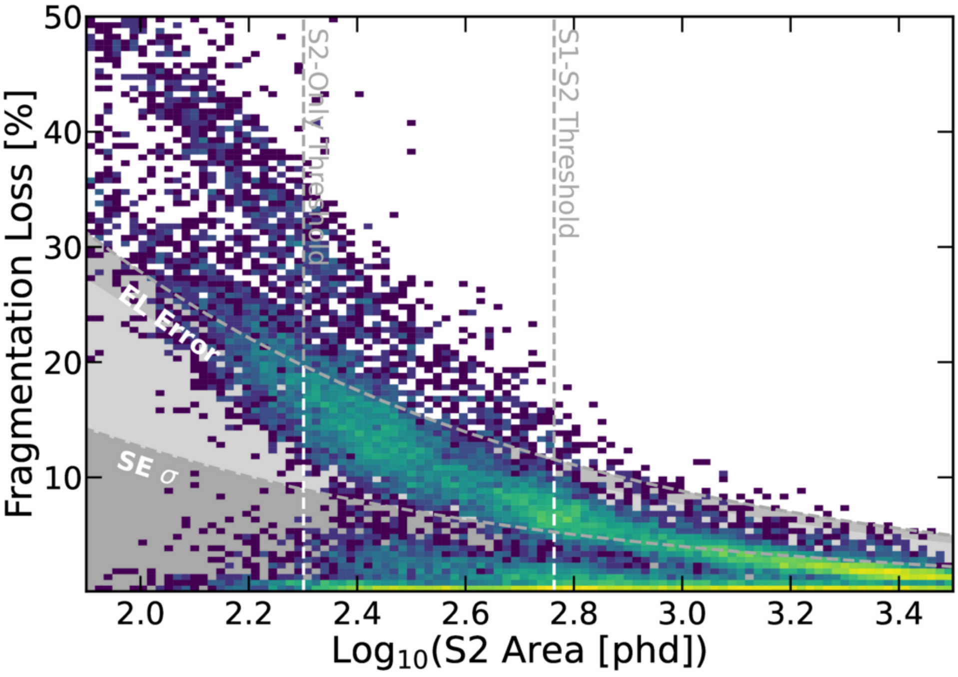


Figure 5.39: Percentage loss of charge in fragmented S2s vs S2 area, for simulated events. S2 pulse area uncertainty from electron lifetime and SE pulse area fluctuations are annotated, for reference. In the S1-S2 ROI, charge loss from fragmentation is subdominant to other sources of pulse area uncertainty. At pulse areas below this, used for low-mass WIMP searches, fragmentation is the dominant source of pulse area uncertainty.

The distribution of pulse area, drift time, and charge loss has been interpolated to produce a coarse PDF of fragmentation. For a given event energy and position, a PDF can be generated of the probability of losing a given amount of pulse area from the recorded S2 in an event. This can be used to minimise low-mass WIMP sensitivity uncertainty as a nuisance parameter in PLR analysis.

This PDF is illustrated in Figure 5.40. A distribution of probability against lost pulse area is shown for a 10 e event of drift time 0 µs to 100 µs and 950 µs to 1050 µs. By sampling from this PDF, PLR analysis can account for the S2 pulse area uncertainty introduced by mis-classification.

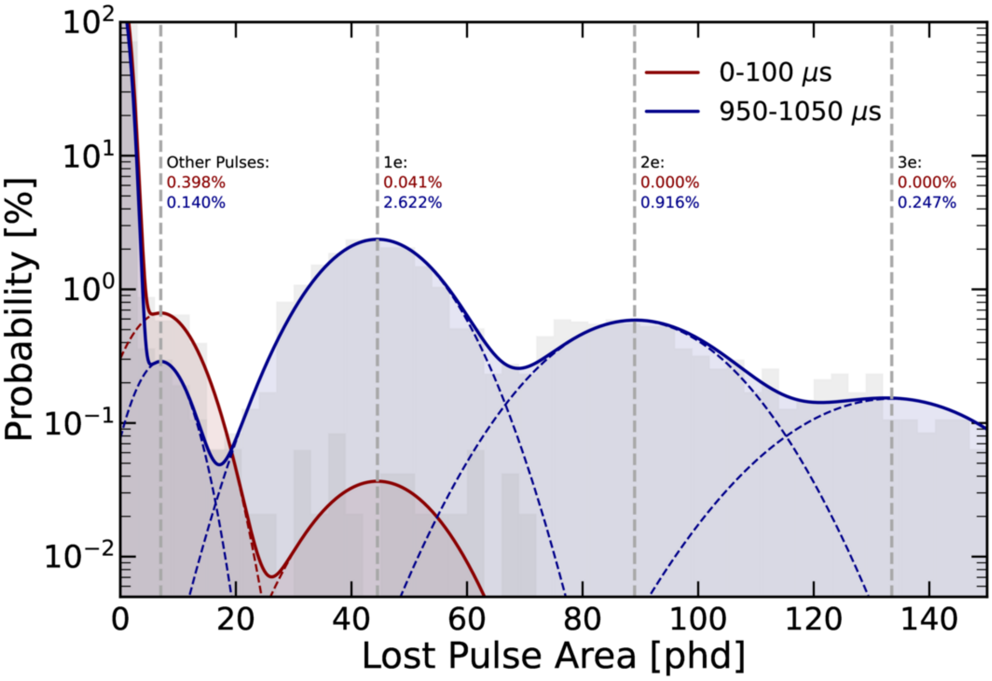


Figure 5.40: PDF of charge loss for simulated 10e S2s, generated at high and low drift times. Data (shown in low opacity) has been fitted to a multi-Gaussian distribution, to reflect the loss of a fragment of charge, 1e, 2e, and 3e. The interpolated distribution of data can be used to introduce the probability of charge loss from pulse mis-classification into PLR analysis.

S2 Splitting

Pulse splitting probability against S2 pulse area can be seen from simulated data in Figure 5.41. Splitting probability increases with decreasing S2 pulse area, to a maximum of $(4.0 \pm 0.8)\,\%$. As with pulse fragmentation, this probability decreases below 120 phd due to S2 detection efficiency. Within the S1-S2 ROI, splitting is more probable than observed in calibration data, reaching a maximum of $(0.5 \pm 0.1)\,\%$. This is a result of the conservative definition of S2 splitting used in simulations, as well as systematic errors in measurements using calibration data. However, in both simulations and calibrations, pulse splitting is minimal in the S1-S2 ROI.

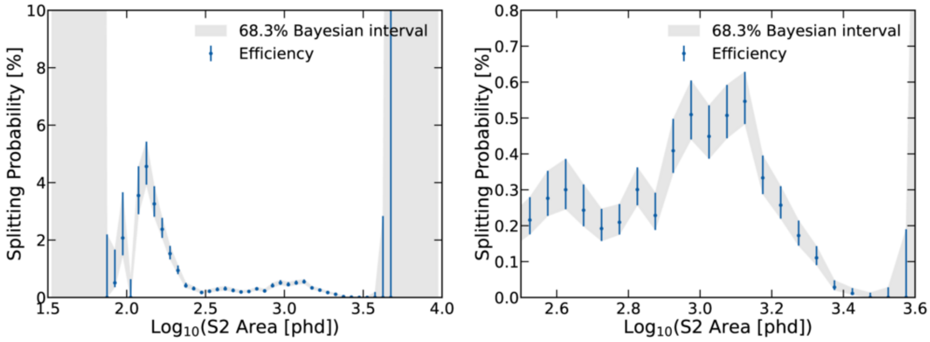


Figure 5.41: Splitting probability against S2 pulse area, using simulation data. Left: Splitting probability across all S2 pulse areas. Splitting becomes more probable with decreasing S2 pulse area, to a maximum of $(4.0 \pm 0.8)\,\%$ at 120 phd. Probabilities below this pulse area decrease due to S2 detection efficiency. Right: Splitting probabilities within the S1-S2 ROI. Splitting probability is greater than that observed in calibration data, because of the more conservative definition of splitting used in this analysis.

The distribution of pulse area and drift time has been interpolated to produce a coarse PDF of splitting. For a given event energy and position, the probability of an event splitting can be measured, alongside 1σ errors. As with the PDF of fragmentation, this can be inserted into PLR analysis as a nuisance parameter.

This PDF is illustrated in Figure 5.42. The detection efficiency (1 - splitting probability) is measured for each simulated S2 size, for slices of drift time. For a fixed S2 pulse area and drift time, data from this graph can be interpolated to identify the probability that this event is lost due to pulse mis-classification.

The PDF of fragmentation presented in this chapter will be used in low-energy analyses. However, fragmentation in S1-S2 analysis has been demonstrated to be of low concern. The PDF of splitting presented in this chapter is currently used in WIMP search analysis, to account for pulse mis-classification as a source of signal inefficiency.

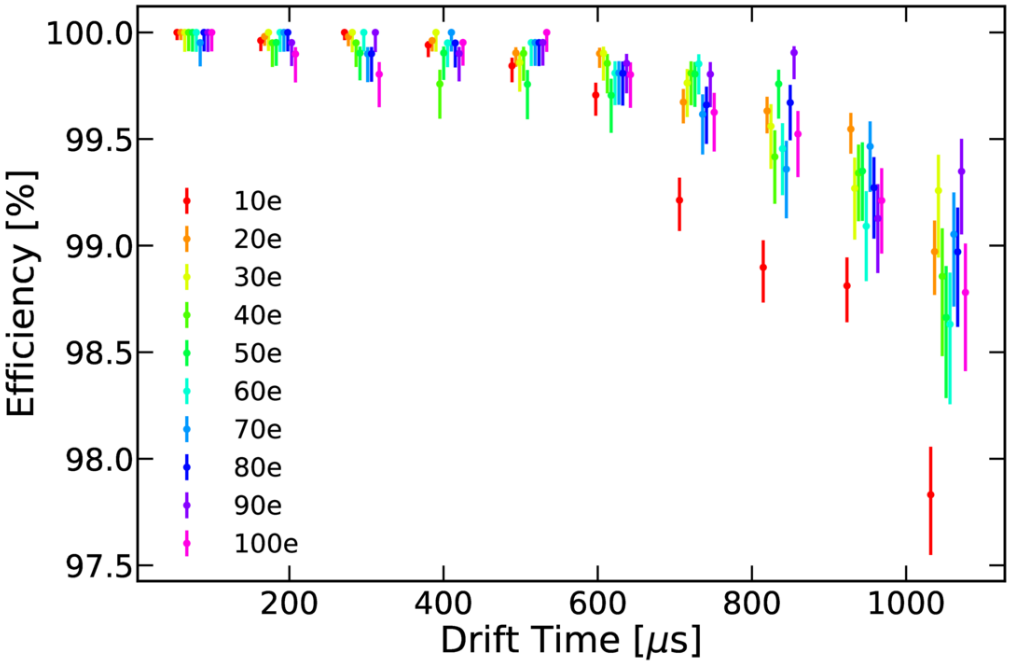


Figure 5.42: Splitting probability vs drift time for varying numbers of electrons, from simulation data.

5.3.4 Summary

This chapter produces measurements of S2 mis-classification in signal-like data in LZ. These measurements can be used to mitigate the bias of both high-energy and low-energy analyses.

Figure 5.19 demonstrates that pulse merging is negligible in the current WIMP S1-S2 ROI. This work can be repeated for iterations of LZap, to ensure anomalies do not arise in WIMP search data. Measurements above the ROI indicate that S2 merging produces a significant signal loss in high-energy analyses, whereas merging was previously anticipated to be more prominent towards the TPC walls and therefore negligible in such an analysis.

Figure 5.39 indicates that fragmentation of S2s is subdominant in the current WIMP S1-S2 ROI to other sources of S2 area resolution, such as the electron lifetime and single electron pulse area uncertainty. It has been decided as a result of this work to neglect fragmentation in current WIMP search analysis. However, fragmentation increases in probability and physics impact at lower S2 areas, which will be of interest in ^{8}B and S2-Only analyses. In this context, a PDF can be designed, as in Figure 5.40, to account for this effect.

Figure 5.42 indicates that splitting of S2s is minimal in the current WIMP S1-S2 ROI. This has been included as a PDF in current analysis, such that the PLR is capable of accounting for this form of signal loss. In future analyses at lower energies, where splitting is a more considerable form of signal loss, this measurement with uncertainties can be incorporated as a nuisance parameter into the PLR.

6

Rejection of Grid Radiogenic Backgrounds in an S2-Only Analysis

This chapter presents the development of analysis cuts to reject grid radiogenic backgrounds in LZ in an electroluminescence-only (S2-Only) analysis. After a clean selection of gate, cathode, and signal events is acquired through the use of unsupervised machine learning, a cut in pulse widths is demonstrated to reject an average of 60% of background events in an S2-Only analysis while preserving 80% of signal data. In upcoming science runs in LZ, where statistics are higher, a Boosted Decision Tree has been trained on grid radiogenic and signal data. This BDT is demonstrated to reject an average of 90% of background events while preserving 80% of signal data.

This chapter then concludes with the application of measured BDT background leakage and signal acceptance to simulated events. It is projected that, with the application of a cut, LZ is capable of discovering a 2 GeV WIMP of cross section $1 \times 10^{-46}\,\mathrm{cm}^2$ to 5σ in 50 live days of data, compared to 175 days with a cut in pulse width, and 480 days without the use of analysis cuts.

6.1 Grid Radiogenic Backgrounds

6.1.1 Overview

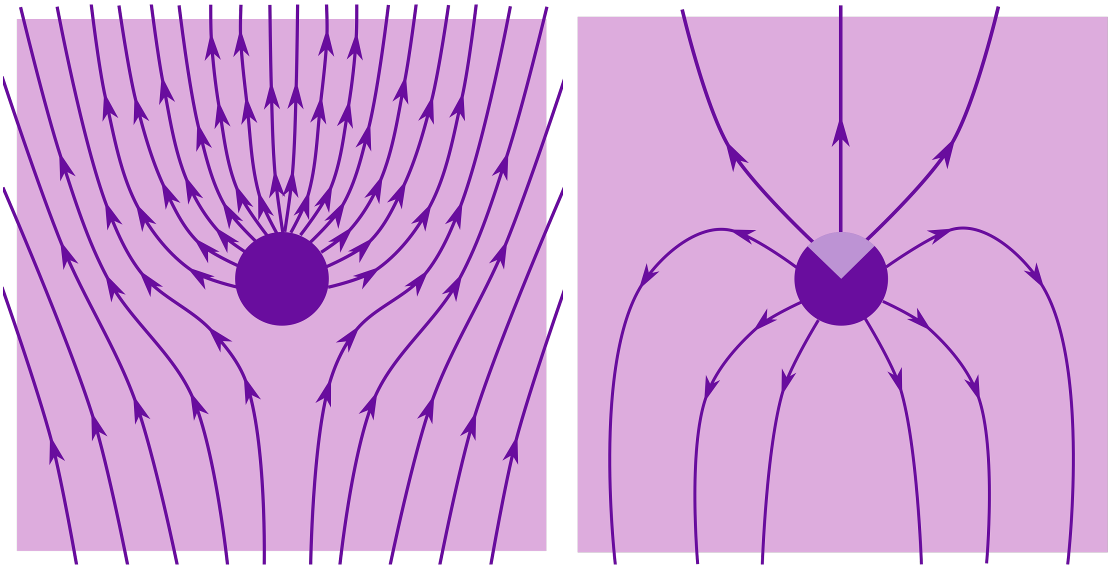
The materials used in the construction of the electric field grids in LZ contain an intrinsic background from the decay chain of ^{238}U and ^{232}Th. During construction, atmospheric ^{222}Rn and its progeny attached to the surface of the wire, in a process referred to as plate-out. While the intrinsic contamination of the grids was minimised through an intense cleanliness control program [126] and atmospheric contamination was minimised during fabrication of the grids [132], ^{222}Rn produces a dominant ER background in LZ data, with a subdominant contribution from ^{238}U and ^{232}Th decay chains [128].

Figure 3.7 shows the decay chain of ^{222}Rn and its progeny. ^{222}Rn on wires quickly decays into ^{210}Pb, which has a half life of 22.3 years. During operation of LZ, this gradually decays into its daughter nuclei, ^{210}Bi and ^{210}Po, before eventually decaying into the stable ^{206}Pb. These decays result in the grid wires remaining a constant source of β decays in LZ.

In analysis, this background is rejected in S1-S2 analysis through the application of a minimum and maximum S1-S2 drift time as part of a fiducial volume. However, as an S1 is not present in S2-Only data, events do not have a measured drift time.

The energy of β decays in the ^{222}Rn decay chain is such that the majority of decays in the TPC bulk produce a detected S1 pulse, producing an S1-S2 background. However, the electric field strength around grid wires is such that the rate of recombination in LXe decreases considerably. This increases the probability of an interaction producing no detected scintillation, resulting in grid radiogenics becoming a dominant contributor to iS2 backgrounds in LZ.

The field strengths around the gate and cathode are illustrated in Figures 6.1(a) and 6.1(b). Separation between field lines near to the gate wire in Figure 6.1(a) indicate that the induced electric field is stronger nearer to the grid wires. The presence of the RFR changes the field around the Cathode wire, illustrated in Figure 6.1(b). The majority of field lines travel towards the RFR, resulting in electrons producing a large iS1 pulse. The rate of iS2s from the gate is therefore higher than that from the cathode. Events which drift upwards towards the extraction region from the cathode still lose some charge to the RFR, resulting in smaller S2 and iS2 pulses.



(a) Gate Field **(b)** Cathode Field

Figure 6.1: Illustrations of electric field lines around the gate and cathode wires in LZ. Left: Gate field lines, digitised from Ref. [161]. Field lines are considerably closer near to the wire. The field here is stronger, increasing the probability of iS2s. Right: Cathode field lines. The RFR results in some electrons drifting downwards, producing an iS1. Charge drifting upwards produces a smaller S2 or iS2. The latter is more frequent than the former, illustrated by a pie chart in the cathode wire. The rate of iS2s from the cathode is therefore lower than that from the gate.

6.1.2 Gate Events

Figure 6.2 shows the distribution of S1-S2 events of a drift time between 2 µs and 6 µs. Events are split into three distinct populations. At large S1 pulse areas, events are produced by the α decay of ^{210}Po. Events with a greater S2:S1 ratio are products of ER backgrounds in LXe near to the gate wires. At maximum S2:S1 ratio, events are produced <100 µm from the gate wires by gate radiogenics. The majority of these events originate from the β decay of ^{210}Pb and ^{210}Bi, and the α decay of ^{210}Po where the α strikes the wires [133].

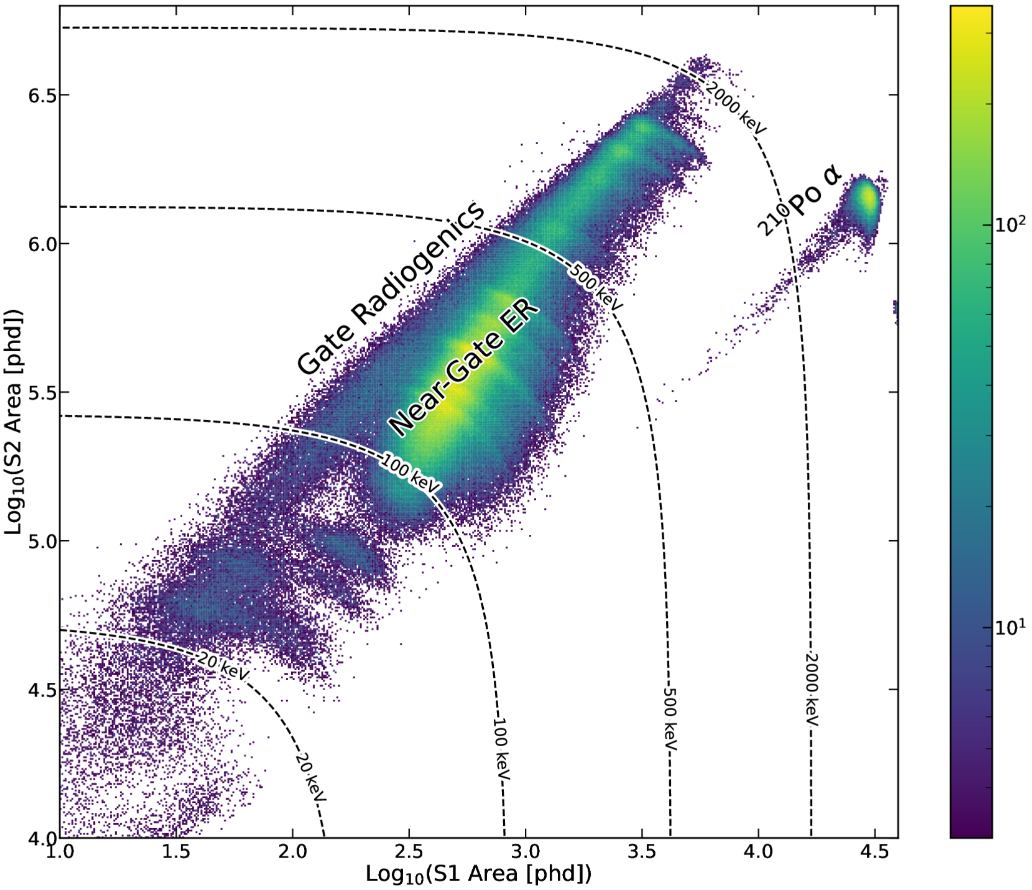


Figure 6.2: S1-S2 distribution of events near the gate array. Populations originating from ^{210}Po α decay are annotated. A broad spectrum of near-gate events, above the ER band, contributes to an S2-Only analysis.

The band of near-gate ER events is similar to that observed in the TPC bulk. However, several additional peaks are present, due to the presence of gammas from the ^{222}Rn decay chain, and events take on a slightly higher S2:S1 ratio due to the elevated field strength. This can be seen in Figure 6.3, in comparison to events from the TPC bulk. At low energies, decays internal to Xe from ^{83m}Kr and ^{125}I are not observed in data near to the gates, which is dominated by low-energy X-rays and gammas from the ^{222}Rn decay chain. At higher energies, however, peaks can be seen in both datasets from ^{129m}Xe and ^{131m}Xe. Several peaks in the ER band near to the gate wire are not present in the ER band in the TPC bulk, and are the product of γ decays in the Rn decay chain.

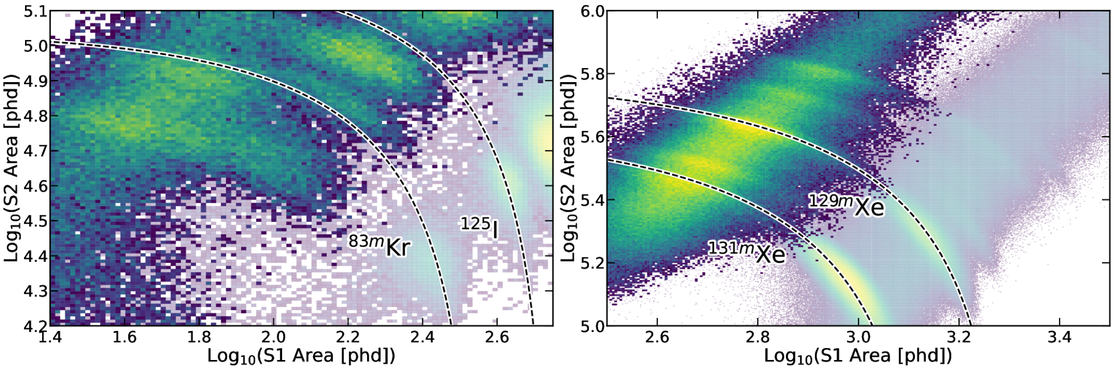


Figure 6.3: S1-S2 distribution of events near the gate array. The S1-S2 distribution of bulk LXe data is displayed at a lower opacity. Mono energetic contours are plotted for the decays of ^{83m}Kr (41.5 keV [162]), ^{125}I (65 keV) [155], ^{131m}Xe (164 keV) [72], and ^{129m}Xe (236 keV) [72]. The latter two are dominant in the band of events near to the gate wires. Additional peaks arise in the near-gate band from ^{210}Pb (46.5 keV), ^{214}Pb (242 keV, 295 keV, 352 keV) and ^{214}Bi (609 keV, 768 keV, 934 keV, 1120 keV, 1238 keV, 1378 keV, 1764 keV, 2204 keV) [127].

The rate of LXe decays decreases when considering distances nearer to the gate wires. This near-gate population is therefore dominant in S1-S2 analysis, but is negligible in iS2 data. In contrast, gate radiogenic events produced on and around the gate wire are less common in S1-S2 analysis, but are dominant in iS2 data [133].

Cathode Events

Figure 6.4 shows the distribution of S1-S2 events of a drift time between 1040 µs and 1050 µs. The same general populations of grid radiogenics, near-grid ER events and large-S1 α decays can be seen in data. The reduced field strength near to the cathode results in an increased rate of recombination, lowering the S2:S1 ratio of all events relative to those from the gate. Peaks also display long vertical 'tails', from the loss of electrons in a given event to the RFR.

As a large fraction of events from the cathode do not drift towards the extraction region, the event rate and iS2 rate are both lower than those from the gate. However, the lower field strength and RFR losses also result in S2 pulses of a lower pulse area.

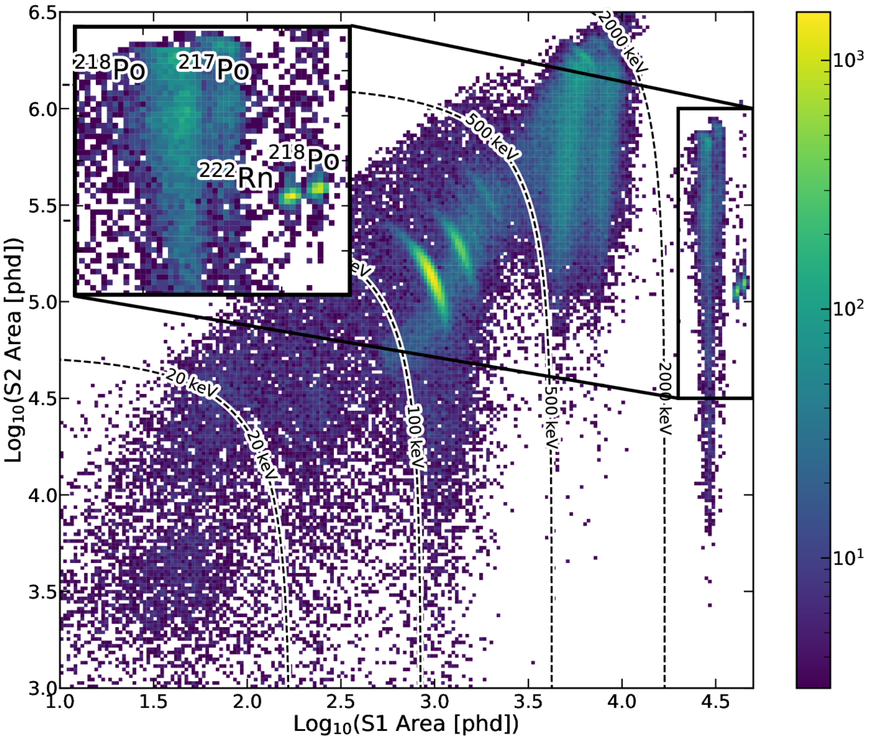


Figure 6.4: The S1-S2 distribution of events near to the TPC cathode. Several α populations are annotated. The majority of events originate from near-wire radiogenic events, which extend into S2-Only data. A 'tail' of events extends to low S2 pulse areas for a given S1 pulse area, because of the RFR.

6.2 S2-Only Analysis

6.2.1 Motivation

The sensitivity of a direct dark matter detector to low-mass WIMPs is heavily dependent on its energy threshold, as discussed in Chapter 2. In LZ, the primary contribution to this threshold is the requirement of the detection of three photons as part of an S1 pulse. In high-mass WIMP searches, this increases sensitivity by allowing for discrimination between ER and NR events, as seen in Figure 3.8. At lower masses, this rejects all iS2 pulses excluding accidental events, which constitute the dominant background. However, below the minimum S1-S2 WIMP search mass of 8 GeV, the requirement of an S1 pulse impacts detection efficiency and therefore WIMP sensitivity.

Removing the requirement of an S1 pulse as part of an electroluminescence-only (S2-Only) analysis increases background from iS2 pulses by over three orders of magnitude, and prevents discrimination of ER and NR events. However, it also allows for sensitivity to events of mass as low as 2 GeV/c^2 in standard analysis, and as low as 80 MeV/c^2 assuming the existence of the Migdal effect in NRs [73].

This concept is demonstrated through simulated signal efficiencies in Figure 6.5. Signal efficiency is two orders of magnitude greater for a dark matter candidate of mass 3 GeV/c^2, and five orders greater for a NR of energy 0.3 keV, allowing for sensitivity to signals of considerably lower energies.

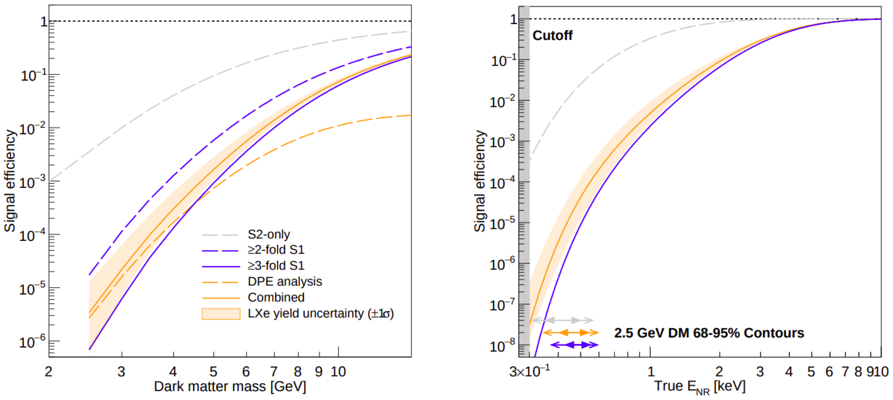


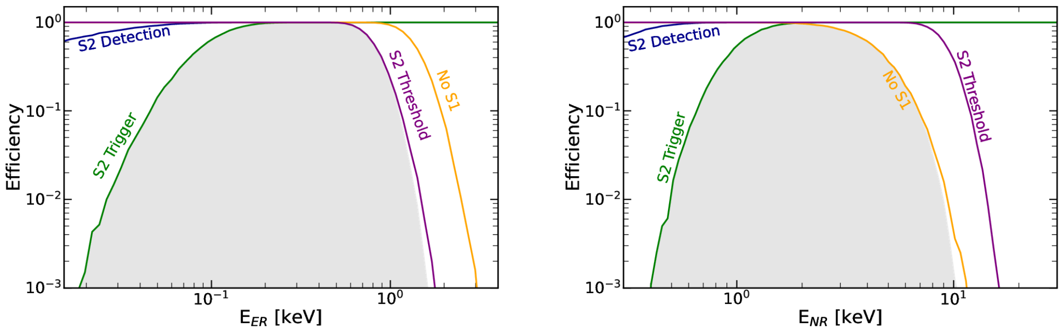
Figure 6.5: NR scatter detection efficiency from simulated data as a function of recoil energy. The signal efficiency in an S2-Only analysis (grey) is considerably higher than that in S1-S2 analysis, particularly for low-energy NRs (right), resulting in sensitivity to lower masses of dark matter (left). Figure taken from Ref. [73].

Standard S1-S2 analysis of LZ data extends to WIMP masses as low as 8 GeV/c^2. However, the total mass range of WIMPs extends to as low as 1 GeV/c^2, as discussed in Chapter 1. This mass range also contains light dark matter candidates, such as asymmetric dark matter, and signals from CEνNS, particularly ^{8}B [145]. By removing the requirement of an S1 pulse, LZ is sensitive to light dark matter candidates of mass as low as 80 MeV/c^2, as illustrated in Figures 2.8 and 2.9.

6.2.2 Backgrounds

The energy range of backgrounds from the TPC bulk accepted in an S2-Only analysis can be calculated using NEST models of the LZ TPC. The baseline energy sensitivity of any direct detection experiment is determined by its trigger efficiency. LZ has been found to trigger with near-100% efficiency on S2 pulses of size 5 e and above [85]. Below this, it is possible for an S2 pulse generated within the TPC to not be detected. For current S2-Only analysis, with this in mind, a threshold of 200 phd has been set for analysis. Other S2 inefficiencies in LZap have been demonstrated to be negligible, as demonstrated in Chapter 5.

Figures 6.6(a) and 6.6(b) show the projected efficiency for ER and NR events respectively in an S2-Only analysis. The lower limit on energies is defined by the 5e threshold imposed by S2 trigger efficiency, whereas the upper limit is imposed through a combination of the requirement of no S1 pulse and an S2 area threshold of 2000 phd. ER backgrounds of energies 0.1 keV to 1 keV are accepted in this analysis, as are NR backgrounds and signals of energies 1 keV to 10 keV.



(a) **(b)**

Figure 6.6: Detection efficiency of (a) ERs and (b) NRs in LZ in an S2-Only analysis, modelled using NEST [119]. Contours indicate the constraints placed by S2 pulse detection (blue), triggering on this S2 (green), not detecting an accompanying S1 (yellow) and this S2 being below a threshold of 2000 phd (purple). The grey shaded region in both plots indicates the overall efficiency of detection, with all factors taken into account. ER events of energy ∼0.1 keV to 1 keV and NR events of energy ∼1 keV to 10 keV are considered in this analysis.

Mono energetic peaks in S2-Only data are produced by auger electrons from two ER sources: ^{37}Ar, which produces L-shell auger electrons of energy 0.27 keV [163], and ^{127}Xe, which produces up to five auger electrons of total cascade energies of 1.1 keV and 186 eV [156]. The broad spectrum of beta emission also reaches energies low enough for an S2-Only analysis, including backgrounds from ^{214}Pb and ^{212}Pb from Rn decay, and ^{85}Kr from surface contamination [128]. However, the rate of these events in the S2-Only ROI is low.

The majority of NR backgrounds in LZ produce events of energies above 10 keV. As a result, the primary background from NR events is CEνNS from ^{8}B at threshold. This leaves iS2 events, particularly iS2s from the gate and cathode, as the dominant background in an S2-Only analysis.

Background events for an S2-Only search have been simulated in BACCARAT [159] and can be seen in Figure 6.7. Radiogenic events from the gate and cathode have been modelled in detail in a separate body of work, and are in close agreement with events seen in data [133]. The S2-Only distribution of a 2 GeV/c^2 WIMP of cross section $1 \times 10^{-46}\,\mathrm{cm}^2$ is simulated using NEST [121], with an interaction rate calculated using WIMPrates [84]. Without cuts, this distribution is dominated by grid radiogenics down to a pulse area of 200 phd, the current threshold for S2-Only analysis.

The remainder of this chapter details the selection of gate and cathode events, and the tuning of cuts to reject these backgrounds in S2-Only analysis. Rejection of these events is the deciding factor as to LZ's sensitivity to low-mass WIMPs [107, 164].

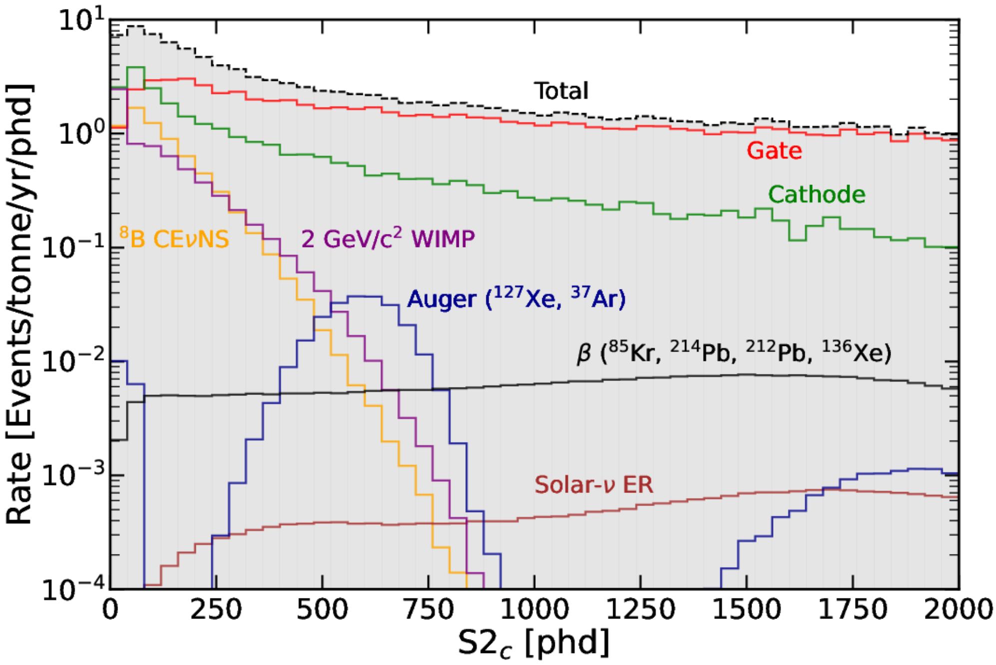


Figure 6.7: Distribution of background populations in an LZ S2-Only search, simulated in BACCARAT [159], and grid background populations from independent empirical simulations [133]. The projected event rate from a 2 GeV/c^2 WIMP is included, generated using NEST [121] with event rate calculated using WIMPrates [84]. Gate and cathode radiogenic events are a dominant background for all S2 pulse areas.

6.3 Event Selection

6.3.1 Classifying Events

The pulse and event classification software LZap is tuned using S1-S2 events in the WIMP ROI. This results in a higher rate of event mis-classification in regions outside of the ROI, as discussed in Chapter 5.

In order to prevent this from impacting S2-Only analysis, a simple event selection pipeline has been designed for S2-Only analysis in LZ, which is agnostic to event classification in LZap.

For an event to be considered S2-Only, there must be no S1 or S2 pulse present up to Tµs before the largest S2 in an event window, where T is the maximum physical drift time. This has been demonstrated to result in minimal signal loss in Chapter 5.

This criterion for event selection will contain some MS events. In order to reject these, a simplified form of the S2 prominence checks discussed in Chapter 5 is applied to data. Events containing more than one prominent S2 are removed from analysis.

The tuning of cuts to reject grid radiogenic backgrounds requires a selection of gate and cathode events, made using an event's drift time. For this work, events are selected using the criterion outlined above, but where an S1 has been identified within the maximum drift time of an S2. The time between the S1 and S2 can then be used to identify gate, cathode, and signal events.

6.3.2 Calibration Data

The AmLi calibration sources in LZ produce neutrons of a spectrum of energies up to an endpoint of 1.5 MeV, as outlined in Table 3.1. This produces a spectrum of S1-S2 events with S2 pulse areas within the S2-Only ROI. For this reason, AmLi data is chosen as a signal dataset for training grid cuts in this chapter.

As AmLi sources are located in the three CSDs inside of the TPC walls, events are biased towards the walls of the TPC, compared to a low-mass WIMP signal, which would be evenly distributed. To account for this, events are given a weight, inversely proportional to the number of events in data at a given position and S2 pulse area.

In order to ensure cuts are not biased to AmLi calibration data, a cross check can be performed using CH_3T calibration data. CH_3T is the lowest energy calibration source injected into LZ. However, the event rate from CH_3T drops below an S2 pulse area of 2000 phd, the maximum S2 pulse area in the S2-Only ROI. For this reason, events are accepted of S2 pulse areas below 4000 phd. Events of pulse area up to 2000 phd are used for the training of cuts in S2 pulse width, whereas events of up to 4000 phd are used for training a BDT.

6.3.3 Preliminary Cuts

In Chapter 4, it was demonstrated that cuts in S2 TBA and XY χ^2 are capable of rejecting iS2 events, particularly those from the glue ring and near to the liquid surface. As these cuts are likely to be implemented in an S2-Only analysis, they have been applied to data, to account for the impact they will have on the number and distribution of gate events. Cuts to live time, to reject periods of low data quality, are also applied [85].

In S1-S2 analysis, a fiducial volume is applied to data, in order to reject ER backgrounds produced by the TPC walls. The decision has been made for this analysis to not apply a radial fiducial volume, as the dominant grid radiogenic background is uniformly distributed in (x, y). However, the final analysis of S2-Only data will likely introduce a small radial cut, in order to reject regions of the TPC wall where measurements of signal loss are less precise.

6.4 Rejection of Above-Anode Events

In S1-S2 analysis, approximately 5% of events are rejected due to their low drift time, in order to reject both gate radiogenic events and a background of what are referred to as above-anode events. Without a dedicated cut, a large fraction of above-anode events leak into the selection of gate radiogenic events, as well as of signal events at low drift times. This biases the tuning of cuts at low S2 pulse areas.

As described in Chapter 3, an S2 is produced by electroluminescence in the TPC extraction region. Upon reaching the gate grid wires, electrons enter a region of increased electric field strength, and continue to travel towards the anode grid wires. While doing so, the electrons pass into a thin layer of GXe at the top of the TPC. As GXe has a lower scintillation threshold than LXe, it is then possible for these electrons to electroluminesce as they drift towards the anode grid. The field strength in this region determines the size of S2 detected, and the gate-anode separation determines the width of the S2 pulse in time.

However, it is also possible for interactions to take place above the gate wires. Events near to the extraction region immediately electroluminesce upon an interaction. This produces a population of 'liquid surface' events where the S1 and S2 are not resolved, discussed in Chapter 4. However, events above the anode grid will instead drift down towards the anode wires. These events are classified as single-scatter events as if they originated from the TPC bulk. An illustration of this region, and the pathologies of events produced in each sub-volume, can be seen in Figure 6.8.

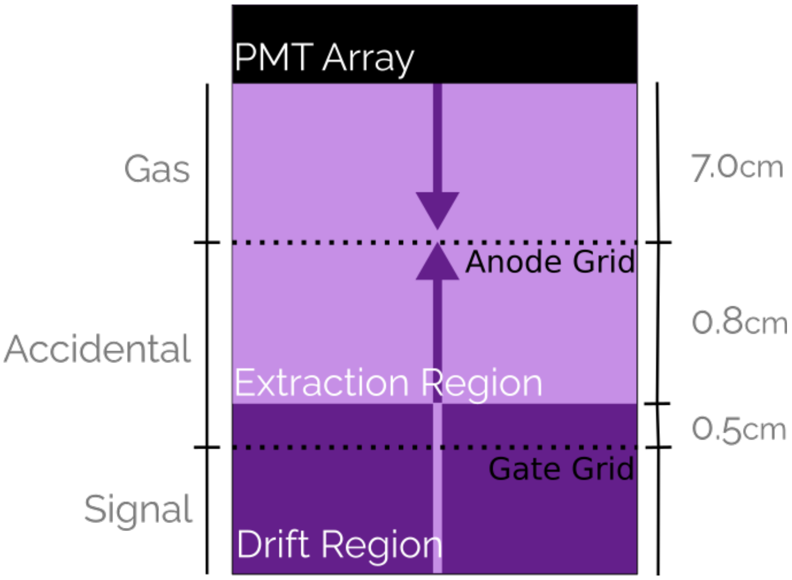


Figure 6.8: Illustration of the top of the LZ TPC, based on Figure 3.1. Arrows indicate the diffusion of ionisation electrons in that region towards the anode. Three distinct populations are annotated: a region of ordinary LXe events (Signal), a region where an S1 and S2 are not resolved (Accidental), and a region of above-anode background events (Gas).

This section details the design of a novel cut using unsupervised learning to reject low-energy above-anode events. This then allows for a cleaner selection of gate and signal events at the top of the TPC, for the tuning of S2-Only cuts.

6.4.1 GXe Event Properties

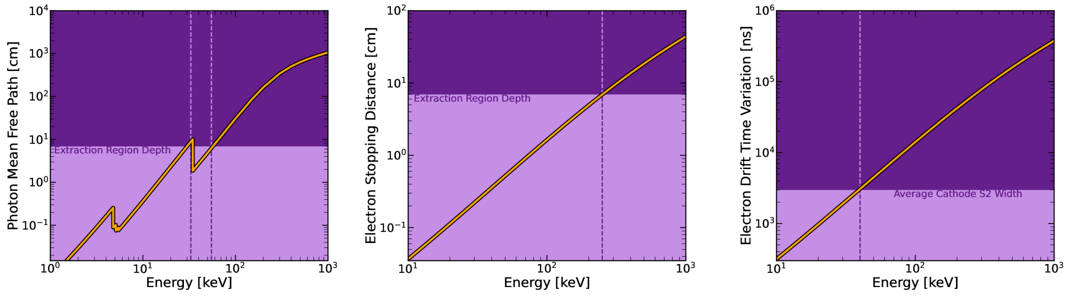
Events from above the anode cannot be rejected using a flat 1-dimensional cut without significant signal loss or background leakage. However, they are distinct in several parameter spaces, due to the different properties of GXe and the position of events relative to the extraction region. These properties can be used in combination to improve the rejection power of a cut targeting above-anode events.

Pulse Areas

Above-anode events are identical in origin to backgrounds in the LZ bulk. However, the lower density of GXe results in less ionisation and more scintillation for a given energy of interaction. This results in events occupying a region in S1-S2 space below the NR band.

Events detected above the anode are also of a lower energy. As GXe is less dense than LXe, its energy threshold is lower, increasing the rate of low-energy signals. This low density also increases the stopping distance for high-energy backgrounds. This increases the probability of these events being mis-classified as MS events, or of entering the extraction region and producing an iS2.

This second concept is demonstrated for γ and β sources in Figure 6.9. 6.9(a) shows the mean free path of high-energy photons, with the height of the above-anode volume annotated. Above an energy of 30 keV, γ radiation produced in this volume is most likely to enter the extraction region or strike the liquid surface, producing liquid surface iS2s. In Figure 6.9(b), repeating this calculation for electrons demonstrates that the equivalent energy for β radiation is at 200 keV. However, when this variation in (z) is scaled to variation in time using the mean GXe drift velocity, the time variation of S2s above 40 keV is found to be significantly higher than the mean S2 pulse width of cathode radiogenic events, as seen in Figure 6.9(c). Events above this energy are therefore likely to be mis-classified as MS events.



(a) **(b)** **(c)**

Figure 6.9: a) Photon attenuation length in GXe vs energy, generated with XCOM [120]. The energy at which photons interact within the extraction region is annotated. Above-anode gamma radiation is likely to escape the GXe region at energies above 30 keV to 60 keV. b) Electron stopping distance vs energy, generated with ESTAR [165]. The energy at which betas escape the extraction region is annotated. Above-anode beta radiation is likely to escape the GXe region at energies above 200 keV. c) Electron stopping distances, scaled by GXe drift velocity into time variation. The mean cathode width is annotated. Above-anode beta radiation is likely to be classified as a multiple scatter event at energies above 40 keV.

Top-Bottom Asymmetry

An above-anode S1 pulse is produced very near to the top PMT array. This produces a higher value of S1 TBA, a parameter discussed in the context of accidentals in Chapter 4. However, GXe events at a higher drift time are produced directly underneath the top PMT array. Events directly between PMTs produce S1s with significantly lower light collection efficiency (LCE), as scintillation light reflects off the PTFE between PMTs, resulting in a lower S1 TBA. Similarly, an event directly underneath a PMT will have a significantly higher value of S1 TBA. This is illustrated in Figure 6.10.

Electroluminescence in above-anode events takes place next to the anode wires, as opposed to throughout the extraction region. This results in a higher S2 TBA in above-anode events than that measured in the TPC bulk. The difference between measured TBA values is lower than that observed in S1s, as the separation between the bottom and top of the extraction region is lower than the separation of the GXe layer from the TPC bulk. However, S2 LCE is constant in drift time.

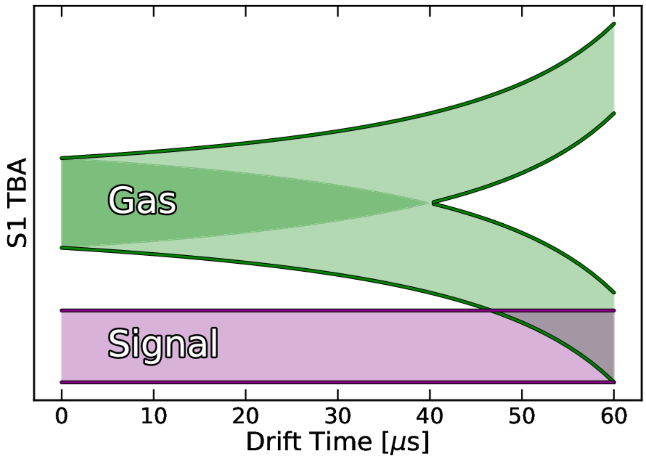


Figure 6.10: Illustration of the relationship between S1 TBA and drift time in above-anode (green) and bulk (purple) data. The two distributions are distinct until high drift times, where above-anode S1 TBA diverges into data underneath PMTs (high TBA) and between PMTs (low TBA). The latter is more difficult to distinguish from LXe data.

The TBA of a small pulse is more susceptible to statistical fluctuations, as described in Chapter 4. This results in weaker discrimination power for small S1s and S2s. As S2s are disproportionately smaller in above-anode events, this has the largest impact on S2 TBA. This is illustrated in Figure 6.11.

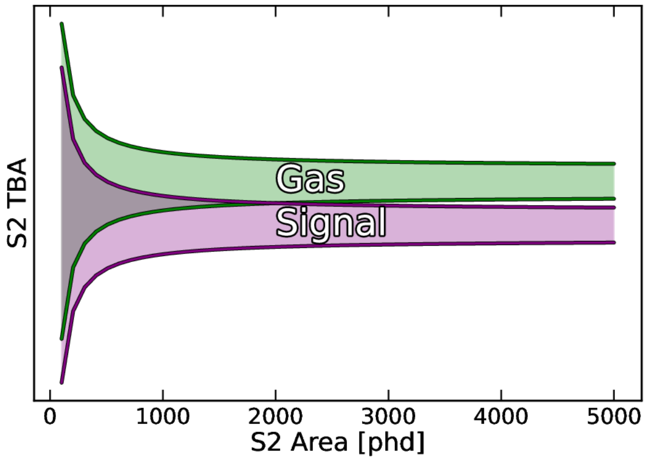


Figure 6.11: Illustration of the relationship between S2 TBA and S2 pulse area in above-anode (green) and bulk (purple) data. Above 1000 phd, the two distributions are distinct. Below this, above-anode events are more difficult to distinguish from LXe data.

Pulse Widths

As described in Chapter 4, the width of S2 pulses in a TPC follows a $\sqrt{Mt + C}$ relationship with varying drift time t. The width of pulses is therefore dependent on the electron pulse width C, as well as the diffusion M.

The single electron width in LXe data is determined by the transport of electrons from the gate towards the anode. In contrast, ionisation above the anode drifts towards and is already in close proximity to the anode. Ionisation which travels directly towards the anode wires accelerates rapidly, such that electroluminscence is detected instead as scintillation light, referred to as a stinger pulse. Events which travel between the anode wires instead enter the higher field strength of the extraction region, and electroluminsce as they circle the anode wires. This produces an S2 pulse of significantly lower width than that observed in LXe, resulting in a lower value of C. This process is illustrated in Figure 6.12.

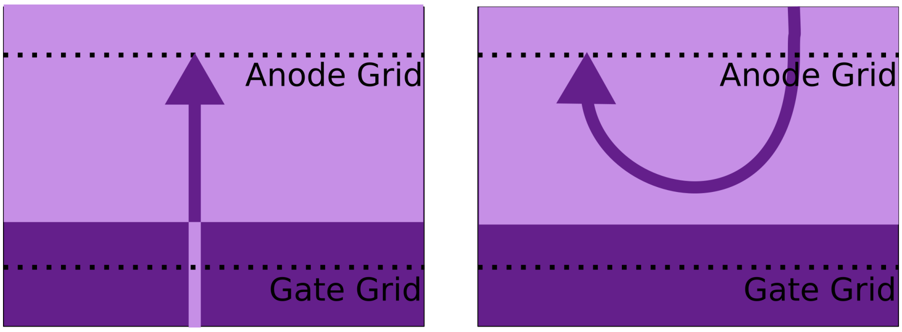


Figure 6.12: Illustration of the electroluminescence of events in the TPC. Left: events originating from the LXe are accelerated through the gas from the gate array to the anode array, emitting photons in the process. Right: events originating from the GXe are accelerated back towards the anode, resulting in a lower pulse width.

The diffusion of electrons in noble gases is dependent on the longitudinal diffusion coefficient D_L. The lower density of GXe results in a significantly higher diffusion coefficient [166]. The width of pulses from events originating in GXe therefore increases more with drift time than those in LXe, resulting in a higher value of M.

As detailed in Section 6.4.1, the density of GXe above the anode is such that events of a high energy will be identified as multiple scatters. Before this point, they will instead be classified as unphysically broad S2 pulses. This will result in a diffuse distribution of broad S2 pulses, in a similar manner to that seen in Figure 5.17.

An illustration of the width-drift time relationship of GXe and LXe events can be seen in Figure 6.13. Events at low drift times near to the gate overlap with the diffuse spectrum of broad above-anode S2s, while events nearer to 60 µs overlap instead with the main width-drift distribution.

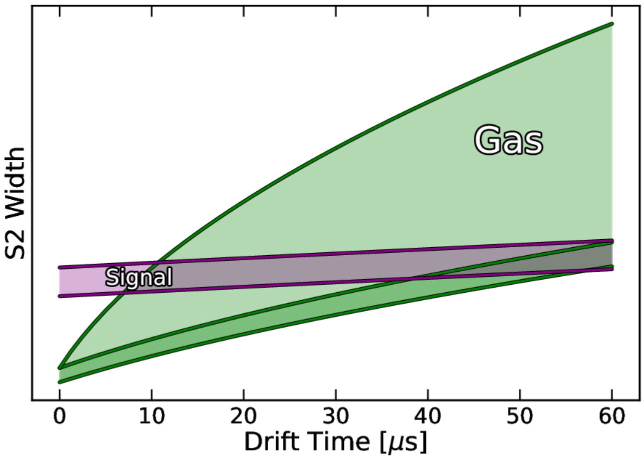


Figure 6.13: Illustration of the relationship between S2 pulse width and drift time in above-anode (green) and bulk (purple) data. Above-anode data is separated into SS (light) and MS pileup (dark) events.
The majority of data is distinct in the two distributions until high drift times. However, MS pileup events mis-classified as a SS are harder to distinguish from LXe data.

Bad Area

The bad area of an event can be defined as the integrated charge of all pulses that are not the S1 and S2 within a given event window. This parameter was used in LUX to reject E-Trains [117]. While the primary tool to reject this background in LZ has been replaced with a cut in live-time following S2 pulses [133], bad area remains a useful tool in rejecting events during periods, or within TPC volumes, of low data quality.

The S2 from an event produced in GXe begins once it has entered the extraction region, where the increased field strength increases electroluminescence. However, a small amount of electroluminescence is still observed from events as they approach the anode wires [167]. This results in a high photon rate before the S2 in above-anode events, in comparison to those in the TPC bulk.

A bad area parameter has been designed using the sum of pulse areas up to 70 µs before an S2. Events of pulse area greater than 5 phd are not considered, to prevent the inclusion of S2 or SE pulses in this metric. The bad area in a 70 µs window ahead of the S2 is subtracted from this measurement as a baseline, in order to ensure the metric is consistent during periods of elevated photon rate in the TPC. An illustration of this can be seen in Figure 6.14.

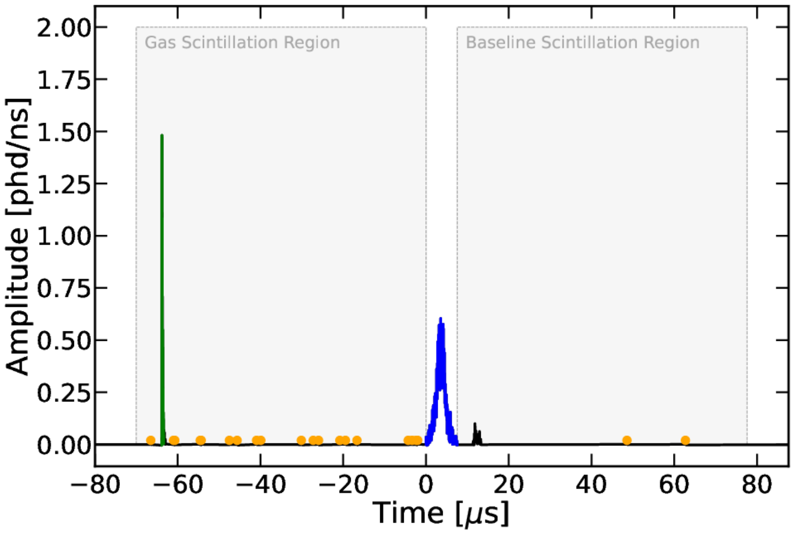


Figure 6.14: An example single scatter from above the TPC anode. Shaded regions behind and ahead of the S2 (blue) indicate the region used to calculate bad area and the baseline subtracted from it, respectively. The photons included in the calculation are highlighted in orange. The small pulse after the S2, from a single electron, is not included in this calculation.

An illustration of the bad area relationship of GXe and LXe events can be seen in Figure 6.15. In signal data, bad area decreases with increasing pulse area, as larger S2s produce more delayed emission in the 70 µs following them, resulting in a larger baseline subtraction. The distribution of bad area in above-anode events is mostly distinct: however, the bad area of events of a lower drift time is reduced, making it more difficult to discriminate above-anode events from events in the TPC bulk.

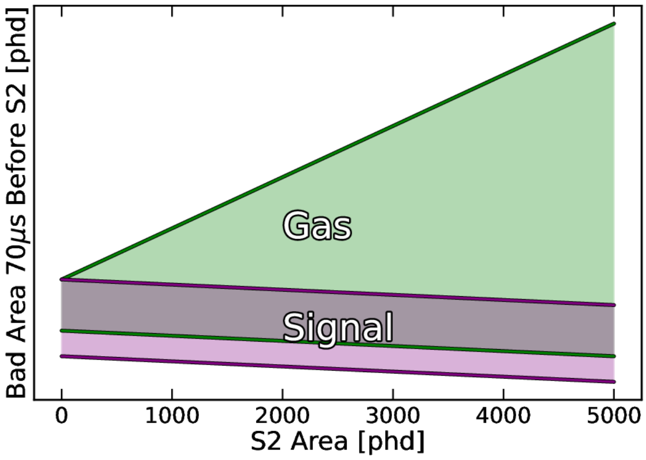


Figure 6.15: Illustration of the relationship between good and bad area in above-anode (green) and bulk (purple) data. The two distributions are mostly distinct. However, above-anode events of a lower drift time will produce less bad area, and are harder to distinguish from LXe data.

Selection of Above-Anode Events

Events in GXe and LXe are distinct in several parameter spaces, but overlap in each individual parameter. This results in any simple cut having considerable signal loss or background leakage. For example, a cut in S2 TBA and S2 pulse area is applied to WIMP search data to reject GXe accidentals, based on the illustration of the signal contour in Figure 6.11 [154]. However, below 2000 phd, the rejection power of this cut is as low as 50%, decreasing for smaller S2s.

A cut in multiple parameter spaces is capable of identifying above-anode events with significantly lower background leakage. As this cut is to be applied in several parameters simultaneously, without a clean selection of signal and background data, this is an analysis which would benefit from unsupervised machine learning techniques. In unsupervised learning, a dataset is given to an algorithm without truth values, and the algorithm clusters data into populations such that covariance in each cluster is minimised.

For this analysis, a Gaussian Mixture Model with 25 clusters has been chosen to reject above-anode events. The nature of a GMM is described in Appendix A, alongside a brief illustration of its strengths.

A mixture of events of drift times below 60 µs are used to train a GMM, using AmLi and CH_3T calibrations as well as WIMP search data. The GMM takes as inputs the FWHM of an event's S2 pulse, its S1 and S2 TBA, its bad area, drift time, as well as the S1 pulse area, S2 pulse area, and XY position. Events are then sorted into 25 clusters. Clusters observed to follow the expected relationships for GXe data are identified and merged, as are clusters following trends observed in LXe data. Remaining above-anode events in LXe data are rejected through the use of the standard S2 TBA cut, as well as a linear cut in bad area.

Results

Events classified as originating from the TPC bulk can be seen in multiple parameters in Figure 6.16. All relationships illustrated in this section are observed, and events are located predominantly towards the TPC walls. Events generally follow the ER and NR band in S1-S2 space, with a small population of accidental-like events at large S1 and small S2 pulse area, produced from the TPC walls.

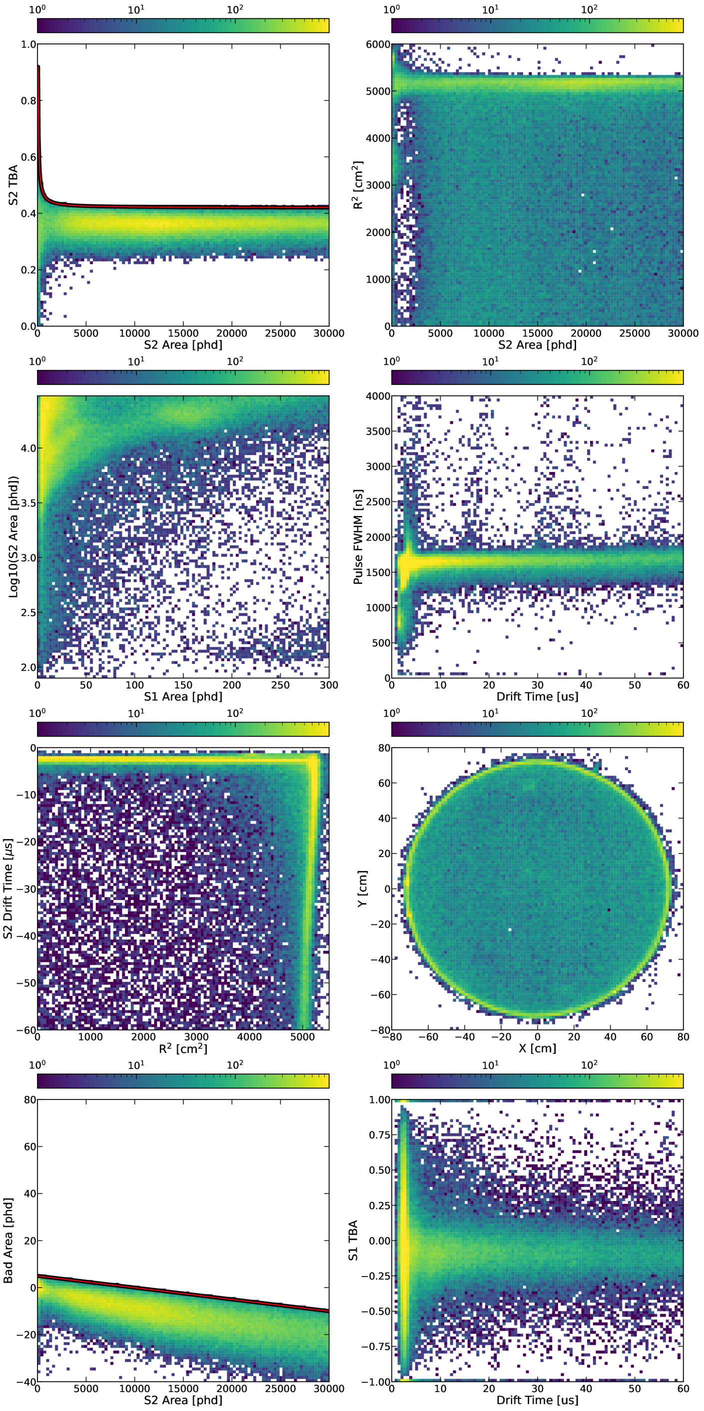


Figure 6.16: Events with a drift time below 60 µs in SR3 and calibration data identified as LXe events by a GMM classifier. Data follows all expected relationships outlined in Section 6.4

Events produced above the TPC anode can be seen in multiple parameters in Figure 6.17. All relationships discussed in this section are observed in data.

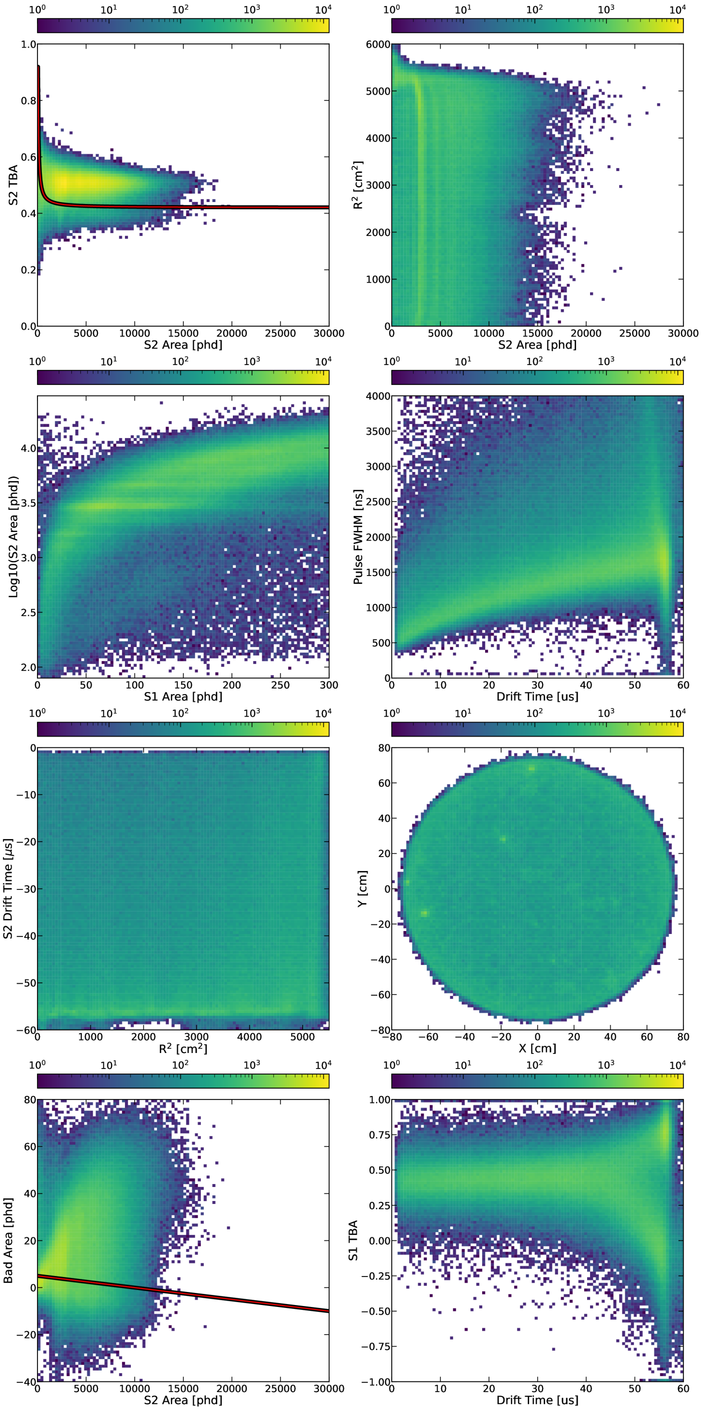


Figure 6.17: Events with a drift time below 60 µs in SR3 and calibration data identified as GXe events by a GMM classifier. Data follows all expected relationships outlined in Section 6.4.

The S2 pulse area of above-anode events forms several distinct bands. These correspond to particular energies of radiogenic background in the GXe volume. Figure 6.18 demonstrates this, by showing the distribution of S2$_c$ against radius squared and drift time for above-anode events. Peaks in S2$_c$ are constant in radius. A lower value of S2$_c$ at high radius is the result of corrections from LXe data being applied to GXe data. Peaks are also constant in drift time. Event rate increases at high drift time, suggesting that these backgrounds are produced by mono energetic radiogenic events in the PMTs.

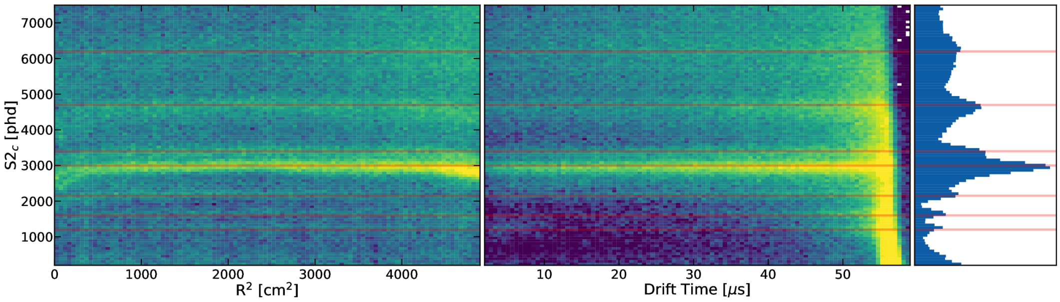


Figure 6.18: Above-anode S2$_c$ against TPC position. Left: S2$_c$ against R^2. Peaks remain constant, suggesting that they do not originate from systematics or the TPC wall. Centre: S2$_c$ against drift time. Peaks increase in counts at high drift times, suggesting that backgrounds originate from the top PMT array. Right: A histogram of S2$_c$.

The value of g_1 and g_2 for energy reconstruction differs in the GXe region compared to the TPC bulk. The distribution of S1 pulse areas of above-anode events is also very broad, due to the variation in light collection efficiency in the volume. As such, energy cannot be precisely reconstructed using the method outlined in Section 3.1.3.

However, some information can be attained by reconstructing the energy of an event from the area of the S2 pulse alone. This can be scaled to an approximate energy through the measurement of the gamma peak in ^{83m}Kr data at 41.5 keV and the peak from internal capture at 32.1 keV, an approach used for energy reconstruction in the GXe NEXT detector [162]. Reconstructed above-anode event energies from WIMP search data and calibrations can be seen in Figure 6.19.

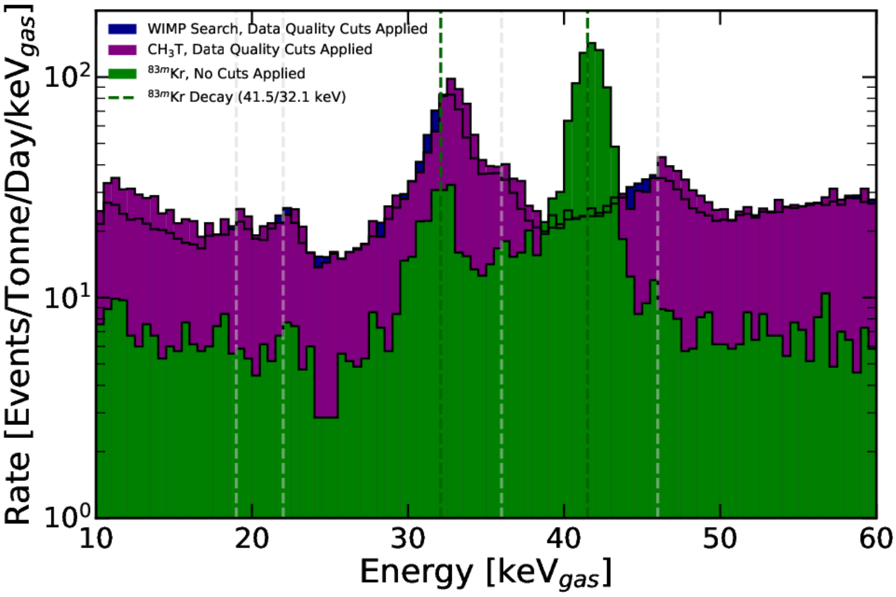


Figure 6.19: Energy distribution of events in the above-anode region for WIMP search data, CH_3T calibration data, and ^{83m}Kr calibration data, in keV$_{gas}$. A pair of vertical lines (green) annotate the 32.1 keV and 41.5 keV peaks of ^{83m}Kr data, used to linearly scale S2 pulse area into keV$_{gas}$. Backgrounds in the above-anode region appear to originate from sources below 50 keV, with distinct peaks at 19, 22, 36, and 46 keV, all annotated. Kr rates have been artificially scaled down, to improve legibility.

The rate of above-anode events is similar in calibration and WIMP search data, with a small low-energy excess during calibration. The majority of data is located at around 32 keV to 36 keV, the expected energy of Xe X-rays [168] also detected in GXe by NEXT [169]. Peaks are also present in data at energies of approximately 19, 22, and 46 keV, produced by x-rays from PMT quartz windows. Above energies of 60 keV, event rate drops exponentially, as expected.

6.4.2 Above-Anode Events in S2-Only Analysis

The absence of an S1 signal makes above-anode events difficult to identify in S2-Only data. However, the two most effective metrics at identifying these events in S1-S2 data, S2 TBA and bad area, can still be used to identify above-anode events. While the rejection power of these cuts in isolation is not as high, they can still be used to mitigate background events, as well as to evaluate the impact of above-anode events in an S2-Only analysis.

Figure 6.20 shows the iS2 data from Section 4.1.2 rejected by the S1-S2 cut in S2 TBA, as well as by a cut in bad area. 559,406 events pass both cuts, the majority of which take on pulse widths indicative of grid backgrounds. 9174 events fail the S2 TBA cut, the majority of which take on a low pulse width indicative of liquid surface events. 5180 events are rejected by a cut in bad area which would otherwise be accepted by a cut in S2 TBA, with the majority of these events at small S2 pulse areas.

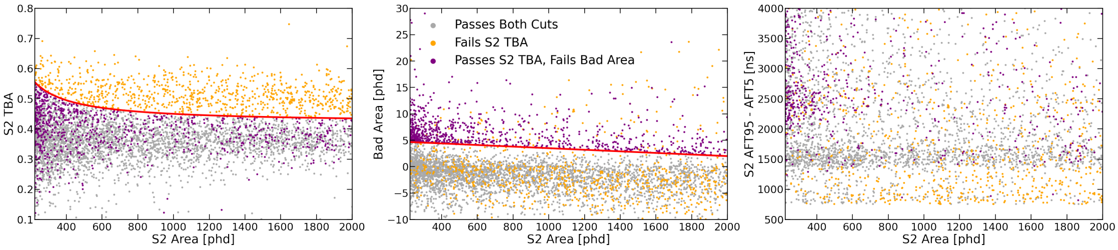


Figure 6.20: iS2 data failing a cut in S2 TBA (orange), passing S2 TBA but failing a linear cut in bad area (purple), and 10% of data passing both cuts (grey), in S2 TBA (left), bad area (centre), and width (right) space. The majority of high-S2TBA events rejected in iS2 data originate from the liquid surface, with low bad area and unphysically low S2 width. Far fewer events pass the S2 TBA cut and fail bad area, indicating that the number of above-anode events in S2-Only data is very low until threshold.

Above-anode events in iS2 data are less frequent than in S1-S2 data, due to the low threshold on scintillation drastically lowering the probability of three photons not being detected for an interaction. Above-anode events in this dataset passing a cut in S2 TBA but failing a cut in bad area are disproportionately located between PMTs and underneath inactive PMTs, where light is not collected. The rate of liquid surface events, which pass a cut in bad area due to their immediate electroluminescence but fail a cut in S2 TBA, is a factor of two higher. Both backgrounds are effectively rejected through a combination of cuts in S2 TBA, bad area, and pulse width. The rate of remaining events is negligible compared to the rate of radiogenic events from the gate and cathode arrays.

6.5 Rejecting Grid Radiogenics Using Pulse Width

6.5.1 Pulse width of gate and cathode events

Chapter 4 and Section 6.4.1 both discuss how S2 pulse width is dependent on the diffusion of electrons through the TPC. This can be exploited in an S2-Only analysis to reject events from the gate and cathode, which typically have the lowest and highest S2 pulse widths respectively. The distribution of widths of S1-S2 events from these populations can be seen in Figure 6.21. Gate events and cathode events can be rejected with a cut on events of S2 pulse widths below and above a threshold, respectively.

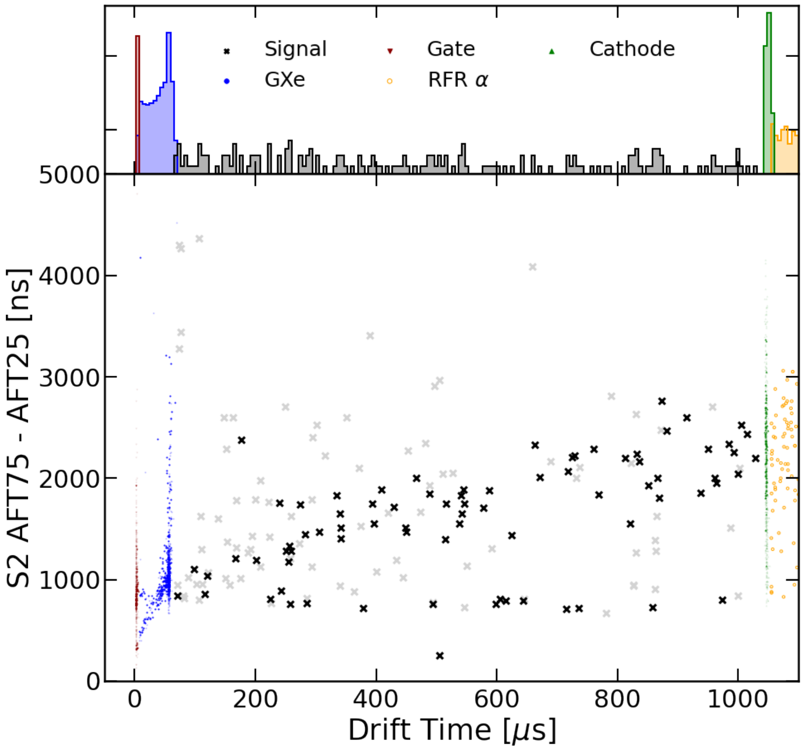


Figure 6.21: Sample S1-S2 events from WIMP search data, distributed in S2 pulse width and drift time. The drift time of an event is used to highlight populations of bulk events (black), gate (red) and cathode (green) radiogenics, above-anode events (blue), and cathode αs (orange). S2s from the gate and cathode are on average narrower and broader than S2s from the TPC bulk. Signal S2s of size 5 e to 20 e are shown at a lower opacity.

Events from the cathode take on a very broad distribution of widths, as their pulse shape is diffusion-dominated. Cathode S2s, as a result, take on very Gaussian-like shapes.

The width of events at low drift times, near to the gate, can be seen in Figure 6.22. A fraction of events in this region are predicted to be from Xe decays, as their drift time indicates they are further from the gate, and they follow the anticipated width-drift relationship discussed in Chapter 4: these are removed from the selection of gate events through a flat cut on pulse widths and drift times.

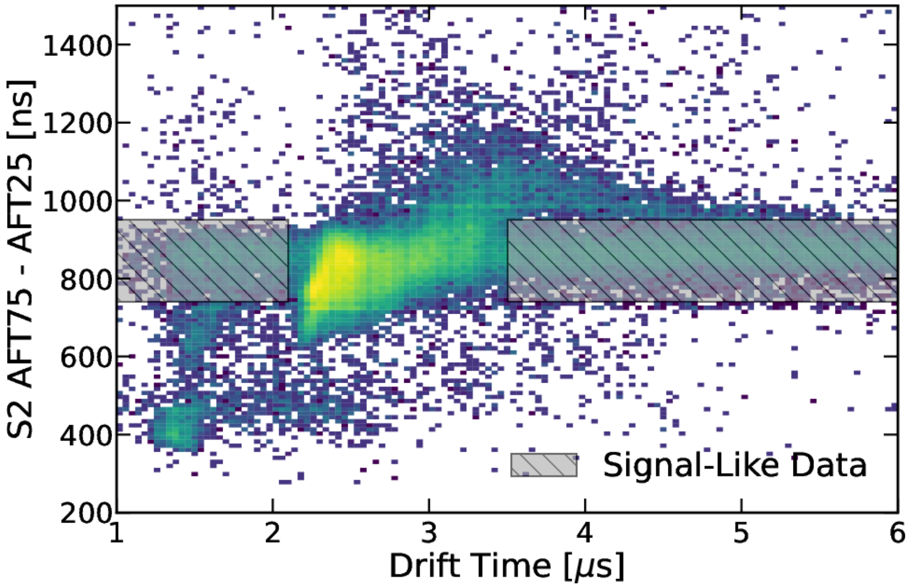


Figure 6.22: Width of S1-S2 events with a drift time below 6 µs. Events within the shaded boxes predominantly originate from near-gate ER events, and are removed from the selection of gate events.

A population of events are present in data between a drift time of 1 µs and 2 µs. These events originate from the glue ring and weir reservoir. These are not removed from the gate selection, as both are backgrounds in iS2 data.

Events in the region 2 µs to 2.5 µs are produced by internal gate radiogenics. Events vary in pulse width and drift time due to grid deflection, where electrostatic forces draw the gate wires towards the liquid surface, resulting in a spread of pulse widths and drift times radially across the TPC. This is illustrated in Figure 6.23.

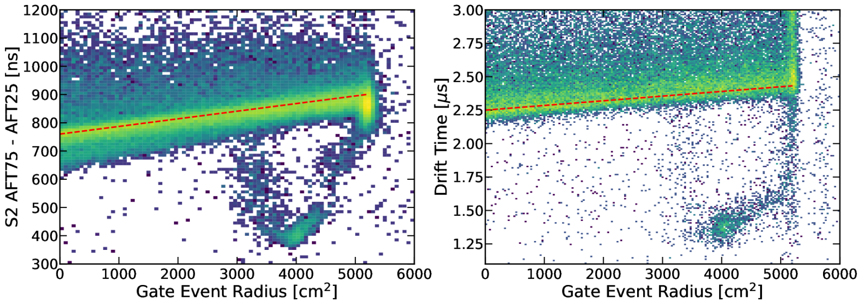


Figure 6.23: Width and drift time of gate radiogenics against radial position. A red dashed line illustrates the typical relationship observed in data. Grid deflection causes the width and drift time of gate radiogenic events to vary radially, introducing variations to S2 pulse shape.

Grid radiogenic events are also observed up to a drift time of 6 µs at higher pulse widths. These events are produced by field fringing, in regions underneath the gate wires where the electric field is non-uniform. Figure 6.1(a) illustrates the concept of fringing around a wire. Around the bottom of the gate wire, field lines are less uniform and homogeneous. The field inhomogeneity around the wires is increased by grid deflection.

When an interaction occurs underneath the gate wires in a region with fringing, the resultant electron cloud is distorted such that it produces an odd S2 pulse shape. When drifting, electrons experience drastic changes in electric field, which cause their velocity to vary, increasing the drift time and pulse width of some S2s. As interactions travel further from the gate wire, field lines become more uniform. Although field lines are longer, resulting in higher drift times, the field uniformity results in electron diffusion comparable to S2s from the TPC bulk.

The equations of pulse width for varying S2 area and drift time, developed in Chapter 4, can be used to model grid radiogenic S2s by selecting a fixed drift time and varying S2 pulse area. Figure 6.24 overlays the parametric distributions for gate, cathode, and signal events. This is in agreement with previous simulations of S2-Only events [118], as well as with S1-S2 events from the cathode and TPC bulk.

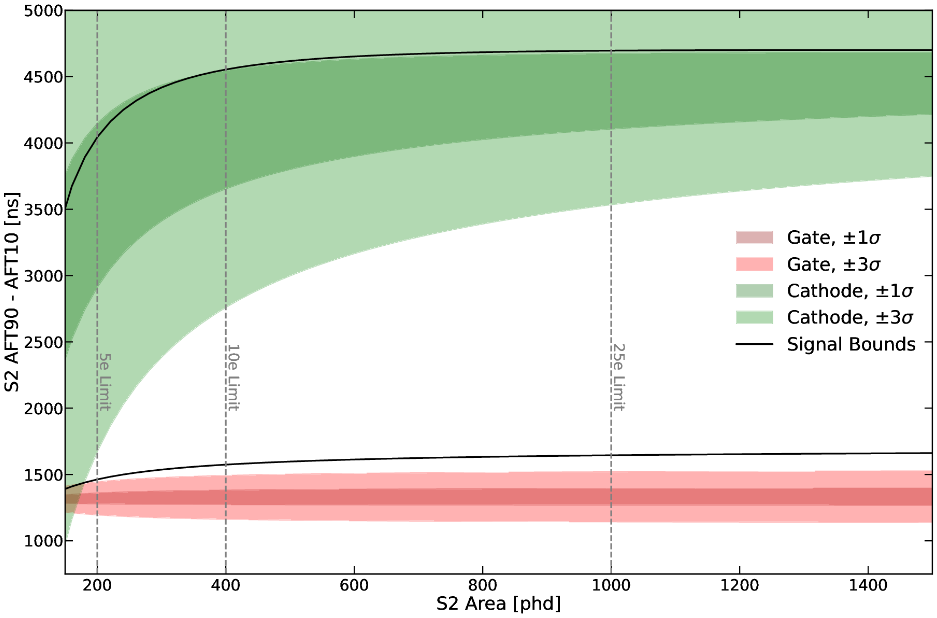


Figure 6.24: S2 Area vs width for LXe signal (black lines), cathode radiogenics (green), and gate radiogenics (red). Dark and light bands indicate 1σ and 3σ bounds.

This contour highlights that grid events in LZ can be rejected through a width-based cut with the loss of a fraction of signal events. The majority of signal is lost in the rejection of cathode events, which are more diffusion-dominated and therefore take up a larger portion of the parameter space. However, this contour accurately reflects the distribution of all cathode events. Gate events are much more densely populated, and the majority can be rejected through the application of a width cut which is constant with respect to S2 pulse area. However, S2 pulses affected by field fringing are not accounted for by these contours, and will require a more conservative cut to be rejected from analysis.

6.5.2 Quantifying Sensitivity

The projected distribution of pulse widths vs area, shown in Figure 6.24, demonstrates that a cut in widths is capable of improving the sensitivity of an S2-Only analysis to signal. However, in order to define the position of such a cut, a metric must be defined which allows the sensitivity of data to low-mass WIMPs to be quantified.

Figure 6.7 shows the simulated rate of gate, cathode, and WIMP events in an S2-Only analysis. In the range of 200 phd to 2000 phd, the rate of events from these sources is estimated to be 603, 1906, and 76 events/tonne/year, respectively.

The ideal signal acceptance and background rejection of S2-Only analysis depend not only on the rate of signal and background events, but also on the statistical inference of data. In one scenario, using a profile likelihood ratio, an assumed background model is subtracted from data. In this context, sensitivity scales as $\frac{s}{\sqrt{b}}$, where s and b denote signal and background rate respectively, as the mean of the background spectrum can be subtracted to leave only background fluctuations. In the second scenario, where background cannot be subtracted and a simple Poisson analysis is used, sensitivity instead scales as $\frac{s}{b}$. Sensitivity will be quantified in both of these metrics, with the final choice of cut dependent on the choice of statistical inference.

The interaction of dark matter in the TPC is equally probable at all positions in (x, y, z), as well as in S2 area, assuming that a dark matter candidate could take any mass. For this reason, signal data used in this model should be evenly weighted in drift time, radius, and S2 area. In order to maintain high statistics in each bin during weighting, it was chosen to use 20, 20, and 10 bins for area, R^2, and drift time respectively.

6.5.3 Tuning a Static Width Cut

A cut using the S2 pulse width of events can improve the sensitivity of S2-Only analysis to signals. In the case of a static width cut, events are selected if they have an AFT90-AFT10 between an upper and lower bound. To select a position for these boundaries which maximises sensitivity, the rate of signal and background events can be calculated from the sum of event weights, and a sensitivity can be calculated for a given upper and lower bound.

Signal and background rate are then calculated from the sum of event weights, from which sensitivity can be quantified and the optimum position of these boundaries selected.

When considering exclusively gate events, which are located at lower S2 widths, the location of an upper bound has negligible impact on background rejection. The distribution of background-subtracted sensitivity in this region, $\frac{s}{\sqrt{b}}$, is shown in Figure 6.25. Maximum signal to background with regards to gate events is achieved with a cut which removes events with a pulse width less than (2300 ± 20) ns.

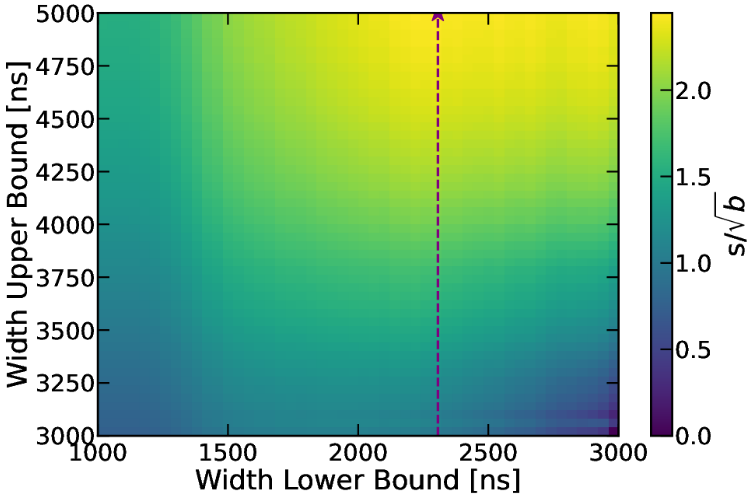


Figure 6.25: Background-subtracted sensitivity for varying maximum and minimum S2 widths, assuming only a gate radiogenic background. Maximum sensitivity is reached with a cut on events of pulse widths below (2300 ± 20) ns.

When considering exclusively cathode events, the lower bound has negligible impact on background rejection. The distribution of background-subtracted sensitivity in this region is shown in Figure 6.26. Maximum sensitivity is achieved with a cut which removes events with a pulse width above (3160 ± 20) ns.

When considering a background-subtracted analysis, the presence of both gate and cathode backgrounds results in the upper and lower bound of a cut on pulse

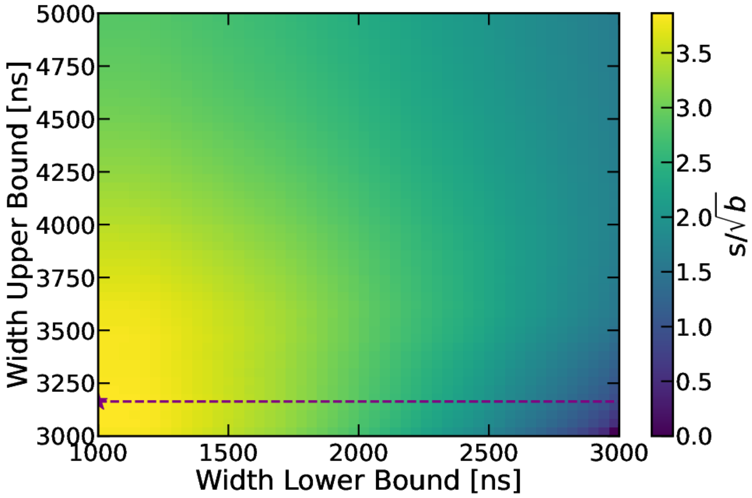


Figure 6.26: Background-subtracted sensitivity for varying maximum and minimum S2 widths, assuming only a cathode radiogenic background. Maximum sensitivity is reached with a cut on events of pulse widths above (3160 ± 20) ns.

width shifting. As the background rate relative to signal rate increases, the upper and lower bound of a cut on pulse width with maximum sensitivity broaden, to increase signal acceptance. As the rate of cathode events is lower than of gate events, the upper bound increases more than the lower bound, resulting in a width range of (1900 ± 20) ns to (4960 ± 20) ns maximising sensitivity. The distribution of sensitivities can be seen in Figure 6.27.

Overall, this cut rejects 67% of gate events and 22% of cathode events, preserving 78% of signal data. The distribution of data relative to the cut can be seen in Figure 6.28.

In a Poisson analysis, the boundaries of a cut become much more narrow, in order to maximise rejection of background events at the expense of signal. A maximum sensitivity is achieved when rejecting pulse widths below (2920 ± 20) ns and above (3450 ± 20) ns. The distribution of Poisson sensitivities can be seen in Figure 6.29.

This narrower width cut has a signal acceptance of just 15%. However, 90% of cathode and 95% of gate events are rejected. The distribution of data relative to

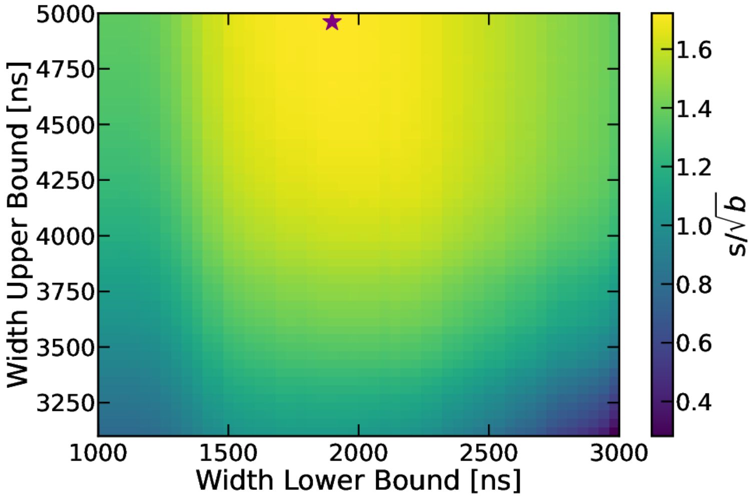


Figure 6.27: Background-subtracted sensitivity for varying maximum and minimum S2 widths. Maximum sensitivity is reached at a width range of (1900 ± 20) ns to (4960 ± 20) ns, annotated with a purple star.

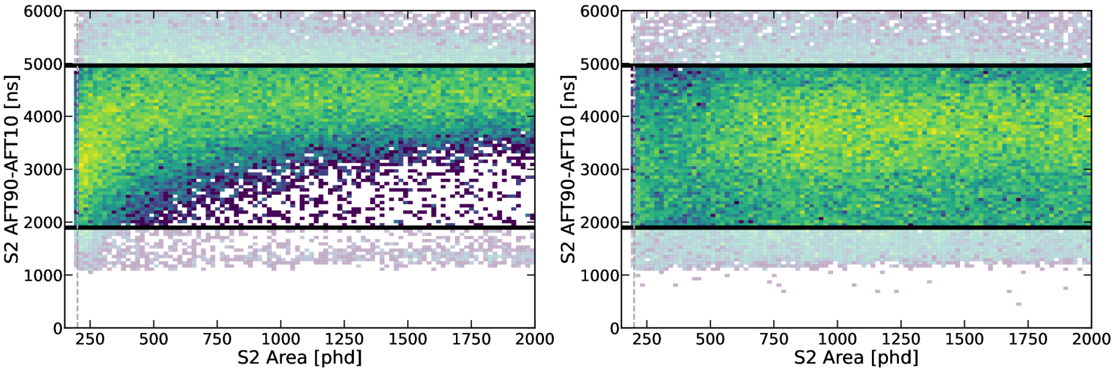


Figure 6.28: S2 pulse area against width of events from the TPC grids (left), and in signal data (right). The upper and lower bounds of a background-subtracted static cut in S2 width are annotated in green and red, respectively. 67% of gate radiogenic events are rejected, improving sensitivity considerably. Only 22% of cathode events can be rejected, in order to maximise acceptance of 78% of signal data.

the cut can be seen in Figure 6.30.

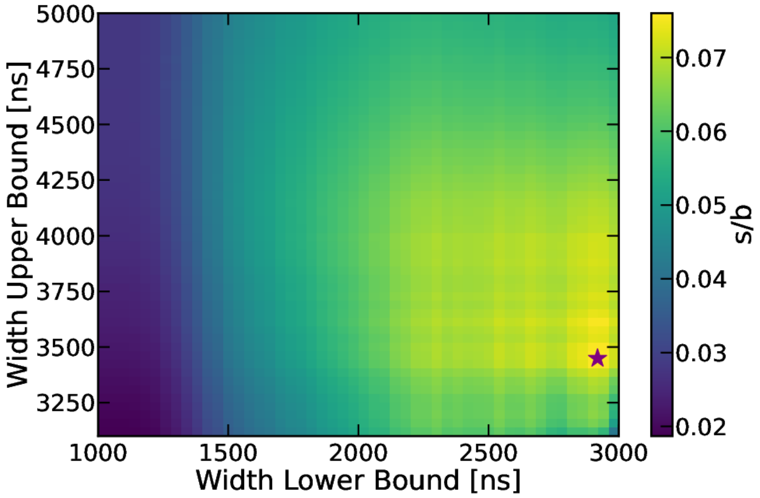


Figure 6.29: Poisson sensitivity for varying maximum and minimum S2 widths. A maximum is reached at a width range of 2918 ns to 3448 ns, annotated with a purple star.

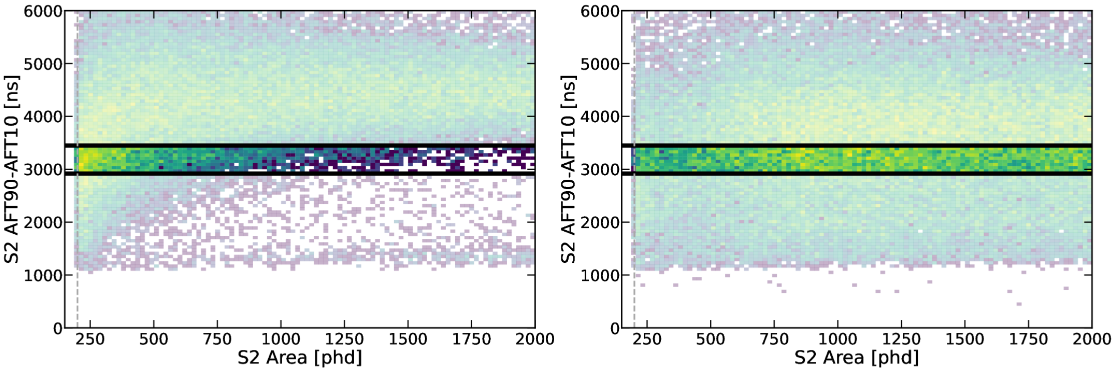


Figure 6.30: S2 pulse area against width of events from the TPC grids (left), and in signal data (right). The upper and lower bounds of a Poisson static cut in S2 width are annotated in green and red, respectively. 95% of gate radiogenic events and 90% of cathode radiogenic events are rejected, improving sensitivity considerably. Only 15% of signal events are accepted, in order to maximise rejection power.

Dynamic Cut in Widths

A static cut in S2 width ensures consistent signal acceptance across all energies. However, this results in decreasing background rejection at lower energies. In contrast, a cut which varies in width with pulse area allows for constant background rejection, at the expense of signal at low energies. This may be desirable for a low-mass WIMP search for masses nearer to 8 GeV/c^2, where recoil energies are greater and signal events therefore extend to higher S2 pulse areas.

Such a cut can be designed using the boundaries shown in Figure 6.24. The same tuning method as with a static width cut is repeated, with varying standard deviation σ from the mean width expected of gate and cathode events.

In a background-subtracted analysis, the optimum boundaries for a dynamic cut in pulse width can be seen in Figure 6.31. 23% of cathode events and 66% of gate events are rejected, with this fraction remaining consistent across S2 area. 80% of signal events are preserved: however, this number decreases towards zero at the analysis threshold of 200 phd.

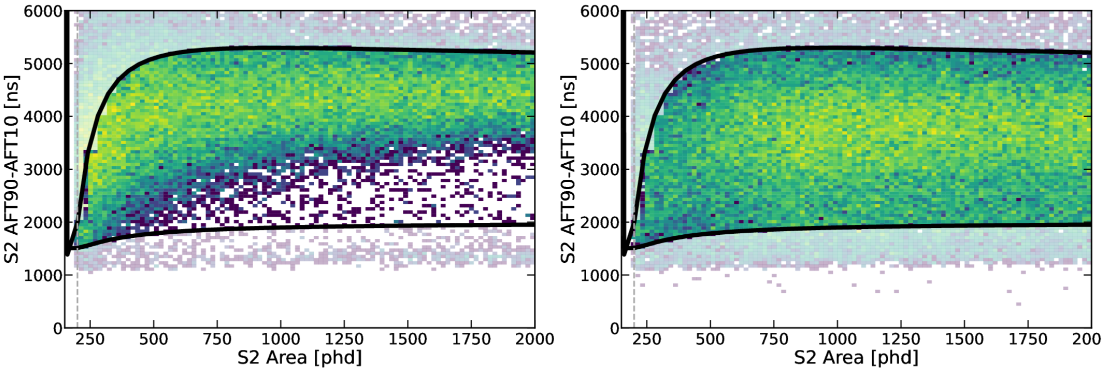


Figure 6.31: S2 pulse area against width of events from the TPC grids (left), and in signal data (right). The upper and lower bounds of a background-subtracted dynamic cut in S2 width are annotated in green and red, respectively. 66% of gate radiogenic events and 23% of cathode radiogenic events are rejected, improving sensitivity. 80% of signal events are preserved, with this fraction decreasing at lower S2 areas.

In a Poisson analysis, the optimum boundaries for a dynamic cut can be seen in Figure 6.32. 99% of cathode and 94% of gate events are rejected from analysis, at the expense of 85% of signal events. Signal acceptance is biased towards higher S2 pulse areas, with acceptance reaching 0 above the analysis threshold.

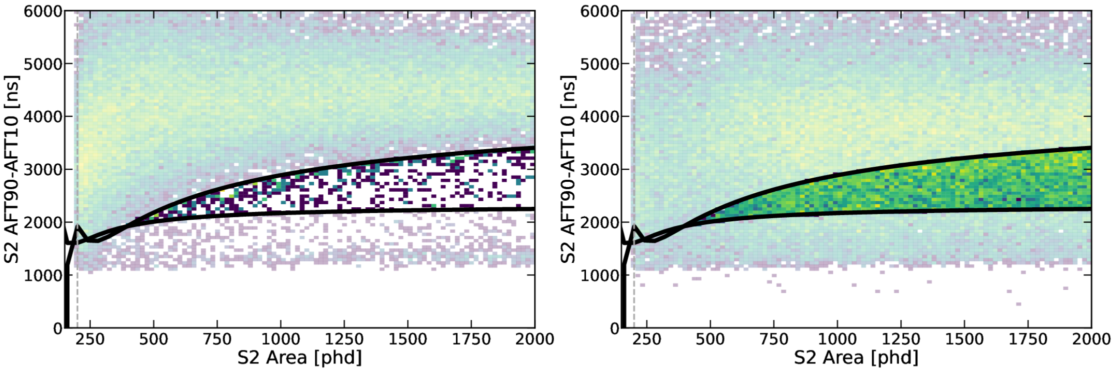


Figure 6.32: S2 pulse area against width of events from the TPC grids (left), and in signal data (right). The upper and lower bounds of a background-subtracted dynamic cut in S2 width are annotated in green and red, respectively. 66% of gate radiogenic events and 23% of cathode radiogenic events are rejected, improving sensitivity. 80% of signal events are preserved, with this fraction decreasing at lower S2 areas.

Figure 6.33 shows the acceptance of the static and dynamic width cut, tuned for a background-subtracted analysis. The rejection of gate events is constant for both a static and dynamic width cut. The rejection of cathode events is also constant in a static width cut. However, below 800 phd, cathode leakage decreases when using a dynamic width cut, down to just 5% at threshold. Signal acceptance is constant when using a static width cut, but decreases below 800 phd using a dynamic cut, reaching just 10% at threshold.

The overall fraction of signal acceptance and background rejection are constant when considering all S2 pulse areas. However, acceptance and leakage are consistent when using a static width cut, whereas they are biased towards higher S2 pulse areas when using a dynamic width cut. The former is therefore best for WIMP searches of mass 2 GeV/c^2, where signal events are predominantly found at threshold, while the latter is best for masses around 8 GeV/c^2, where signal extends to higher S2 pulse areas.

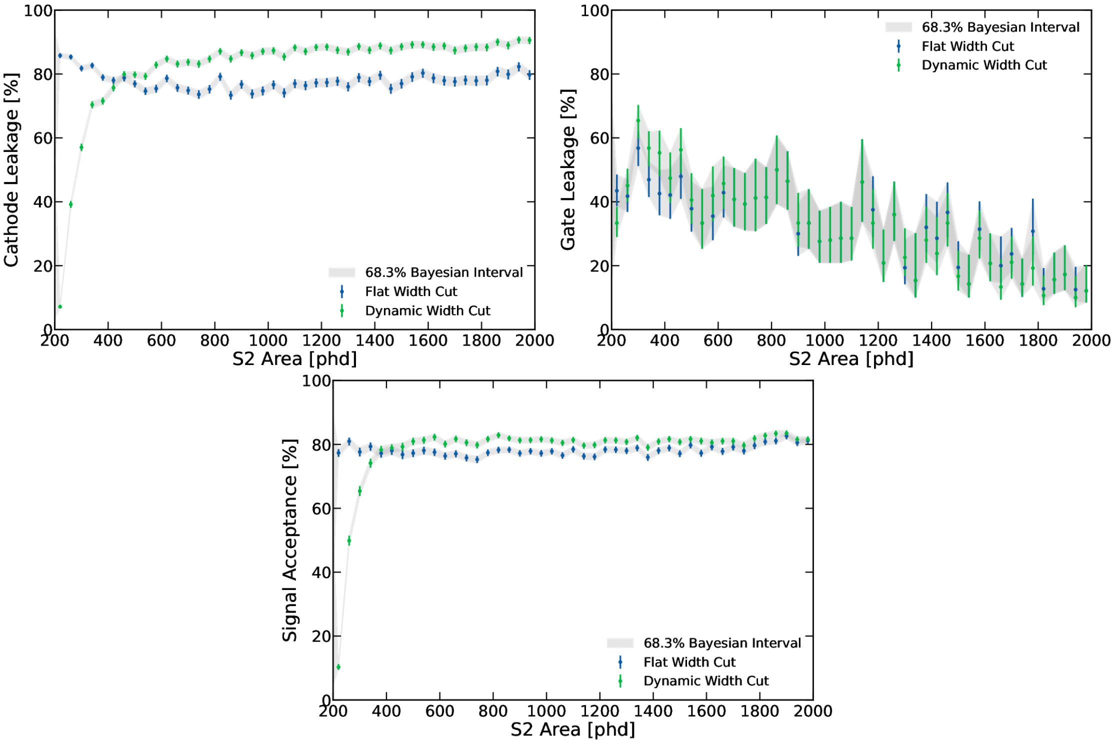


Figure 6.33: The percentage of background leakage from cathode (left) and gate (centre) events, as well as the signal acceptance (right) for a static and dynamic width cut, assuming a Poisson analysis. Signal acceptance decreases for lower S2 areas when using a dynamic width cut, whereas backgrounds increase when using a static width cut. In both scenarios, sensitivity to low-energy signal is impacted.

Versions of these cuts have been produced for analysis of data from the first science run of LZ. A cut can be chosen on the basis of the range of WIMP masses of interest, as well as the choice of either a cut-and-count approach or a profile likelihood ratio approach to measuring sensitivity.

The number of gate and cathode events in the first science run are limited, preventing a more elaborate cut from being tuned. However, the use of machine learning techniques can considerably improve the sensitivity of an S2-Only analysis, particularly at threshold.

6.6 Rejection of Grid Backgrounds with a Boosted Decision Tree

6.6.1 Introduction

A cut on S2 pulse widths is capable of improving sensitivity to low-mass dark matter in an S2-Only analysis. However, such a cut leverages only the difference in diffusion of events from the gate and cathode compared to the TPC bulk. However, the field fringing around gate and cathode wires also results in pulses taking on an asymmetric shape. This shape can be effectively probed using multiple AFT metrics in combination to increase sensitivity beyond that of a simple width cut.

An illustration of this concept can be seen in Figure 6.34. Waveforms of gate, cathode, and signal events are shown: the AFTs of these events, annotated in red, provide information on the overall pulse shape of the S2, allowing grid radiogenics impacted by field fringing to be better distinguished from bulk events.

Several methods of probing the width of an S2 pulse also exist, including multiple AFT-based definitions, the Full Width at Half Maximum (FWHM), and Root Mean Square (RMS) width. These metrics in combination are capable of measuring the Gaussian shape of an S2 with greater confidence, and can therefore be used to better discriminate between cathode events and signal.

Each of these parameters is dependent on the size and position of an S2 pulse: as a result, by including these as input parameters, the boundaries of a cut can be made narrower while preserving signal. By weighting events by position and pulse size, this can be achieved without biasing a cut towards particular energies or TPC positions, as WIMP signals are expected to be homogeneous across the detector.

Several other metrics can also be used to reject grid events, such as the precision of event reconstruction (XY χ^2), the S2 TBA, and the nature of the surrounding event.

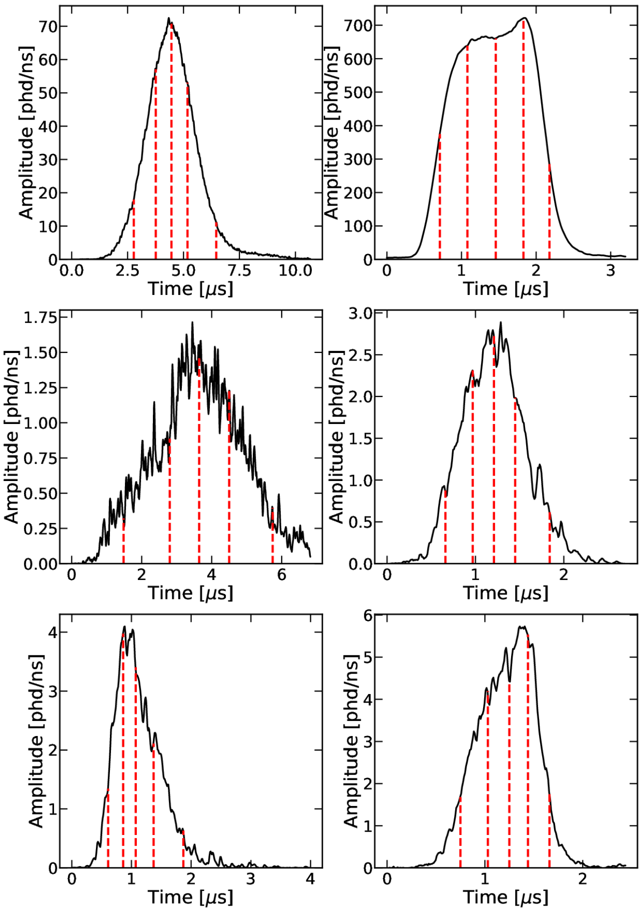


Figure 6.34: Example waveforms, plotted alongside their AFT 10, 25, 50, 75, and 90 (red). Top row: square pulses from the gate (right) show a different distribution of AFTs than more Gaussian S2s (left). Centre row: multiple measurements of width using AFTs can increase the rejection power of events on the basis of their pulse width. Bottom row: AFTs can probe the skewness of grid (left) S2s, to distinguish from bulk (right) S2s.

The signal and background datasets obtained contain several relationships in each of these parameters which are difficult to see by eye. Such relationships, with the statistics available in LZ, are best suited to a machine-learning-based cut. As events are divided into distinct populations of signal and background, a cut can be trained through supervised learning, where an algorithm selects the ideal cut for a selection of events of known origin.

Several methods exist of rejecting events using multiple input parameters through supervised learning. However, the decision has been made to use a Boosted Decision Tree (BDT) for this analysis, due to the ability to easily quantify the importance of input parameters post-training, and use these to draw conclusions about data.

6.6.2 Boosted Decision Trees

A decision tree is a set of individual checks which take a range of input data and divide it recursively based on the features of this data. However, a single decision tree would produce a very simplistic binary division between input data. Instead, an ensemble of decision trees $h(x)$ are created, and an output is given in the form

$$y(x) = \Sigma_i w_i h_i(x). \tag{6.1}$$

Such a form allows for a more continuous distribution of output, allowing not only for many metrics to be used in determining the nature of an event but also for the confidence of such a classification to be quantified.

After each iteration of training, data is given a weight w based on its accuracy: the more often an event is misclassified, the more it is weighted in future iterations. This allows for simple relationships to remain consistent throughout training, and for a BDT to eventually train itself to reject specific clusters of data.

In order to evaluate the success of a BDT, it is compared to a function taking the form

$$O(x) = \Sigma_i l(Y, y) + \Sigma_t \Omega(f) \tag{6.2}$$

The first component here is the loss function - the distance between the truth value in data and the prediction of samples. This is lowest when the prediction of the BDT most closely represents the data. The second component here is a regularisation function - a factor specifically designed to punish training on more complex trees, as a bottleneck [170].

Initially, few decision trees are present in $y(x)$. As training continues, more trees are introduced. In doing so, each new layer is tuned such that it reduces the residual errors of the previous layers, a process referred to as boosting. This is repeated using a gradient descent approach, where the weight of successive layers is selected depending on the variation of the loss function.

6.6.3 BDT Structure

A BDT has several input parameters, which can be varied to ensure optimum performance for a given input dataset. The maximum number of nodes in an individual decision tree is referred to as ’depth’. A greater depth allows for more complex relationships to be modelled in data. However, a large number of nodes can lead to overfitting, where a BDT trains to reject specifically training data, and will not generalise to other datasets. For this work, a depth of 5 nodes was found to maximise performance while preventing overfitting of data. The maximum number of decision trees was set at 50, with the same motivation.

The choice of loss function can also affect training, altering the final BDT as well as the rate at which it is reached during training. The BDT used in this analysis was developed using the popular python module scikit-learn [171]. Scikit-learn’s SAMME algorithm [172] is shown to achieve a low test error with fewer iterations, for a classification BDT, and has been chosen for this work.

The minimum sample split refers to the minimum number of events which reach a given node of a decision tree before it produces another layer. This prevents overfitting of data, and allows for the complexity of a BDT to be constrained. As the number of layers is already minimal, the minimum sample split has been set at just five events.

Finally, events are divided into training and testing data, in order to measure the extent of overfitting. For this work, 75% of events are used for training, with the remainder being used for validation.

6.6.4 Input Structure

The data and weighting used when tuning a static and dynamic cut in pulse widths can also be used in training a BDT. In order to ensure the final trained BDT has not overfitted to exclusively AmLi signal, input data extends to 4000 phd, such that the output of CH_3T calibration data can be compared to training data.

The input RQs into the BDT are:

AFTs: Allows for the measurement of widths as well as the overall pulse shape

FWHM: Helps to probe the shape of the S2 peak

RMS Width: Allows for a more general width metric, to compare to AFTs

Amplitude: Gate events can be tagged by their high amplitude

Peak Time: The peak of an S2 waveform relative to the pulse start time is generally earlier for gate events and later for cathode events

TBA: The TBA of gate events will vary slightly from signal data for events very near to the liquid surface

XY χ^2: Distorted gate events are more likely to display poor XY reconstruction

Area and Radius: While not exploited, due to the weighting of the events, relationships can be found between these and other parameters

The pulse area in an event which is not produced by the main S1 or S2 can also be used to reject grid radiogenic events, due to the high field strength around grid wires also elevating the rate of scintillation in an interaction. However, the data quality in AmLi calibrations is somewhat lower than that in WIMP search data. As such, a BDT with event pulse area as input is capable of identifying signal events in a manner which does not generalise to WIMP search data. Photoionisation of the gate and glue ring is more likely to be classified as a separate S2 pulse in events of low drift time, as discussed in Chapter 5, but the shape and size of S2s following the main S2 pulse cannot be used as BDT input for the same reason.

Each of the chosen input parameters will show a unique correlation for cathode and gate data. However, for small S2s, these trends will overlap significantly with signal. The use of trends in multiple parameter spaces allows background events to be rejected with the preservation of a greater fraction of signal events. This concept of using multiple parameter spaces is illustrated in Figure 6.35: background and signal data overlap in each of these metrics, but are distinct when considered in multiple parameters simultaneously.

The impact of each individual RQ can be quantified using the Kendall rank correlation coefficient, a test statistic used to establish whether two variables are statistically dependent [173]. A coefficient of 1 or -1 indicates that two variables are perfectly correlated: as such, one of these inputs can be dropped from the BDT with no impact on performance, to improve training time. A coefficient of 0 indicates that two variables are not correlated, and will therefore be used in a BDT independently from one another. A coefficient between these two values indicates a physical relationship where two parameters are somewhat correlated, which can be exploited in an individual decision tree to reject background events.

The Kendall coefficient of input variables relative to one another for signal data can be seen in Figure 6.36. This can be compared to the same distribution for Cathode and Gate events, in Figures 6.37 and 6.38. AFTs in close proximity (for example, AFT5 and AFT10) are closely correlated in signal, with this correlation decreasing for more distant AFTs, as a result of AFTs naturally following one another in a pulse. However, AFTs are less correlated in gate radiogenic events,

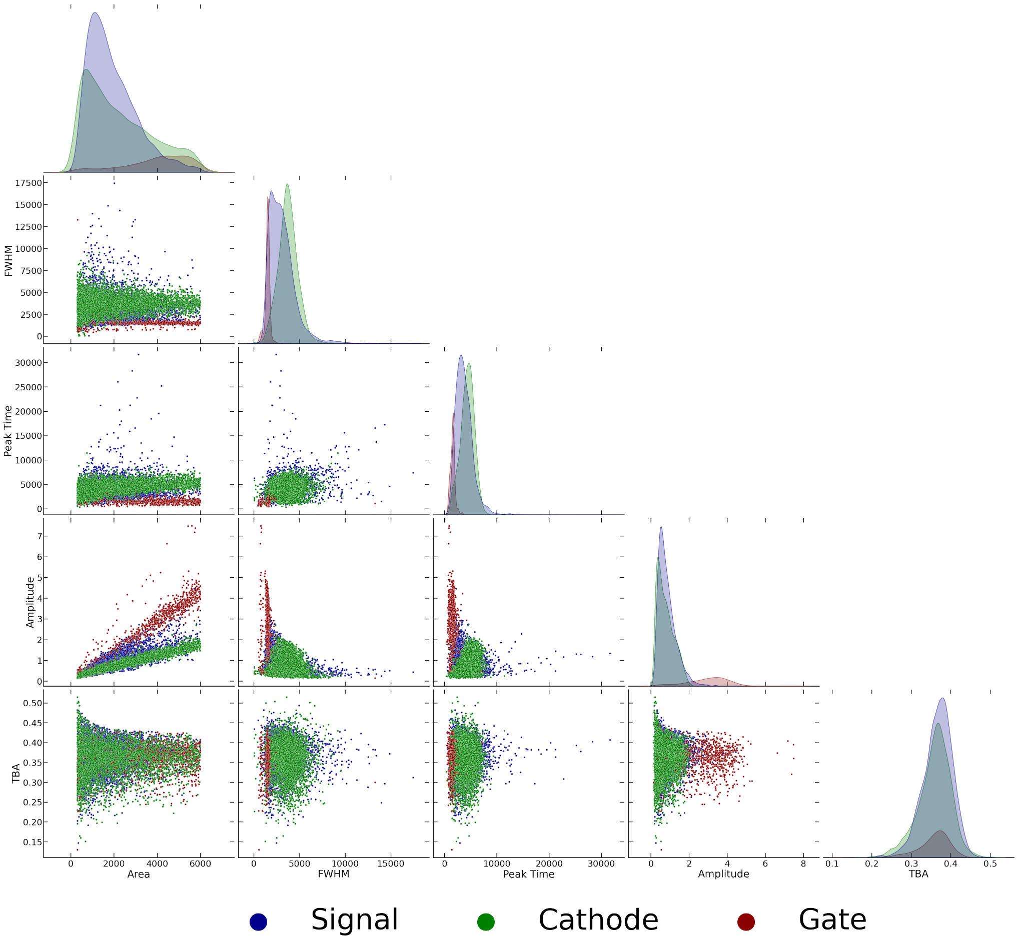


Figure 6.35: A pair plot of signal (blue), cathode (green), and gate (red) events, plotted in area, FWHM, peak time, amplitude, and S2 TBA. Background and signal data is distinct at high S2 areas. At lower S2 areas, data is not distinct in individual parameters, but can be discriminated through the use of multiple parameters.

due to field fringing distorting pulse shapes. Similarly, FWHM and RMS width are found to correlate in signal data, but do not correlate as strongly in cathode events, due to the broad shape of S2s from this position. Similar differences can be found in peak amplitude and peak time, due to the increase of rise time with increasing drift time. S2 TBA and χ^2 are found to not correlate with any other parameters, and can be used in isolation to reject events from the gate array which are near to the liquid surface.

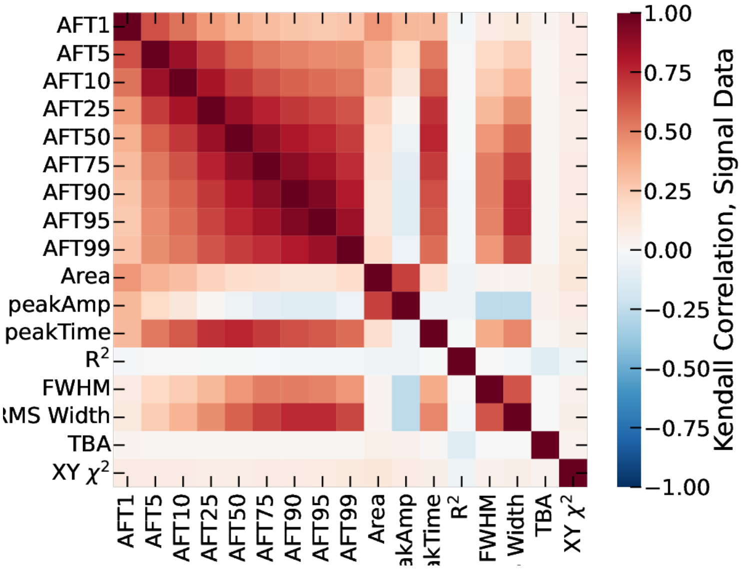


Figure 6.36: Kendall correlation matrix for signal events, for all relevant RQs. Close correlation is found between AFTs, RMS width, and FWHM, with limited correlation in S2 TBA and χ^2.

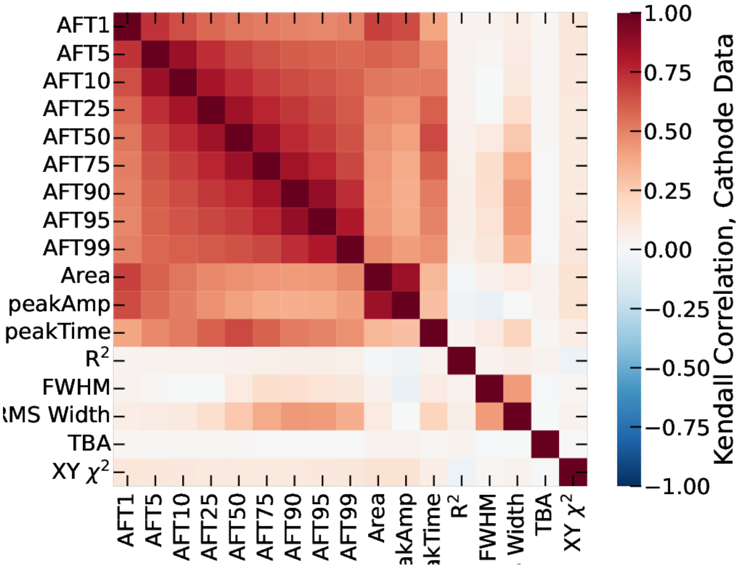


Figure 6.37: Kendall correlation matrix for cathode radiogenic events, for all relevant RQs. Correlations between width metrics are lower than in signal data, allowing for their use in rejecting cathode events.

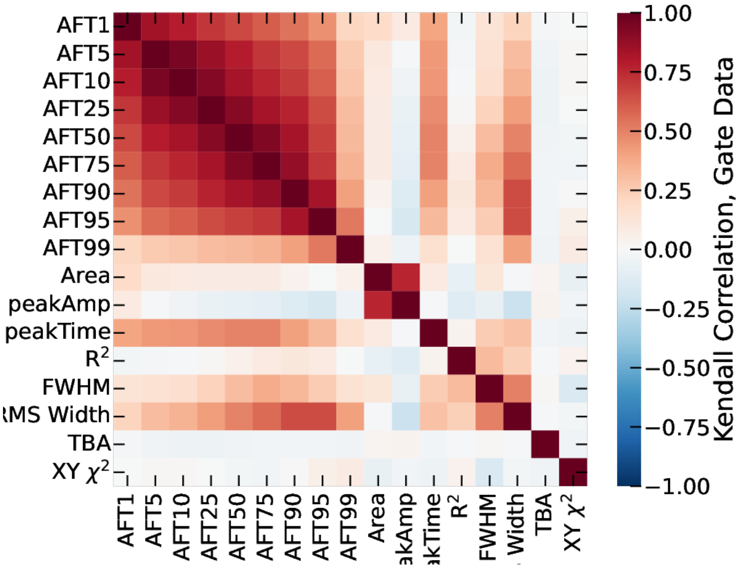


Figure 6.38: Kendall correlation matrix for gate radiogenic events, for all relevant RQs. While width metrics are more closely correlated than in signal data, less correlation is seen in amplitude, peak time, S2 TBA and χ^2, allowing for these inputs to be effective in rejecting gate events.

6.6.5 Training of the BDT

An S2-Only BDT is developed using an Adaboost (Adaptive Boosting) Classifier [172] in the python module scikit-learn [171], using the properties described in Section 6.6.3. The BDT was trained for 50 iterations, at which point overfitting was observed.

Post-training, the BDT shows distinct distributions for signal and background data. The training and testing data for both signal and background are in agreement with one another, indicating that the BDT has not overfitted on training data. The distribution of these results for training and testing data can be seen in Figure 6.39.

To ensure the BDT is not training on features exclusive to AmLi calibration data, this output is compared to that of CH_3T calibration data in Figure 6.40. The distribution of background events in CH_3T is identical to background events

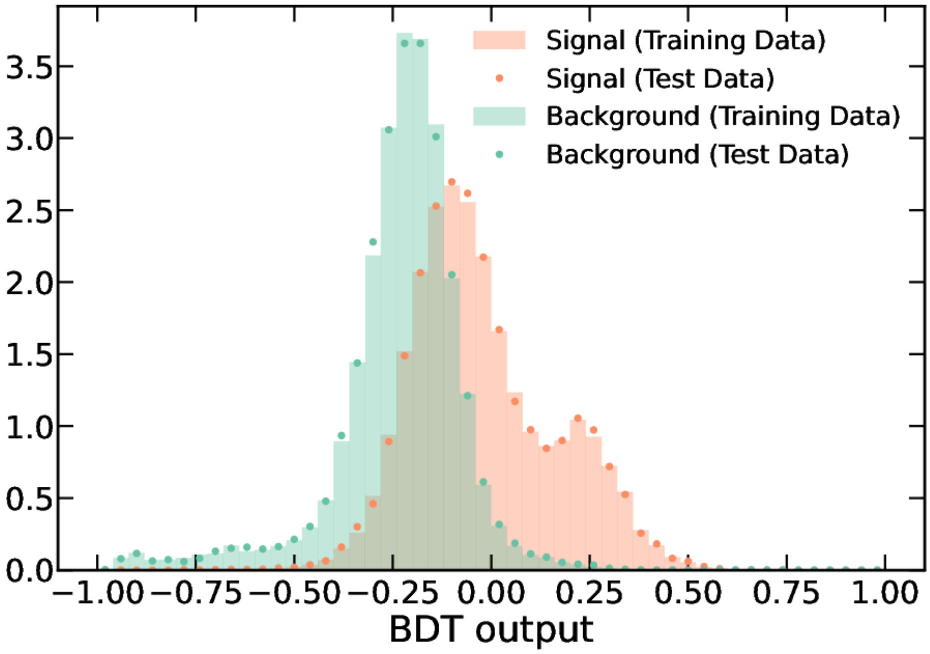


Figure 6.39: BDT Output for signal and background events. Training and testing data are represented by bars and points, respectively. Signal and background take on distinct distributions. The distribution of training data (bars) is similar to that of testing data (points), indicating that the BDT is not overfitting.

selected from WIMP search data. The distribution of signal events differs: however, data takes the same general shape of two Gaussian-like peaks, with these peaks being more symmetric for CH_3T. Events of a higher BDT score are found to be located towards the centre of the TPC in z at higher S2 pulse areas. Events in this position of this energy are more common in CH_3T than in AmLi, explaining this discrepancy. The trained BDT can therefore be generalised to S2-Only data.

A key method of evaluating BDT performance is a Receiver Operator Characteristic (ROC) curve [174], where the amount of background leakage and signal efficiency are plotted against one another for varying cuts in BDT output. The area underneath a given position on the curve relative to 100% leakage and 0% efficiency can then be used to determine the optimum position for a cut.

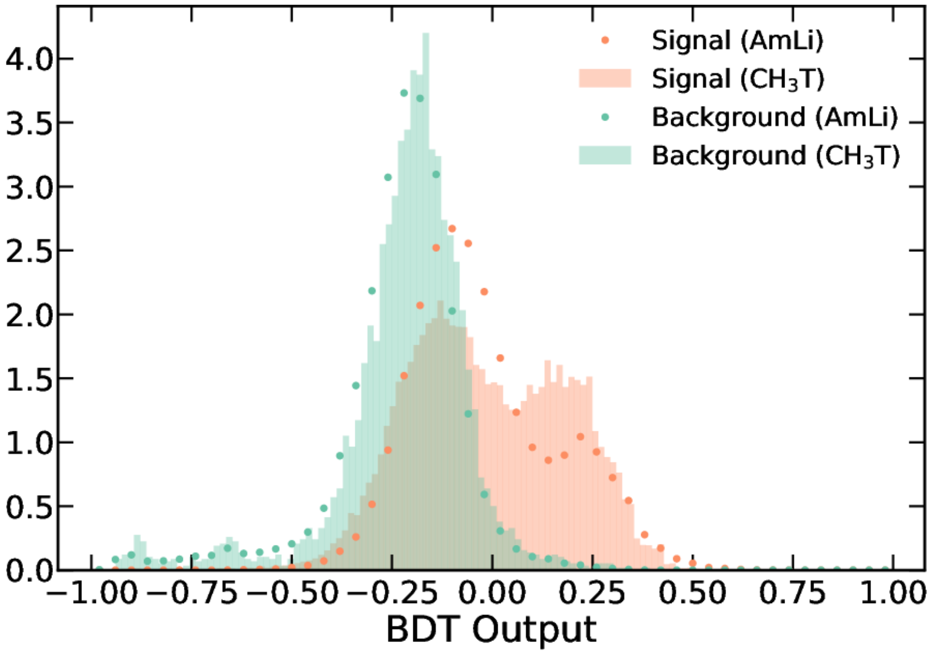


Figure 6.40: BDT Output for signal and background events. CH_3T and AmLi signal are represented by bars and points, respectively. The distribution of background data is the same for both datasets. The distribution of signal data takes the same general form, but tends towards higher output values in CH_3T due to its generally higher S2 pulse areas.

The average ROC curve for the BDT trained in this chapter can be seen in Figure 6.41, in comparison to the results of LZ width cuts and a similar study conducted in LUX [164]. Grey lines indicate the position of constant $\frac{s}{\sqrt{b}}$ and constant $\frac{s}{b}$, to gauge the improvement in sensitivity in a background-subtracted and Poisson analysis. When accounting for the rate of gate and cathode events, the static and dynamic width cuts both had an approximate signal acceptance of around 78% and 15% for background-subtracted and Poisson analysis, and a background leakage of 44% and 5%, which are also annotated. For a given signal acceptance, the ROC curve for the trained BDT reaches a lower level of background leakage compared to both width cuts in LZ and the use of a BDT in LUX. This demonstrates the utility of a BDT in S2-Only analysis, and the improved conditions in LZ for S2-Only sensitivity compared to LUX.

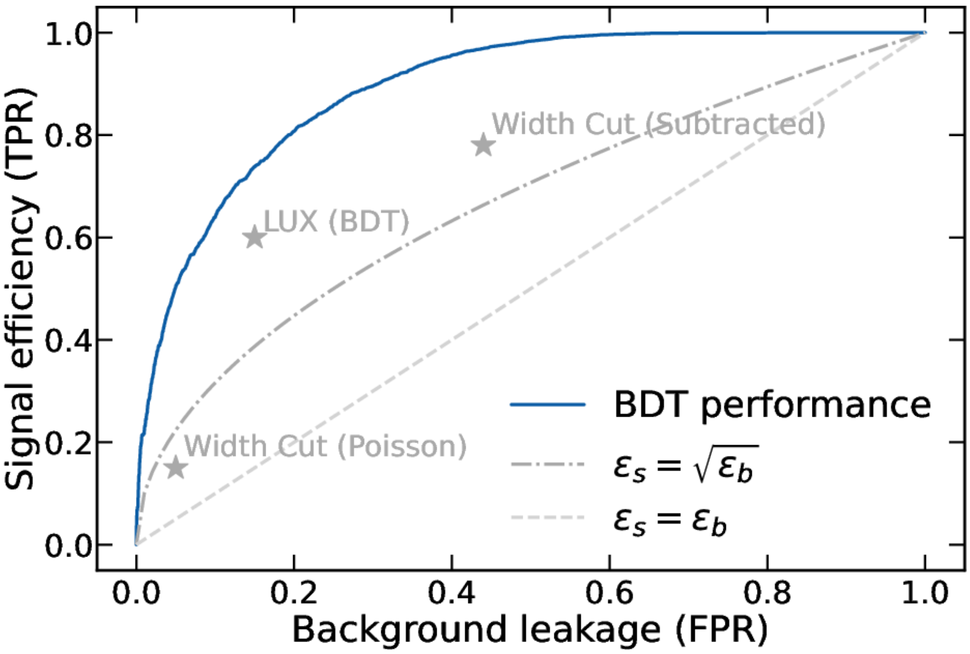


Figure 6.41: Average Receiver Operating Characteristic (ROC) curve across all training data. Points are annotated for the average position of a width cut for background-subtracted analysis, for a Poisson analysis, and for similar analysis from LUX [164].

BDT output can also be shown for slices of S2 area. The results of this can be seen in Figure 6.42. For each slice, a star indicates the position of best signal-to-background, calculated as the product of signal efficiency and background rejection. These points are consistently in the region of 80 to 85% signal acceptance, considerably higher than that achieved with a simple width cut and a marked improvement on the 65% acceptance observed in LUX [164]. Background leakage is found to decrease in analysis of larger S2 pulses, due to the increasing precision of input parameters with which to identify relationships in data.

The weight of each parameter in a BDT can be used to measure the ’importance’ of each parameter individually. This provides a second application for a BDT, in providing an improved understanding of the physics influencing input data which can then be applied to non-ML analysis cuts.

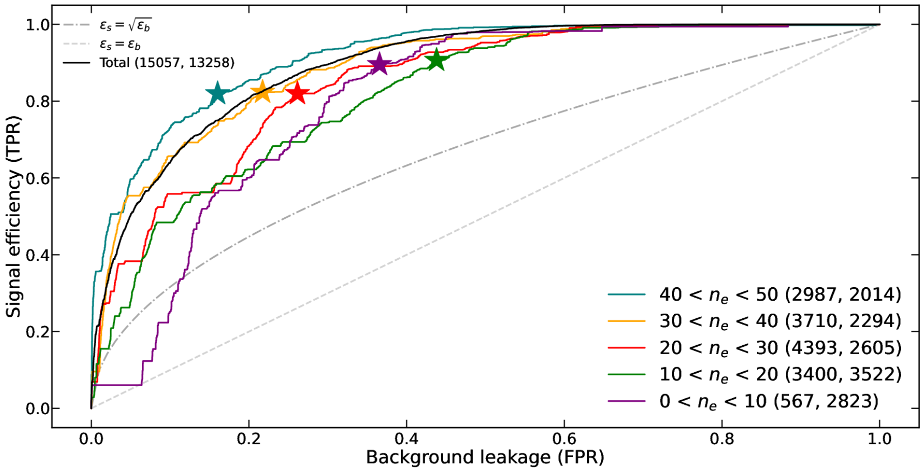


Figure 6.42: Average Receiver Operating Characteristic (ROC) curve for varying slices of S2 area in training data. Points annotate the maximum ROC area for a given curve, indicating an optimum cut position.

The importance of RQs in the BDT trained here can be seen in Figure 6.43. The BDT places high importance on the area and radius of an event: while events are weighted such that no physical relationships can be identified in these RQs alone, the relationship between these RQs and other parameters is exploited by the BDT in order to improve the precision of its cuts. The RMS width is also of high priority, likely because of its precise measurements of the width of an S2. Priority is then placed on early AFT metrics (effectively a pulse's rise time), pulse amplitude, and χ^2, all of which are used to reject gate events influenced by field fringing. Metrics related to other AFTs and TBA are of lower priority, but still allow for more precise cuts of cathode events to be made.

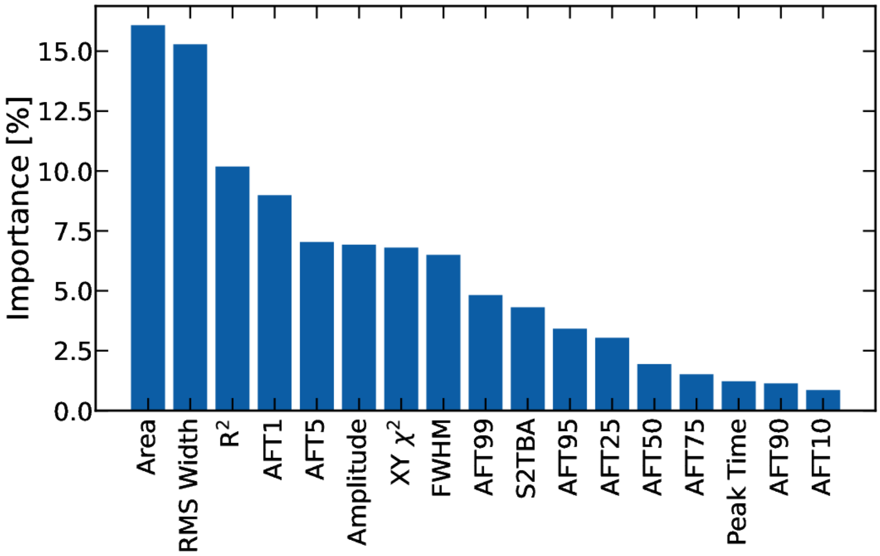


Figure 6.43: The importance of RQs used in the BDT to reject grid radiogenics. RQs with higher importance are generally more effective at rejecting grid radiogenic events.

The points in Figure 6.42 assume the same rate of signal and background events in analysis. However, signal counts are considerably lower than background counts in this analysis. As such, a choice of BDT output threshold must be selected in order to maximise sensitivity.

This concept is demonstrated in Figure 6.44. Background leakage and signal acceptance both drop with increasing BDT output threshold. However, gate events drop off faster than cathode events, which drop off faster than signal data. This leads to a peak in background-subtracted sensitivity at a BDT cut point of -0.18, where sensitivity is maximised.

This optimum cut position varies with changing S2 pulse area. For this reason, the procedure outlined in Figure 6.44 is repeated for different S2 pulse areas, for both a background-subtracted and Poisson analysis. The optimum position of a BDT cut is found to fit well to a linear function in the former scenario, whereas the latter is found to remain consistently at a threshold of 0.1. These thresholds, as well as contours of BDT output for each dataset against S2 area, can be seen in Figure 6.45.

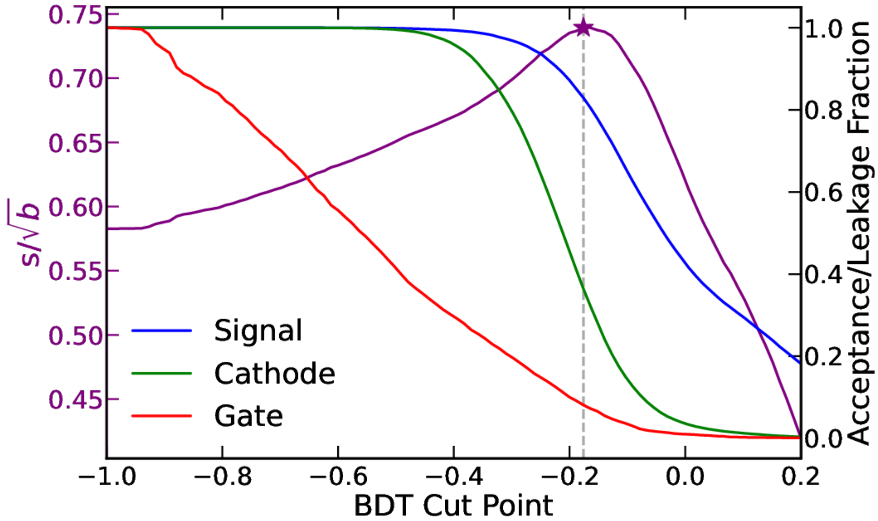


Figure 6.44: Measurement of the ideal BDT cut position using expected gate, cathode, and signal rates. The weighted acceptance of signal (blue), cathode radiogenics (green) and gate radiogenics (red) are plotted against background-subtracted sensitivity.

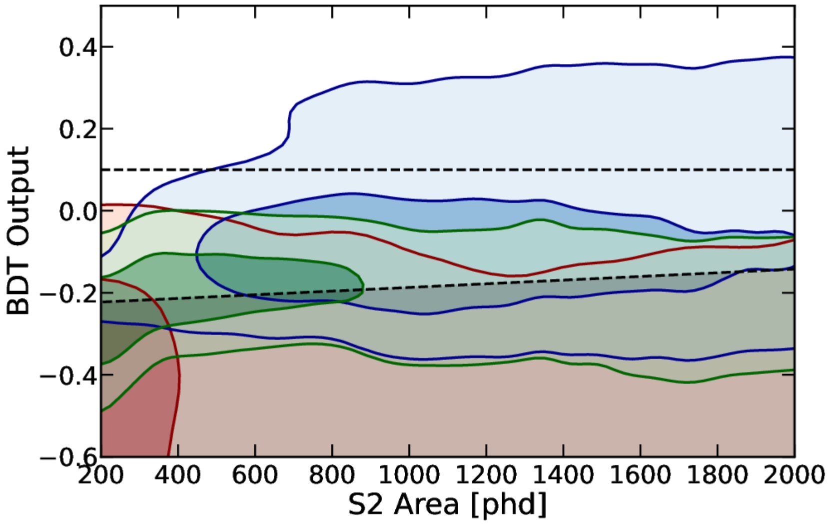


Figure 6.45: Contours of BDT output against S2 area for gate (red), cathode (green) and signal (blue) data. Black dashed lines annotate the chosen thresholds for a background-subtracted analysis at ~ 0.2 and for a Poisson analysis at 0.1, respectively. In the former, background events above threshold increase at low S2 area: in the latter, signal counts above threshold decrease, limiting sensitivity.

Signal acceptance and background leakage can be seen for a background subtracted analysis in Figure 6.46. Cathode leakage is as low as 30% at 2000 phd, increasing to a plateau of 70% at low S2 areas and decreasing to 30% at threshold, where pulse shapes approach those of single electrons. Gate leakage shows a similar relationship, starting at just 5% and increasing to 30%, before decreasing to 15% at threshold. In contrast, signal acceptance remains consistently at 80%, only decreasing to 60% at threshold.

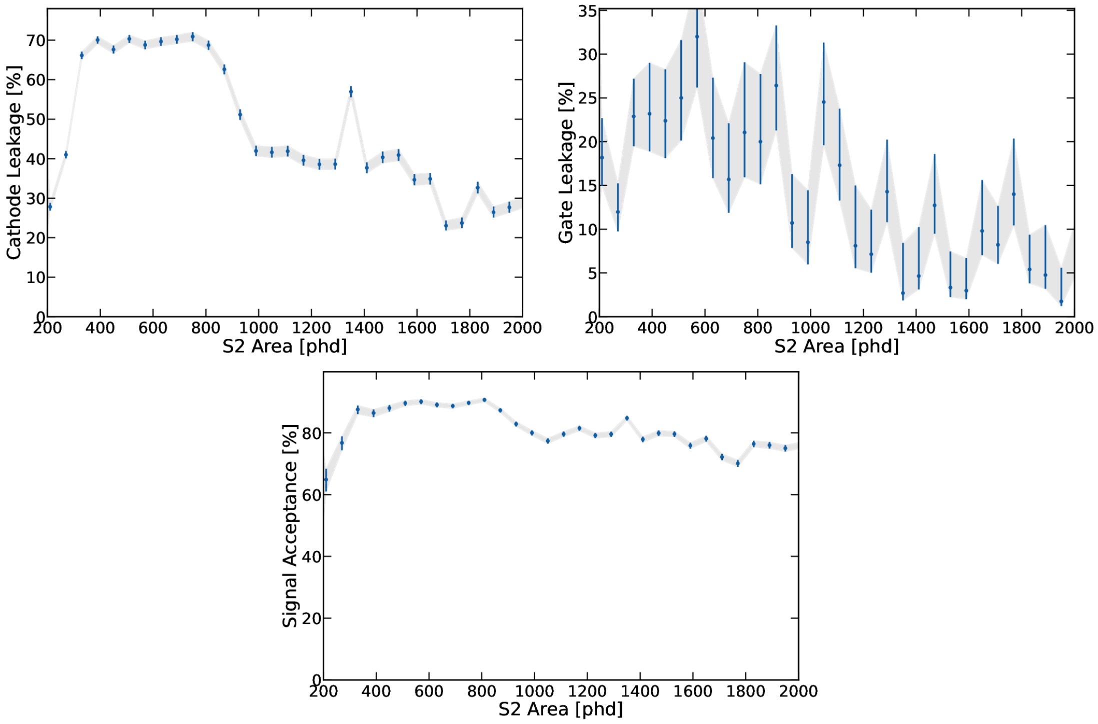


Figure 6.46: The fraction of events rejected by an S2-Only BDT in a background-subtracted analysis, for background (left) and signal (right) populations.

The BDT signal acceptance and background leakage can be seen for a Poisson analysis in Figure 6.47. In contrast to the background-subtracted analysis, cathode leakage remains consistently below 2% and gate leakage below 5%, with this decreasing at threshold. In order to achieve this rejection, signal acceptance is 30% at 2000 phd, decreasing to just 3% at threshold.

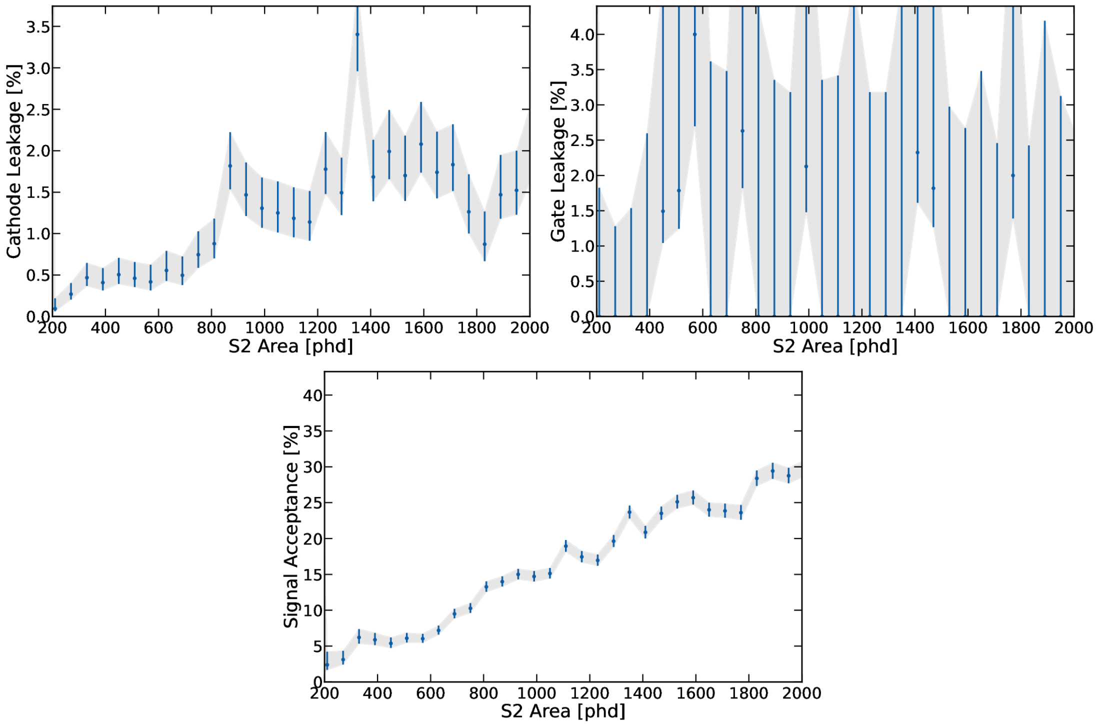
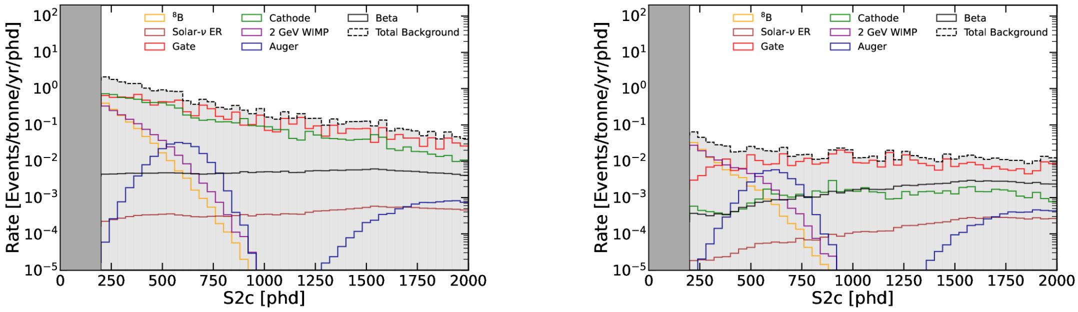


Figure 6.47: The fraction of events rejected by an S2-Only BDT in a Poisson analysis, for background (left) and signal (right) populations.

The improvement in sensitivity achieved through the application of a BDT can be seen in Figures 6.48(a) and 6.48(b). Figure 6.48(a) shows the distribution of signal and background in Figure 6.7 scaled by the signal acceptance and background leakage of the BDT in a background-subtracted analysis. At threshold, the ratio of gate and cathode events to signal is considerably lower, increasing sensitivity. Figure 6.48(b) shows this same distribution assuming a Poisson analysis. While event rates are considerably lower, the rate of cathode and gate events is subdominant to that from signal at threshold.



(a) Background-subtracted **(b)** Poisson

Figure 6.48: Distribution of background events in an S2-Only analysis after the application of a BDT, tuned to a background-subtracted and Poisson analysis. The expected distribution of a 2 GeV/c^2 WIMP of cross section 1×10^{-46} cm^2 is also included, modelled using NEST [121] with rates measured using WIMPrates [84]. In the background-subtracted analysis, the ratio of background to signal is improved at threshold. In the Poisson analysis, almost all background events are rejected at threshold, at the expense of signal.

Similar projections can be made for the static and dynamic width cuts outlined earlier in this chapter. The projected rate of backgrounds and a 2 GeV/c^2 WIMP for all analysis cuts for the rejection of grid events is given in Table 6.1.

Table 6.1: Projected rates of grid radiogenic events and a 2 GeV/c^2 WIMP of cross section 1×10^{-46} cm^2 in the range 200 phd to 2000 phd in an S2-Only analysis. Rates are given for all analysis cuts outlined in this chapter, tuned for both a background-subtracted and Poisson analysis.

		Rate (Events/tonne/year)		
Analysis	Cut	Gate	Cathode	2 GeV/c^2 WIMP
N/A	None	1906	603	76
Background Subtracted	Static Width	328	276	59
	Dynamic Width	328	238	23
	BDT	89	179	68
Poisson	Static Width	5.1	1.6	1.46
	Dynamic Width	4.55	0.04	0.05
	BDT	0.56	0.03	0.38

The choice of BDT threshold will ultimately depend on the choice of analysis pipeline for an S2-Only analysis. However, a simple calculation of sensitivity can be made assuming sensitivity scales as in a background-subtracted analysis, using the rates in Table 6.1. The distribution of $\frac{s}{\sqrt{b}}$ for increasing live time can be seen in

Figure 6.49. Assuming other events are rejected with confidence, LZ is projected to be capable of discovering a WIMP of mass 2 GeV/c^2 and cross section 1×10^{-46} cm^2 to 3σ with 200 live days of data, with a 5σ confirmation in 450 days. With the application of a static width cut, 5σ is reached in 175 live days of data. However, the application of a BDT decreases the required number of live days to just 25 for 3σ and 50 for 5σ.

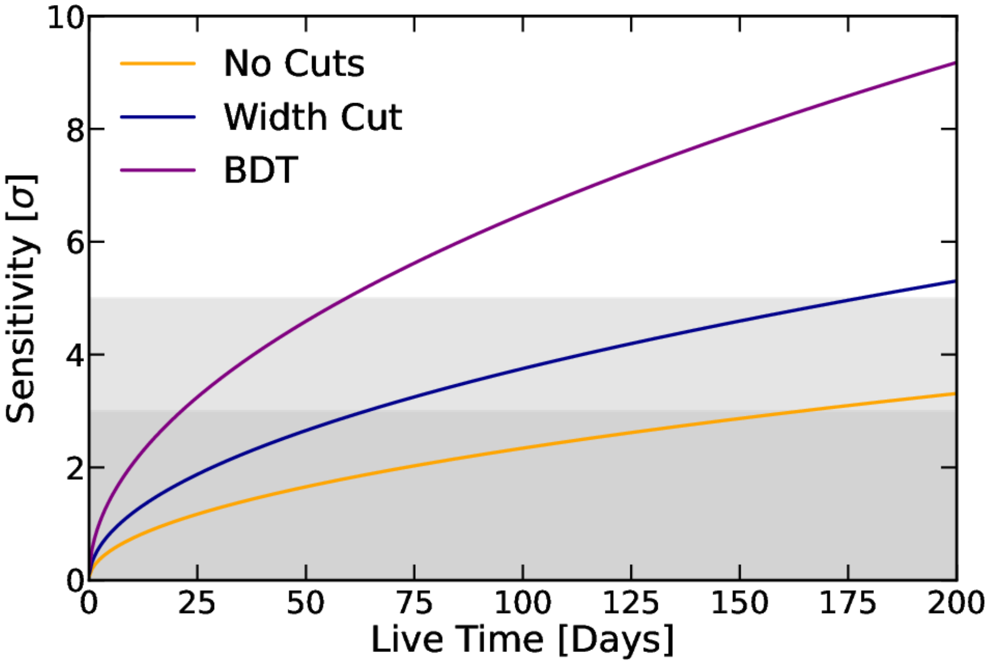


Figure 6.49: Projected discovery limits of an S2-Only analysis in LZ to a 2 GeV/c^2 WIMP of cross section 1×10^{-46} cm^2. Limits are placed using the projected rates of gate and cathode backgrounds and WIMP signal with no cuts (orange), with a static width cut (blue), and with the application of a BDT (purple). Without cuts, LZ data is projected to discover a 2 GeV/c^2 WIMP of this cross section with 3σ confidence using 200 live days of data. With a simple width cut, this decreases to 175 live days. However, with the application of a BDT, this decreases to just 25 days of data, with a 5σ confirmation using 50 live days.

7

Conclusion and Outlook

LZ is designed for the detection of WIMPs of masses between 8 GeV/c^2 and 100 TeV/c^2. However, the technology employed in LZ is versatile enough to be used in a variety of low-energy searches, for the detection of both light dark matter candidates and various other interesting phenomena. The work presented in this thesis maximises the sensitivity of LZ to low-mass WIMPs, through improvements in both rejection of low-energy backgrounds and measurements of acceptance of low-energy signal.

One of the primary backgrounds in LZ, in both primary and alternative S1-S2 analysis, is accidental coincidence. These small pulses, mimicking a low-energy NR in LZ, are irreducible and heavily impact sensitivity to low-mass WIMPs. Chapter 4 outlined analysis cuts designed to reject 95% of accidental events, preserving more than 98% of signal. Through measurements of diffusion, it is projected that LZ's current drift field of 97 V/cm is near to the optimum for mitigating accidental events, which are the dominant background in low-mass WIMP and ^{8}B CEνNS searches.

A limiting factor for sensitivity to low-energy signals in LZ is the S2 detection efficiency. The mis-classification of S2 pulses in LZap is the dominant source of signal detection inefficiency in S1-S2 analysis, and a dominant source of inefficiency and S2 pulse area uncertainty in low-mass WIMP searches. Chapter 5 determines the probability of these effects with precision for the first time in LZ calibration data, then demonstrates that these measurements are in agreement with simulated electron clouds. It is demonstrated that mis-classification of S2s in software does not exceed 3% for any given TPC position or energy, an improvement on the detection efficiency of small signals in LXe TPCs of this scale.

The sensitivity of LZ to low-mass dark matter is maximised in an ionisation-only (S2-Only) analysis, where events without an associated scintillation signal (S1) are considered. Chapter 6 presents a series of data quality cuts to be applied to S2-Only data in order to maximise sensitivity through the rejection of the dominant S2-Only background, radiogenic events from the TPC gate and cathode. Unsupervised learning is used to reject above-anode events from S1-S2 data, improving the quality of signal and background selection for the tuning of these cuts, and these events are demonstrated to have limited impact on S2-Only sensitivity. This clean sample of events is then used to train a series of width-based cuts for analysis of data from the first science run of LZ, as well as a Boosted Decision Tree for later science runs. The BDT is demonstrated to be capable of rejecting upwards of 80% of background events while preserving upwards of 80% of signal, reducing the amount of live days of data required for discovery of a $2\,\mathrm{GeV/c^2}$ WIMP of cross section $1 \times 10^{-46}\,\mathrm{cm}^2$ from 450 days to just 50.

Analysis of S2-Only data in LZ after the application of the cuts presented in this thesis is ongoing. A full PLR analysis of S2-Only data, using S2 pulse area and either pulse width or BDT output, is projected to have unprecedented sensitivity to low-mass WIMPs.

While the analysis methods developed in this thesis will have immediate use in LZ searches for low-mass WIMPs and ^{8}B CEνNS, these topics will become even more important in XLZD. Accidental events are projected to be the dominant background to WIMP sensitivity in a larger TPC. Cuts targeting the correlated drift time of events and their pulse shape will therefore be essential.

The larger TPC size will also result in signal detection efficiency becoming a larger concern, requiring precise measurements of pulse mis-classification. Chapter 5 motivates alternative pulse classification algorithms in XLZD, as well as precise measurements of pulse mis-classification to improve low-mass WIMP sensitivity.

As high-mass WIMP searches approach the neutrino fog in XLZD, low-mass searches become the primary method of achieving sensitivity to dark matter. An S2-Only analysis in XLZD will build on the cuts developed in this thesis, and may achieve sensitivity to ^{7}Be CEνNS.

Appendices

Appendix A: Gaussian Mixture Models

The GMMs used in this thesis are implemented via scikit-learn [171]. This fits N multi-dimensional Gaussians to data, and data is categorised as part of the cluster it is nearest to in multidimensional space. The covariance matrix of each cluster is fitted such that the covariance of the data within the cluster is minimised. With enough clusters, more complicated structures can be successfully imitated. This is illustrated with a 2D toy model dataset in Figure A.1.

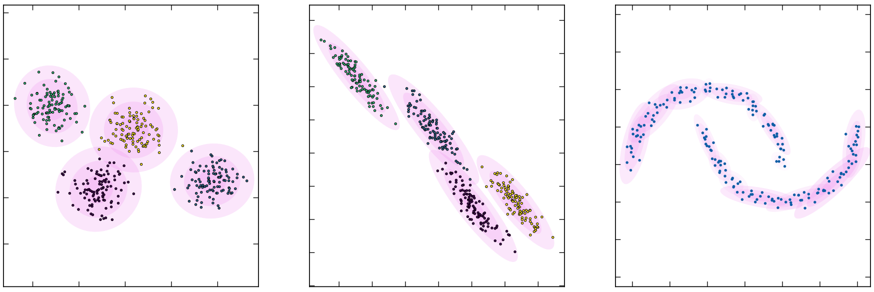


Figure A.1: Illustrations of a Gaussian Mixture Model (GMM) classifying different clusters of 2D toy data. Left: classifying data evenly distributed in space. Centre: classifying data unevenly distributed in space. Right: using multiple clusters to identify more complex distributions of data.

A GMM is useful for identifying overlapping multivariate relationships in data where a simple 2D parameter cut would result in significant signal loss and background leakage. For this reason, a GMM is used in Chapter 4 to identify iS1 and iS2 populations in data, in Chapter 5 to differentiate merged MS S2s from saturated and distorted pulses where possible, to differentiate split S2s from poorly-resolved MS S2s, and in Chapter 6 to identify above-anode events in S1-S2 data.

Appendix B: S2 Toy Modelling

B.1 Waveform Modelling

As discussed in Ref. [141], a single electron can be accurately modelled as a uniform distribution, as the light emitted through electroluminescence is mostly consistent across the extraction region. Electrons arrive at the extraction region following a Gaussian distribution, of standard deviation σ defined by the longitudinal diffusion coefficient, the drift velocity, and the drift time of a given event.

It is therefore possible to model an S2 waveform through the following steps:

- N points are sampled from a normal distribution of mean 0 and standard deviation $\frac{2D_L\sqrt{t}}{v_d}$, as outlined in Section 4.3.2. Each of these N points represents an electron.
- The width of these electrons is sampled from a normal distribution of mean C and standard deviation dC. The number of photons n in each electron is sampled from a similar normal distribution.
- Each electron is a uniform distribution. n photon times are sampled uniformly across the width of each electron pulse, with each photon represented as a narrow Gaussian.
- The photons in each electron pulse are summed, to give the overall S2 pulse.

Examples of waveforms generated using this method can be seen in Figure B.1. This method is used in Chapter 5, to estimate the impact of grid deflection and S2 size on splitting probability.

B.2 Width Modelling

This approach is capable of modelling the waveform of an S2 pulse to some degree of precision. However, it is computationally expensive. The same approximations

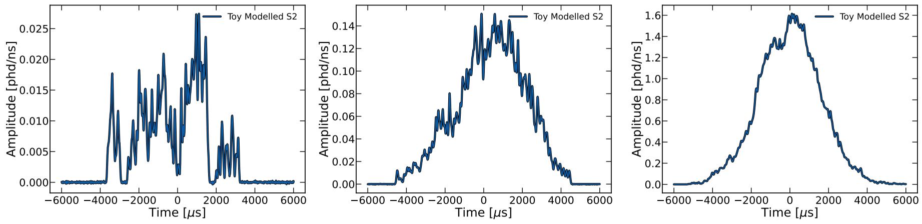


Figure B.1: Examples of toy model S2s generated via the steps listed in Section 5.1.4. Waveforms are generated at a drift time of 1050 µs using 10, 100, and 1000 electrons, from left to right.

can instead be used to calculate only the width of a pulse, without calculating the full waveform.

The integral of a boxcar can be expressed as

$$I(a, \mu, \sigma) = \min\left(1, \max\left(0, \frac{a - \mu + \sigma}{2\sigma}\right)\right), \tag{1}$$

where I is the integral between $-\inf$ and a for a given mean μ and width σ.

An S2 can be modelled as a sum of n boxcar functions, with means μ_i and widths σ_i. The integral of an S2 can be expressed as

$$\sum_{i=0}^{n} I(\mathrm{AFT}_A, \mu_i, \sigma_i) = An, \tag{2}$$

where A denotes the area fraction and the expression can be computationally solved for the area fraction time AFT_A. Values of μ_i can be sampled from a normal distribution of mean 0 and standard deviation from Equation 4.12 to represent diffusion, while widths σ_i can be sampled from a normal distribution of electron pulse widths observed in data.

This method of modelling an S2's width for a given area and drift time were used to identify the relationships in S2 width resolution used in Chapter 4, as well as to evaluate the deviation of pulse widths from expectations for small S2s.

References

[1] Mark Thomson. *Modern Particle Physics*. 1st ed. University Printing House, 2013.

[2] Marco Cirelli, Alessandro Strumia, and Jure Zupan. “Dark Matter”. In: (June 2024).

[3] B J Mount, S Hans, and R Rosero. *LUX-ZEPLIN (LZ) Technical Design Report*. 2017. DOI: `10.48550/ARXIV.1703.09144`. URL: `arxiv.org/abs/1703.09144`.

[4] E. E. Barnard and A. C. Ranyard. “Structure of the Milky Way”. In: *Knowledge: An Illustrated Magazine of Science* 17 (1894), p. 253. URL: `https://nrs.lib.harvard.edu/urn-3:fcor.wolbach:2449481`.

[5] A Secchi. “L’Astronomia in Roma nel pontificato di Pio IX: memoria”. In: (1877). URL: `https://id.lib.harvard.edu/curiosity/expeditions-and-discoveries/38-990084213080203941`.

[6] F. W. Bessel. “On the variations of the proper motions of Procyon and Sirius”. In: 6 (1844), pp. 136–141. DOI: `10.1093/mnras/6.11.136`.

[7] J. C. Kapteyn. “First Attempt at a Theory of the Arrangement and Motion of the Sidereal System”. In: 55 (1922), p. 302. DOI: `10.1086/142670`.

[8] J H Jeans. “The Motions of Stars in a Kapteyn-Universe”. In: *Monthly Notices of the Royal Astronomical Society* 82.3 (1922), pp. 122–132. ISSN: 0035-8711. DOI: `10.1093/mnras/82.3.122`. URL: `https://doi.org/10.1093/mnras/82.3.122`.

[9] Per Olof Lindblad. “Bertil Lindblad’s early work: the two-dimensional classification of stellar spectra at low dispersion”. In: *Journal of Astronomical History and Heritage* 4.2 (2001), pp. 163–170. DOI:

`10.3724/SP.J.1440-2807.2001.02.05`. URL: `https://ui.adsabs.harvard.edu/abs/1978PhDT.......195B`.

[10] J. J. Condon and A. M. Matthews. “ΛCDM Cosmology for Astronomers”. In: (Apr. 2018). DOI: `10.1088/1538-3873/aac1b2`.

[11] Edwin Hubble. “A relation between distance and radial velocity among extra-galactic nebulae”. In: *Proceedings of the National Academy of Sciences* 15.3 (Mar. 1929), pp. 168–173. ISSN: 0027-8424. DOI: `10.1073/pnas.15.3.168`.

[12] Edwin Hubble and Milton L Humason. “The Velocity-Distance Relation among Extra-Galactic Nebulae”. In: 74 (1931), p. 43. DOI: `10.1086/143323`.

[13] Horace W. Babcock. “The rotation of the Andromeda Nebula”. In: *Lick Observatory Bulletins* 19 (1939), pp. 41–51. ISSN: 0075-9317. DOI: `10.5479/ADS/bib/1939LicOB.19.41B`.

[14] Morton S. Roberts. “A High-Resolution 21-CM Hydrogen-Line Survey of the Andromeda Nebula”. In: *The Astrophysical Journal* 144 (May 1966), p. 639. ISSN: 0004-637X. DOI: `10.1086/148645`.

[15] Vera C Rubin and Jr. Ford W. Kent. “Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions”. In: 159 (1970), p. 379. DOI: `10.1086/150317`.

[16] V. C. Rubin, N. Thonnard, and Jr. Ford W. K. “Extended rotation curves of high-luminosity spiral galaxies. IV - Systematic dynamical properties, SA through SC”. In: *The Astrophysical Journal* 225 (Nov. 1978), p. L107. ISSN: 0004-637X. DOI: `10.1086/182804`.

[17] Albert Bosma. “The distribution and kinematics of neutral hydrogen in spiral galaxies of various morphological types”. PhD thesis. University of Groningen, 1978. URL: `https://ned.ipac.caltech.edu/level5/March05/Bosma/Bosma_contents.html`.

[18] D. H. Rogstad and G. S. Shostak. “Gross Properties of Five Scd Galaxies as Determined from 21-CENTIMETER Observations”. In: *The Astrophysical Journal* 176 (Sept. 1972), p. 315. ISSN: 0004-637X. DOI: `10.1086/151636`.

[19] K. C. Freeman. “On the Disks of Spiral and so Galaxies”. In: *The Astrophysical Journal* 160 (June 1970), p. 811. ISSN: 0004-637X. DOI: `10.1086/150474`.

[20] Dante Minniti, Edward W. Olszewski, and Marcia Rieke. “The bulge of M33”. In: *The Astrophysical Journal* 410 (June 1993), p. L79. ISSN: 0004-637X. DOI: `10.1086/186884`.

[21] Julio F Navarro, Carlos S Frenk, and Simon D M White. “The Structure of Cold Dark Matter Halos”. In: *The Astrophysical Journal* 462 (1996), p. 563. DOI: `10.1086/177173`. URL: `https://doi.org/10.1086%2F177173`.

[22] Andrés Granados et al. “GalRotpy: an educational tool to understand and parametrize the rotation curve and gravitational potential of disk-like galaxies”. In: (May 2017). DOI: `10.1016/j.newast.2020.101456`.

[23] Liddle AR and Lyth DH. “The Quasi-Linear Regime”. In: *Cosmological Inflation and Large-Scale Structure*. Cambridge University Press, 2000. Chap. 11.

[24] Heinz Andernach and Fritz Zwicky. *English and Spanish Translation of Zwicky’s (1933) The Redshift of Extragalactic Nebulae*. 2017. DOI: `10.48550/arXiv.1711.01693`.

[25] Ian N. Evans et al. “The Chandra Source Catalog”. In: (May 2010). DOI: `10.1088/0067-0049/189/1/37`.

[26] Douglas Clowe et al. “A Direct Empirical Proof of the Existence of Dark Matter”. In: *The Astrophysical Journal* 648.2 (Sept. 2006), pp. L109–L113. ISSN: 0004-637X. DOI: `10.1086/508162`.

[27] Andrew Liddle. “The Cosmic Microwave Background”. In: *An introduction to modern cosmology*. 2003. Chap. 10. ISBN: 0470848359.

[28] A. A. Penzias and R. W. Wilson. “A Measurement of Excess Antenna Temperature at 4080 Mc/s.” In: *The Astrophysical Journal* 142 (July 1965), p. 419. ISSN: 0004-637X. DOI: `10.1086/148307`.

[29] George F. Smoot. “COBE observations and results”. In: *Conference on 3K cosmology*. ASCE, 1999, pp. 1–10. ISBN: 1563968479. DOI: `10.1063/1.59326`.

[30] C. L. Bennett et al. “Nine-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Final Maps and Results”. In: (Dec. 2012). DOI: `10.1088/0067-0049/208/2/20`.

[31] Planck Collaboration et al. “Planck 2018 results. VII. Isotropy and Statistics of the CMB”. In: (June 2019). DOI: `10.1051/0004-6361/201935201`.

[32] N Aghanim and Y Akrami. “Planck 2018 results. VI. Cosmological parameters”. In: *Astronomy &amp Astrophysics* 641 (2020), A6. DOI: `10.1051/0004-6361/201833910`.

[33] Michael Boylan-Kolchin et al. “Resolving cosmic structure formation with the Millennium-II Simulation”. In: *Monthly Notices of the Royal Astronomical Society* 398.3 (Sept. 2009), pp. 1150–1164. ISSN: 00358711. DOI: `10.1111/j.1365-2966.2009.15191.x`.

[34] C M Baugh. “A primer on hierarchical galaxy formation: the semi-analytical approach”. In: *Reports on Progress in Physics* 69.12 (Dec. 2006), pp. 3101–3156. ISSN: 0034-4885. DOI: `10.1088/0034-4885/69/12/R02`.

[35] Arthur Kosowsky, Milos Milosavljevic, and Raul Jimenez. “Efficient cosmological parameter estimation from microwave background anisotropies”. In: *Physical Review D* 66.6 (Sept. 2002), p. 063007. ISSN: 0556-2821. DOI: `10.1103/PhysRevD.66.063007`.

[36] Raelyn M. Sullivan and Douglas Scott. “The CMB Dipole: Eppur Si Muove”. In: (Nov. 2021). DOI: `10.48550/arXiv.2111.12186`.

[37] Roger G. Clowes et al. “Two close Large Quasar Groups of size = 350 Mpc at z = 1.2”. In: (Aug. 2011). DOI: `10.1111/j.1365-2966.2011.19719.x`.

[38] M. Einasto et al. “The Sloan Great Wall. Morphology and galaxy content”. In: (May 2011). DOI: `10.1088/0004-637X/736/1/51`.

[39] Elcio Abdalla et al. “Cosmology Intertwined: A Review of the Particle Physics, Astrophysics, and Cosmology Associated with the Cosmological Tensions and Anomalies”. In: (Mar. 2022). DOI: `10.1016/j.jheap.2022.04.002`.

[40] W. J. G. de Blok. “The Core-Cusp Problem”. In: (Oct. 2009). DOI: `10.1155/2010/789293`.

[41] Anatoly Klypin et al. “Where Are the Missing Galactic Satellites?” In: *The Astrophysical Journal* 522.1 (Sept. 1999), pp. 82–92. ISSN: 0004-637X. DOI: `10.1086/307643`.

[42] Paul Bode, Jeremiah P. Ostriker, and Neil Turok. “Halo Formation in Warm Dark Matter Models”. In: *The Astrophysical Journal* 556.1 (July 2001), pp. 93–107. ISSN: 0004-637X. DOI: `10.1086/321541`.

[43] David N. Spergel and Paul J. Steinhardt. “Observational Evidence for Self-Interacting Cold Dark Matter”. In: *Physical Review Letters* 84.17 (Apr. 2000), pp. 3760–3763. ISSN: 0031-9007. DOI: `10.1103/PhysRevLett.84.3760`.

[44] Mustafa A. Amin and Mehrdad Mirbabayi. “A lower bound on dark matter mass”. In: (Nov. 2022). DOI: `10.48550/arXiv.2211.09775`.

[45] James M. Cline. “Dark atoms and composite dark matter”. In: (Aug. 2021). DOI: `10.21468/SciPostPhysLectNotes.52`.

[46] David M. Jacobs, Glenn D. Starkman, and Bryan W. Lynn. “Macro Dark Matter”. In: (Oct. 2014). DOI: `10.1093/mnras/stv774`.

[47] Luca Caloni, Martina Gerbino, and Massimiliano Lattanzi. “Updated cosmological constraints on Macroscopic Dark Matter”. In: (May 2021). DOI: `10.1088/1475-7516/2021/07/027`.

[48] Robert J. Scherrer and Michael S. Turner. “On the relic, cosmic abundance of stable, weakly interacting massive particles”. In: *Physical Review D* 33.6 (Mar. 1986), pp. 1585–1589. ISSN: 0556-2821. DOI: `10.1103/PhysRevD.33.1585`.

[49] Diego Restrepo and Julian Calle. *WIMP*. 2018. URL: `https://github.com/restrepo/WIMP?tab=BSD-3-Clause-1-ov-file#readme`.

[50] Giorgio Arcadi et al. “The Waning of the WIMP? A Review of Models, Searches, and Constraints”. In: (Mar. 2017). DOI: `10.1140/epjc/s10052-018-5662-y`.

[51] Benjamin W Lee and Steven Weinberg. “Cosmological Lower Bound on Heavy-Neutrino Masses”. In: *Phys. Rev. Lett.* 39.4 (1977), pp. 165–168. DOI: `10.1103/PhysRevLett.39.165`. URL: `link.aps.org/doi/10.1103/PhysRevLett.39.165`.

[52] Juri Smirnov and John F. Beacom. “TeV-Scale Thermal WIMPs: Unitarity and its Consequences”. In: (Apr. 2019). DOI: `10.1103/PhysRevD.100.043029`.

[53] Anne M. Green. “Primordial Black Holes as a dark matter candidate – a brief overview”. In: (Feb. 2024). DOI: `10.48550/arXiv.2402.15211`.

[54] C Alcock, R A Allsman, D R Alves, et al. “The MACHO Project Hubble Space Telescope Follow-Up: Preliminary Results on the Location of the Large Magellanic Cloud Microlensing Source Stars”. In: *The Astrophysical Journal* 552.2 (2001), pp. 582–590. DOI: `10.1086/320554`. URL: `https://doi.org/10.1086%2F320554`.

[55] Riccardo Scarpa. “Modified Newtonian Dynamics, an Introductory Review”. In: *AIP Conference Proceedings*. AIP, 2006, pp. 253–265. DOI: `10.1063/1.2189141`.

[56] Richard Brito, Vitor Cardoso, and Paolo Pani. “Superradiance – the 2020 Edition”. In: (Jan. 2015). DOI: `10.1007/978-3-030-46622-0`.

[57] Ciaran O’Hare. *Gitlab, "How to Make a Plot"*. URL: `https://github.com/cajohare/HowToMakeAPlot/tree/main?tab=readme-ov-file`.

[58] Kiwoon Choi, Eung Jin Chun, and Hyungdo Kim. “Supersymmetry hierarchy problems and anomalous horizontal U(1) symmetry”. In: *Physics Letters B* 394.1-2 (1997), pp. 89–98. DOI: `10.1016/s0370-2693(96)01674-7`. URL: `https://doi.org/10.1016%2Fs0370-2693%2896%2901674-7`.

[59] Th. Kaluza. “On the Unification Problem in Physics”. In: (Mar. 2018). DOI: `10.1142/S0218271818700017`.

[60] L Horoto and F G Scholtz. "A New Perspective on Kaluza-Klein Theories". In: (Apr. 2024). DOI: `10.48550/arXiv.2404.05302`.

[61] Nima Arkani-Hamed, Andrew G. Cohen, and Howard Georgi. "Electroweak symmetry breaking from dimensional deconstruction". In: *Physics Letters B* 513.1-2 (July 2001), pp. 232–240. ISSN: 03702693. DOI: `10.1016/S0370-2693(01)00741-9`.

[62] Martin Schmaltz, Daniel Stolarski, and Jesse Thaler. "The Bestest Little Higgs". In: (June 2010). DOI: `10.1007/JHEP09(2010)018`.

[63] Jodi Cooley et al. "Report of the Topical Group on Particle Dark Matter for Snowmass 2021". In: (Sept. 2022). DOI: `10.48550/arXiv.2209.07426`.

[64] David E. Kaplan, Markus A. Luty, and Kathryn M. Zurek. "Asymmetric Dark Matter". In: (Jan. 2009). DOI: `10.1103/PhysRevD.79.115016`.

[65] Kalliopi Petraki and Raymond R. Volkas. "Review of asymmetric dark matter". In: (May 2013). DOI: `10.1142/S0217751X13300287`.

[66] C. Boehm and P. Fayet. "Scalar Dark Matter candidates". In: (May 2003). DOI: `10.1016/j.nuclphysb.2004.01.015`.

[67] Eric Kuflik et al. "Elastically Decoupling Dark Matter". In: (Dec. 2015). DOI: `10.1103/PhysRevLett.116.221302`.

[68] Nicolás Bernal et al. "The Dawn of FIMP Dark Matter: A Review of Models and Constraints". In: (June 2017). DOI: `10.1142/S0217751X1730023X`.

[69] Igor G. Irastorza. "An introduction to axions and their detection". In: (Sept. 2021). DOI: `10.21468/SciPostPhysLectNotes.45`.

[70] Marco Fabbrichesi, Emidio Gabrielli, and Gaia Lanfranchi. "The Dark Photon". In: (May 2020). DOI: `10.1007/978-3-030-62519-1`.

[71] Basudeb Dasgupta and Joachim Kopp. "Sterile Neutrinos". In: (June 2021). DOI: `10.1016/j.physrep.2021.06.002`.

[72] The LZ Collaboration et al. "A search for new physics in low-energy electron recoils from the first LZ exposure". In: (July 2023). DOI: `10.1103/PhysRevD.108.072006`.

[73] D S Akerib, A K Al Musalhi, S K Alsum, et al. *Enhancing the sensitivity of the LUX-ZEPLIN (LZ) dark matter experiment to low energy signals.* 2021. DOI: `10.48550/ARXIV.2101.08753`. URL: `arxiv.org/abs/2101.08753`.

[74] ATLAS Collaboration. "Search for invisible Higgs boson decays in vector boson fusion at sqrt(s) = 13 TeV with the ATLAS detector". In: (Sept. 2018). DOI: `10.1016/j.physletb.2019.04.024`.

[75] CMS Collaboration. "Dark sector searches with the CMS experiment". In: (May 2024).

[76] LHCb collaboration et al. "Search for dark photons produced in 13 TeV {pp} collisions". In: (Oct. 2017). DOI: `10.1103/PhysRevLett.120.061801`.

[77] Felix Kling and Pablo Quílez. "ALP Searches at the LHC: FASER as a Light Shining through Walls Experiment". In: (Apr. 2022). DOI: `10.1103/PhysRevD.106.055036`.

[78] Giulio Aielli et al. "The Road Ahead for CODEX-b". In: (Mar. 2022). DOI: `10.48550/arXiv.2203.07316`.

[79] Carlos Pérez de los Heros. "Status, Challenges and Directions in Indirect Dark Matter Searches". In: (Aug. 2020). DOI: `10.3390/sym12101648`.

[80] Marc Schumann. "Direct Detection of WIMP Dark Matter: Concepts and Status". In: (Mar. 2019). DOI: `10.1088/1361-6471/ab2ea5`.

[81] 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 (Dec. 1996), pp. 87–112. ISSN: 09276505. DOI: `10.1016/S0927-6505(96)00047-3`.

[82] Julien Billard et al. "Direct Detection of Dark Matter – APPEC Committee Report". In: (Apr. 2021). DOI: `10.1088/1361-6633/ac5754`.

[83] Richard H. Helm. “Inelastic and Elastic Scattering of 187-Mev Electrons from Selected Even-Even Nuclei”. In: *Physical Review* 104.5 (Dec. 1956), pp. 1466–1475. ISSN: 0031-899X. DOI: `10.1103/PhysRev.104.1466`.

[84] Jelle Aalbers et al. *JelleAalbers/wimprates: v0.5.0.* Feb. 2023. DOI: `10.5281/zenodo.7636982`. URL: `https://doi.org/10.5281/zenodo.7636982`.

[85] J Aalbers, D S Akerib, C W Akerlof, et al. *First Dark Matter Search Results from the LUX-ZEPLIN (LZ) Experiment.* 2022. DOI: `10.48550/ARXIV.2207.03764`. URL: `arxiv.org/abs/2207.03764`.

[86] D Baxter, I M Bloch, E Bodnia, et al. “Recommended conventions for reporting results from direct dark matter searches”. In: *The European Physical Journal C* 81.10 (2021). DOI: `10.1140/epjc/s10052-021-09655-y`. URL: `https://doi.org/10.1140%2Fepjc%2Fs10052-021-09655-y`.

[87] Robert James. “Signals, backgrounds and statistical inference for dark matter direct detection experiments”. PhD thesis. London: University College London, 2023. URL: `https://discovery.ucl.ac.uk/id/eprint/10190030/`.

[88] XENON Collaboration et al. “XENONnT WIMP Search: Signal & Background Modeling and Statistical Inference”. In: (June 2024). DOI: `10.48550/arXiv.2406.13638`.

[89] Francis Froborg and Alan R Duffy. “Annual Modulation in Direct Dark Matter Searches”. In: (Mar. 2020). DOI: `10.1088/1361-6471/ab8e93`.

[90] Anne M. Green and Ben Morgan. “The median recoil direction as a WIMP directional detection signal”. In: (Feb. 2010). DOI: `10.1103/PhysRevD.81.061301`.

[91] J. Billard, F. Mayet, and D. Santos. “Assessing the discovery potential of directional detection of Dark Matter”. In: (Oct. 2011). DOI: `10.1103/PhysRevD.85.035006`.

[92] Ciaran A. J. O'Hare. "Theoretical prospects for directional WIMP detection". In: (Oct. 2015). DOI: `10.48550/arXiv.1510.04079`. URL: `https://discovery.ucl.ac.uk/id/eprint/10190030`.

[93] Glen Cowan et al. "Power-Constrained Limits". In: (May 2011). DOI: `10.48550/arXiv.1105.3166`.

[94] A L Read. "Presentation of search results: the CLs technique". In: *Journal of Physics G: Nuclear and Particle Physics* 28.10 (Oct. 2002), pp. 2693–2704. ISSN: 0954-3899. DOI: `10.1088/0954-3899/28/10/313`.

[95] R S James, J Palmer, A Kaboth, et al. "FlameNEST: explicit profile likelihoods with the Noble Element Simulation Technique". In: *Journal of Instrumentation* 17.08 (Aug. 2022), P08012. ISSN: 1748-0221. DOI: `10.1088/1748-0221/17/08/p08012`. URL: `http://dx.doi.org/10.1088/1748-0221/17/08/P08012`.

[96] The DarkSide Collaboration et al. "Low-Mass Dark Matter Search with the DarkSide-50 Experiment". In: (Feb. 2018). DOI: `10.1103/PhysRevLett.121.081307`.

[97] XENON Collaboration et al. "First Dark Matter Search with Nuclear Recoils from the XENONnT Experiment". In: (Mar. 2023). DOI: `10.1103/PhysRevLett.131.041003`.

[98] PandaX-II Collaboration et al. "Dark Matter Results From 54-Ton-Day Exposure of PandaX-II Experiment". In: (Aug. 2017). DOI: `10.1103/PhysRevLett.119.181302`.

[99] J Aalbers, S S AbdusSalam, K Abe, et al. "A next-generation liquid xenon observatory for dark matter and neutrino physics". In: *Journal of Physics G: Nuclear and Particle Physics* 50.1 (2022), p. 13001. DOI: `10.1088/1361-6471/ac841a`. URL: `https://doi.org/10.1088%2F1361-6471%2Fac841a`.

[100] C. E. Aalseth et al. "DarkSide-20k: A 20 Tonne Two-Phase LAr TPC for Direct Dark Matter Detection at LNGS". In: (July 2017). DOI: `10.1140/epjp/i2018-11973-4`.

[101] Ciaran A. J. O'Hare. "Fog on the horizon: a new definition of the neutrino floor for direct dark matter searches". In: (Sept. 2021). DOI: `10.1103/PhysRevLett.127.251802`.

[102] E. Aprile et al. "First Indication of Solar $
8$B Neutrinos via Coherent Elastic Neutrino-Nucleus Scattering with XENONnT". In: (Aug. 2024). DOI: `10.1103/PhysRevLett.133.191002`.

[103] PandaX Collaboration et al. "First Indication of Solar $
8$B Neutrino Flux through Coherent Elastic Neutrino-Nucleus Scattering in PandaX-4T". In: (July 2024).

[104] Ioannis Katsioulas. "NEWS-G, Light dark matter search with a Spherical Proportional Counter, First results and Future prospects". In: (Sept. 2018). DOI: `10.48550/arXiv.1809.02485`.

[105] CRESST Collaboration et al. "First results from the CRESST-III low-mass dark matter program". In: (Mar. 2019). DOI: `10.1103/PhysRevD.100.102002`.

[106] P Agnes, I F M Albuquerque, and T Alexander. *Search for dark matter-nucleon interactions via Migdal effect with DarkSide-50.* 2022. DOI: `10.48550/ARXIV.2207.11967`. URL: `arxiv.org/abs/2207.11967`.

[107] E Aprile, J Aalbers, F Agostini, et al. "Light Dark Matter Search with Ionization Signals in XENON1T". In: *Physical Review Letters* 123.25 (2019). DOI: `10.1103/physrevlett.123.251801`. URL: `doi.org/10.1103%2Fphysrevlett.123.251801`.

[108] E. Aprile et al. "Search for Light Dark Matter Interactions Enhanced by the Migdal effect or Bremsstrahlung in XENON1T". In: (July 2019). DOI: `10.1103/PhysRevLett.123.241803`.

[109] CRESST Collaboration et al. "Results on sub-GeV Dark Matter from a 10 eV Threshold CRESST-III Silicon Detector". In: (Dec. 2022). DOI: `10.1103/PhysRevD.107.122003`.

[110] Jean-Marie Coquillat. "Latest results from the NEWS-G dark matter experiment". In: *Proceedings of XVIII International Conference on Topics in Astroparticle and Underground Physics — PoS(TAUP2023)*. Trieste, Italy: Sissa Medialab, Jan. 2024, p. 042. DOI: `10.22323/1.441.0042`.

[111] L. Balogh et al. "Exploring light dark matter with the DarkSPHERE spherical proportional counter electroformed underground at the Boulby Underground Laboratory". In: (Jan. 2023). DOI: `10.48550/arXiv.2301.05183`.

[112] Masahiro Ibe, Wakutaka Nakano, Yutaro Shoji, et al. "Migdal effect in dark matter direct detection experiments". In: *Journal of High Energy Physics* 2018.3 (2018). DOI: `10.1007/jhep03(2018)194`. URL: `doi.org/10.1007%2Fjhep03%282018%29194`.

[113] H M Araújo, S N Balashov, and J E Borg. F M Brunbauer. *The MIGDAL experiment: Measuring a rare atomic process to aid the search for dark matter.* 2022. DOI: `10.48550/ARXIV.2207.08284`. URL: `arxiv.org/abs/2207.08284`.

[114] M. F. Albakry et al. "A Search for Low-mass Dark Matter via Bremsstrahlung Radiation and the Migdal Effect in SuperCDMS". In: (Feb. 2023). DOI: `10.1103/PhysRevD.107.112013`.

[115] SENSEI Collaboration et al. "SENSEI: First Direct-Detection Results on sub-GeV Dark Matter from SENSEI at SNOLAB". In: (Dec. 2023). DOI: `10.48550/arXiv.2312.13342`.

[116] Laura Baudis. "Dual-phase xenon time projection chambers for rare-event searches". In: (Nov. 2023). DOI: `10.48550/arXiv.2311.05320`.

[117] D.S. Akerib, S Alsum, H.M. Araújo, et al. "Investigation of background electron emission in the LUX detector". In: *Physical Review D* 102.9 (2020). DOI:

`10.1103/physrevd.102.092004`. URL: `doi.org/10.1103%2Fphysrevd.102.092004`.

[118] Andrew Stevens. "Development of the ionisation signal analysis and the thermometry system for the LUX-ZEPLIN direct dark matter detection experiment". PhD thesis. University of Oxford, 2021. URL: `https://ora.ox.ac.uk/objects/uuid:07b7ffcd-97a4-48c3-865e-190a2c37dcf2`.

[119] M. 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: (Feb. 2021). DOI: `10.3390/instruments5010013`.

[120] Martin Berger et al. *XCOM: Photon Cross Section Database (version 1.2)*. en. Mar. 1999. URL: `http://physics.nist.gov/xcom`.

[121] M Szydagis, N Barry, K Kazkaz, et al. "NEST: a comprehensive model for scintillation yield in liquid xenon". In: *Journal of Instrumentation* 6.10 (2011), P10002–P10002. DOI: `10.1088/1748-0221/6/10/p10002`. URL: `https://doi.org/10.1088%2F1748-0221%2F6%2F10%2Fp10002`.

[122] Laura Baudis, Patricia Sanchez-Lucas, and Kevin Thieme. "A measurement of the mean electronic excitation energy of liquid xenon". In: (Sept. 2021). DOI: `10.1140/epjc/s10052-021-09834-x`.

[123] Tadayoshi Doke et al. "Absolute Scintillation Yields in Liquid Argon and Xenon for Various Particles". In: *Japanese Journal of Applied Physics* 41.Part 1, No. 3A (Mar. 2002), pp. 1538–1545. ISSN: 0021-4922. DOI: `10.1143/JJAP.41.1538`.

[124] Hamamatsu. *Photomultiplier Tubes and Assemblies*. URL: `https://www.hamamatsu.com/content/dam/hamamatsu-photonics/sites/documents/99_SALES_LIBRARY/etd/High_energy_PMT_TPMZ0003E.pdf`.

[125] B J Mount, S Hans, R Rosero, et al. *LUX-ZEPLIN (LZ) Technical Design Report*. 2017. DOI: `10.48550/arXiv.1703.09144`.

[126] D S Akerib, C W Akerlof, and D Yu. Akimov. "The LUX-ZEPLIN (LZ) radioactivity and cleanliness control programs". In: *The European Physical Journal C* 80.11 (2020). DOI: `10.1140/epjc/s10052-020-8420-x`. URL: `doi.org/10.1140%2Fepjc%2Fs10052-020-8420-x`.

[127] IAEA Nuclear Data Section. *Live Chart of Nuclides.* URL: `https://www-nds.iaea.org/relnsd/vcharthtml/VChartHTML.html`.

[128] J Aalbers, D S Akerib, A K Al Musalhi, et al. *Background Determination for the LUX-ZEPLIN (LZ) Dark Matter Experiment.* 2022. DOI: `10.48550/ARXIV.2211.17120`. URL: `arxiv.org/abs/2211.17120`.

[129] N. Ackerman et al. "Observation of Two-Neutrino Double-Beta Decay in Xe-136 with EXO-200". In: (Aug. 2011). DOI: `10.1103/PhysRevLett.107.212501`.

[130] KamLAND-Zen Collaboration. "Measurement of the double-beta decay half-life of 136Xe with the KamLAND-Zen experiment". In: (Jan. 2012). DOI: `10.1103/PhysRevC.85.045504`.

[131] D. Akimov et al. "Observation of Coherent Elastic Neutrino-Nucleus Scattering". In: (Aug. 2017). DOI: `10.1126/science.aao0990`.

[132] R Linehan, R L Mannino, A Fan, 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 (2022), p. 165955. DOI: `10.1016/j.nima.2021.165955`. URL: `doi.org/10.1016%2Fj.nima.2021.165955`.

[133] R Linehan. "High Voltage Electrode Development and the LZ Experiment's WIMP Search". PhD thesis. Stanford University, 2022. URL: `https://searchworks.stanford.edu/view/14310527`.

[134] B Edwards, H M Araújo, V Chepel, et al. "Measurement of single electron emission in two-phase xenon". In: *Astroparticle Physics* 30.2 (2008), pp. 54–57. ISSN: 0927-6505. DOI:

`https://doi.org/10.1016/j.astropartphys.2008.06.006`. URL: `https://www.sciencedirect.com/science/article/pii/S0927650508000789`.

[135] E Santos, B Edwards, V Chepel, et al. “Single electron emission in two-phase xenon with application to the detection of coherent neutrino-nucleus scattering”. In: *Journal of High Energy Physics* 2011.12 (Dec. 2011). ISSN: 1029-8479. DOI: `10.1007/jhep12(2011)115`. URL: `http://dx.doi.org/10.1007/JHEP12(2011)115`.

[136] E Aprile, M Alfonsi, K Arisaka, et al. “Observation and applications of single-electron charge signals in the XENON100 experiment”. In: *Journal of Physics G: Nuclear and Particle Physics* 41.3 (Feb. 2014), p. 35201. DOI: `10.1088/0954-3899/41/3/035201`. URL: `https://dx.doi.org/10.1088/0954-3899/41/3/035201`.

[137] E Aprile, K Abe, F Agostini, et al. “Emission of single and few electrons in XENON1T and limits on light dark matter”. In: *Physical Review D* 106.2 (July 2022). ISSN: 2470-0029. DOI: `10.1103/physrevd.106.022001`. URL: `http://dx.doi.org/10.1103/PhysRevD.106.022001`.

[138] P Barrow, L Baudis, D Cichon, et al. “Qualification tests of the R11410-21 photomultiplier tubes for the XENON1T detector”. In: *Journal of Instrumentation* 12.01 (Jan. 2017), P01024–P01024. ISSN: 1748-0221. DOI: `10.1088/1748-0221/12/01/p01024`. URL: `http://dx.doi.org/10.1088/1748-0221/12/01/P01024`.

[139] I Khurana. “Mitigating Accidental Coincidence Backgrounds in the LZ experiment: demonstration of a Machine Learning Approach using first data”. PhD thesis. University College London, 2023. URL: `https://discovery.ucl.ac.uk/id/eprint/10185176/`.

[140] G R Araujo, T Pollmann, and A Ulrich. “Photoluminescence response of acrylic (PMMA) and polytetrafluoroethylene (PTFE) to ultraviolet light: Limits on low-intensity photoluminescence in support materials of rare-event search

experiments”. In: *The European Physical Journal C* 79.8 (Aug. 2019). ISSN: 1434-6052. DOI: `10.1140/epjc/s10052-019-7152-2`. URL: `http://dx.doi.org/10.1140/epjc/s10052-019-7152-2`.

[141] Peter Sorensen. “Anisotropic diffusion of electrons in liquid xenon with application to improving the sensitivity of direct dark matter searches”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 635 (2011), pp. 41–43. DOI: `10.1016/j.nima.2011.01.089`. URL: `doi.org/10.1016%2Fj.nima.2011.01.089`.

[142] The LZ Collaboration et al. “LUX-ZEPLIN (LZ) Conceptual Design Report”. In: (Sept. 2015). DOI: `10.48550/arXiv.1509.02910`.

[143] LUX Collaboration et al. “Improved Limits on Scattering of Weakly Interacting Massive Particles from Reanalysis of 2013 LUX data”. In: (Dec. 2015). DOI: `10.1103/PhysRevLett.116.161301`.

[144] M. Szydagis et al. “A Review of NEST Models, and Their Application to Improvement of Particle Identification in Liquid Xenon Experiments”. In: (Nov. 2022). DOI: `10.48550/arXiv.2211.10726`.

[145] D.S Akerib, C.W. Akerlof, S.K. Alsum, et al. “Projected WIMP sensitivity of the LUX-ZEPLIN dark matter experiment”. In: *Physical Review D* 101.5 (2020). DOI: `10.1103/physrevd.101.052002`. URL: `https://doi.org/10.1103%2Fphysrevd.101.052002`.

[146] O Njoya, T Tsang, M Tarka, 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. DOI: `10.1016/j.nima.2020.163965`. URL: `doi.org/10.1016%2Fj.nima.2020.163965`.

[147] E Aprile and T Doke. “Liquid xenon detectors for particle physics and astrophysics”. In: *Reviews of Modern Physics* 82.3 (2010), pp. 2053–2097. DOI: `10.1103/revmodphys.82.2053`. URL: `doi.org/10.1103%2Frevmodphys.82.2053`.

[148] D S Akerib, S Alsum, H M Araújo, et al. “Discrimination of electronic recoils from nuclear recoils in two-phase xenon time projection chambers”. In: *Physical Review D* 102.11 (Dec. 2020). ISSN: 2470-0029. DOI: `10.1103/physrevd.102.112002`. URL: `http://dx.doi.org/10.1103/PhysRevD.102.112002`.

[149] J. Aalbers et al. “The Data Acquisition System of the LZ Dark Matter Detector: FADR”. In: (May 2024). DOI: `10.48550/arXiv.2405.14732`.

[150] P. Brás et al. “A machine learning-based methodology for pulse classification in dual-phase xenon time projection chambers”. In: (Jan. 2022). DOI: `10.1140/epjc/s10052-022-10502-x`.

[151] 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: 01689002. DOI: `10.1016/j.nima.2011.11.063`.

[152] Aaron Manalaysay. “Response of Liquid Xenon to Low-Energy Ionizing Radiation and its use in the XENON10 Dark Matter Search”. PhD thesis. University of Florida, 2009. URL: `https://ufl-flvc.primo.exlibrisgroup.com/permalink/01FALSC_UFL/6ad6fc/alma990256641850306597`.

[153] Vance W. Berger and YanYan Zhou. “Kolmogorov–Smirnov Test: Overview”. In: *Wiley StatsRef: Statistics Reference Online*. Wiley, Sept. 2014. DOI: `10.1002/9781118445112.stat06558`.

[154] Ryan Smith. “Searches for weakly interacting and multiply interacting massive particles with LZ, and characterization of nuclear recoils in helium”. PhD thesis. University of California, Berkeley, 2023.

[155] A K Al Musalhi. “Searches for rare nuclear decays in the LUX-ZEPLIN experiment”. PhD thesis. University of Oxford, 2023. URL: `https://ora.ox.ac.uk/objects/uuid:cf8cd7b7-6782-49e8-8862-a65fa702295e`.

[156] LUX Collaboration. “Ultra-Low Energy Calibration of LUX Detector using 127Xe Electron Capture”. In: (Sept. 2017). DOI: `10.1103/PhysRevD.96.112011`.

[157] Maris Arthurs. “Radon Reduction and the First Science Results of the LZ Experiment”. PhD thesis. University of Michigan, 2022. DOI: `10.7302/5960`. URL: `https://deepblue.lib.umich.edu/handle/2027.42/174229`.

[158] Tyler Anderson. “The LZ Dark Matter WIMP Search and Treatment of Fundamental Signals”. PhD thesis. Stanford University, 2023. URL: `https://searchworks.stanford.edu/view/in00000001612`.

[159] D S Akerib, C W Akerlof, A Alqahtani, et al. “Simulations of events for the LUX-ZEPLIN (LZ) dark matter experiment”. In: *Astroparticle Physics* 125 (Feb. 2021), p. 102480. ISSN: 0927-6505. DOI: `10.1016/j.astropartphys.2020.102480`. URL: `http://dx.doi.org/10.1016/j.astropartphys.2020.102480`.

[160] J. Allison et al. “Recent developments in Geant4”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 835 (Nov. 2016), pp. 186–225. ISSN: 01689002. DOI: `10.1016/j.nima.2016.06.125`.

[161] Luigi Rolandi, Werner Riegler, and Walter Blum. *Particle Detection with Drift Chambers*. Vol. 0. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008. ISBN: 978-3-540-76683-4. DOI: `10.1007/978-3-540-76684-1`.

[162] G. Martínez-Lema et al. “Calibration of the NEXT-White detector using 83mKr decays”. In: (Apr. 2018). DOI: `10.1088/1748-0221/13/10/P10014`.

[163] The LZ Collaboration. “Cosmogenic production of 37Ar in the context of the LUX-ZEPLIN experiment”. In: (Jan. 2022). DOI: `10.1103/PhysRevD.105.082004`.

[164] D S Akerib, S Alsum, H M Araújo, et al. "Improving sensitivity to low-mass dark matter in LUX using a novel electrode background mitigation technique". In: *Physical Review D* 104.1 (2021). DOI: `10.1103/physrevd.104.012011`. URL: `doi.org/10.1103%2Fphysrevd.104.012011`.

[165] Martin Berger, J Coursey, and M Zucker. *ESTAR, PSTAR, and ASTAR: Computer Programs for Calculating Stopping-Power and Range Tables for Electrons, Protons, and Helium Ions (version 1.21)*. en. Mar. 1999. URL: `http://physics.nist.gov/Star`.

[166] NEXT Collaboration et al. "Electron drift properties in high pressure gaseous xenon". In: (Apr. 2018). DOI: `10.1088/1748-0221/13/07/P07013`.

[167] C. A. O. Henriques et al. "Neutral Bremsstrahlung emission in xenon unveiled". In: (Feb. 2022). DOI: `10.1103/PhysRevX.12.021005`.

[168] R.D. Deslattes et al. *X-ray Transition Energies*. 2005. URL: `http://physics.nist.gov/XrayTrans`.

[169] J. Renner et al. "Initial results on energy resolution of the NEXT-White detector". In: (Aug. 2018). DOI: `10.1088/1748-0221/13/10/P10020`.

[170] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. "Boosting and Additive Trees". In: 2009, pp. 337–387. DOI: `10.1007/978-0-387-84858-7{\_}10`.

[171] F Pedregosa et al. "Scikit-learn: Machine Learning in Python". In: *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830. DOI: `10.48550/arXiv.1201.0490`.

[172] Trevor Hastie et al. "Multi-class AdaBoost". In: *Statistics and Its Interface* 2.3 (2009), pp. 349–360. ISSN: 19387989. DOI: `10.4310/SII.2009.v2.n3.a8`.

[173] "Kendall Rank Correlation Coefficient". In: *The Concise Encyclopedia of Statistics*. New York, NY: Springer New York, pp. 278–281. DOI: `10.1007/978-0-387-32833-1{\_}211`.

[174] Andrew P. Bradley. "The use of the area under the ROC curve in the evaluation of machine learning algorithms". In: *Pattern Recognition* 30.7 (July 1997), pp. 1145–1159. ISSN: 00313203. DOI: `10.1016/S0031-3203(96)00142-2`.