Mitigating Accidental Coincidence Backgrounds in the LZ experiment: demonstration of a Machine Learning Approach using first data.

Ishan Khurana

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

of

University College London.

Department of Physics and Astronomy

University College London

October 20, 2023

I, Ishan Khurana, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the work.

Dedicated to my ma Annie,
for battling the world to bring us here,
and to my late father Pankaj Khurana,
you would have enjoyed the graduation.

Acknowledgements

It takes a village to write a PhD. And I can't thank enough the lovely people in my village.

Firstly, my supervisor Jim Dobson. A brilliantly organised, tenacious and inspiring physicist who was a joy to work with. I extend my heartfelt gratitude for the rigorous technical discussions, regular feedback, support, and a commitment to professionalism that should set the standard.

I thank Tim Scanlon who, as my masters supervisor and as the second supervisor on my PhD, helped me develop the analytical vision required to deliver results.

I started at UCL in 2014 and finish now in 2023. During my nine years at this institution, I found inspiration, encouragement, and much needed laughter from many in the physics department, as well as support for my various projects that have invariably shaped the way I understand the world. I would like to give special thanks to Nick Nicolaou and Kelvin Vine.

I thank Peter for all his support over the years. I can't begin to enumerate all the ways in which you have put my mind at ease. Hail Hail!

All of my friends have in some way been an anchor as I've worked on this PhD, and I thank each and every one of you. A special mention goes to Jax, who's consistent support, chat and love for music provided sustenance. To my dear pal Chairman Mais, who does great bits, in a world where not enough people do bits.

I thank Gurminder K Bhambra for her friendship, mentorship, and for the opportunity to work on the Connected Sociologies Project.

Lastly, I thank Lukas Kikuchi, for everything we've worked on and everything we plan to work on.

Abstract

Understanding the nature of dark matter has been one of the pre-eminent problems in particle physics, given the compelling body of astrophysical and cosmological evidence pointing towards the existence and abundance of an elusive, invisible form of matter. A theoretically well motivated candidate is the Weakly Interacting Massive Particle (WIMP). LUX-ZEPLIN (LZ) is a direct detection experiment searching for WIMP interactions with xenon nuclei. Based in the Sanford Underground Research Facility (SURF) in Lead, South Dakota, it monitors a 7 tonne active volume of liquid xenon and began its first science run (SR1) on the 23rd of December 2021 consisting of a 60 live day exposure using a fiducial mass of 5.5 tonnes. First results from the SR1 data report a world-leading limit for spin-independent scattering at $36\,\mathrm{GeV/c^2}$, rejecting cross sections above $\sigma_{\mathrm{SI}} = 9.2 \times 10^{-48}\,\mathrm{cm}^2$ at the 90% confidence level.

With increased access to computing power, another area of research that has grown significantly in the last decade is that of applied machine learning (ML). Realising the potential gains in physics analyses that can be made using ML techniques requires a meshing of expertise in both detector physics and ML tools. The research reported in this thesis presents a framework for background rejection in the LZ data using Boosted Decision Trees (BDT). The effectiveness of the approach is demonstrated on accidental coincidence events, though it can be generalised to other analyses or ML models. Crucially, the BDT based rejection of accidental coincidence events reduces the rate of such backgrounds by a factor of four compared with the official SR1 analysis, at a relatively small loss in signal efficiency. Thus it brings the accidentals rate to a level required for LZ to reach its projected sensitivity.

Impact Statement

In times when our universities are under increasing pressure to operate like businesses with curricula dictated by the market, it is imperative to make the case for fundamental research regardless of its material impact in a profit driven society. At the time of writing, several humanities departments across the UK are being shut down as their economic impact on society is deemed insufficient. However, like history, philosophy, or sociology, particle physics and cosmology contribute to humanity's understanding of the world around us. Insisting that these immaterial contributions, that do not directly lead to start-ups, are valued is essential to protecting the integrity and the funding of fundamental research.

The work conducted during this PhD has made significant contributions to the LZ dark matter experiment, a world-leading research project aiming to answer fundamental questions in particle physics and modern cosmology. The methods developed to apply machine learning techniques to the LZ data provide a foundation for researchers to build upon and improve future analyses. Furthermore, through LZ's operations at SURF in Lead, South Dakota science is becoming a part of the local community culture, adding to its vibrancy and vitality.

Making publicly funded research accessible to all is integral to creating a well educated citizenry, and long term outreach work contributes to this. Given the knowledge gathered over the course of the PhD in both particle physics and data intensive sciences, I was able to deliver courses in statistics and inference in local high schools, which included a critical discussion on the uses of statistics to make knowledge claims in popular discourse. Additionally, I contributed to the dissemination of LZ's research through a number of talks in schools in East London.

Lastly, the data science skills developed whilst undertaking the research presented in this thesis allowed me to make contributions to economic policy through my work at the think tank Autonomy. Much of the work conducted at Autonomy researching the impact of the COVID-19 pandemic was regularly covered in major media outlets (Financial Times, Guardian, ITV etc). Notably, I also developed a microsimulation model to evaluate the impact of changes to the tax-benefit system on the UK's household income distribution, child poverty figures, and inequality statistics. This was used to assess the viability of a universal basic income (UBI) in Wales, and the findings were presented to the Welsh Assembly.

Contents

List of Figures

List of Tables

Glossary & List of Acronyms

WIMP: Weakly interacting massive particle.

TPC: Time Projection Chamber. The central vessel containing the detector target material.

S1: Primary scintillation light produced through scatters in the TPC.

S2: Secondary signal produced through the detection of ionisation electrons created by scatters in the TPC.

g1: The light collection efficiency i.e. the fraction of S1 photons produced that are detected by the PMT arrays.

g2: The charge collection efficiency i.e. the fraction of S2 photons produced that is detected by the PMT arrays.

S1c: S1 signal corrected for spatial differences in light collection efficiency. The signal is normalised to the centre of the TPC.

S1c: S2 signal corrected for spatial differences in charge collection efficiency. The signal is normalised to the centre of the liquid surface.

VUV: Vacuum Ultraviolet.

NR: Nuclear recoil.

ER: Electron recoil.

RQ: Reduced quantity. A parameterisation of the characteristics of a pulse or event. For example, the full width at half maximum of a pulse or position of an event.

DT: Drift time. The time taken to drift ionisation electrons from the interaction site out of the liquid and into the gas phase.

MERCURY: A position reconstruction algorithm for dual phase Xenon TPCs.

UDT: Unphysical drift time. A drift time corresponding to a length that is longer than the length of the TPC.

FWHM: Full width at half maximum.

AFT: Area fraction time. Used as, for instance, AFT50 to denote the time at which 50% of the total pulse area has been recorded.

RMS: Root mean square. Used to denote the RMS width of a pulse.

TBA: Top Bottom Asymmetry. A value between -1 and 1 used to measure the distribution of light between the top and bottom PMT arrays. A TBA of -1 (1) corresponds to all the light being detected in the bottom (top) array.

Isolated S1 (S2): An S1 (S2) signal without an associated S2 (S1) signal.

Accidental Event: An accidental coincidence of an isolated S1 and isolated S2 signal which can appear as an event originating from a scatter in the TPC.

Chopstitch Accidental: A synthetic event generated by ”stitching“ together an isolated S1 and an isolated S2 event to look like an accidental event.

Chopstitch NR: A synthetic NR event generated by ”stitching“ together an S1 and an S2 from calibration data.

BDT: Boosted decision tree.

PLR: Profile Likelihood Ratio.

Chapter 1

Introduction

There is a significant body of evidence motivating a cold non-luminous, non-baryonic "dark matter" particle that accounts for the majority of the matter in the universe. A leading class of Beyond the Standard Model particles that meets the requirements to make up this dark matter are the Weakly Interacting Massive Particles (WIMPs). These are a set of stable, massive particles that interact with ordinary matter via gravity, and an additional force with an interaction rate that is at least as weak as the Weak force.

Searches for WIMP dark matter are currently underway at the Sanford Underground Research Facility (SURF) in Lead South Dakota by the LUX-ZEPLIN (LZ) collaboration. The LZ experiment is a dual-phase liquid xenon (LXe) time projection chamber (TPC) built in an ultra-clean environment to operate as a low background rare event astroparticle observatory. Completed during the COVID-19 pandemic, it successfully started taking data in December 2019 and plans to record a 1000 days worth of data during the course of its lifetime.

The focus of this thesis is on the original work undertaken to reject an important background for the LZ experiment, the accidental coincidence background.

Chapter 2 gives an overview of the evidence for dark matter at the galaxy, galaxy cluster and cosmological scale. Viable dark matter candidates that meet the constraints set by the astrophysical evidence are discussed with a focus on WIMPs. This is followed by an overview of the current experimental landscape, with a emphasis on WIMP direct detection. Lastly, the operating principles of a dual phase

LXe detector are discussed, including an overview of the mechanism by which scintillation light is produced in xenon.

Chapter 3 introduces the LZ detector and gives an overview of its various subsystems. The two main components of the TPC, the PMT arrays and the electric field, are discussed. The two optically separated veto detectors, the *Skin* and the *Outer Detector*, that are used to enhance the discrimination of neutron and gamma ray backgrounds are described. And the chapter concludes with an overview of LZ's planned calibration programme as well as a discussion on the dominant backgrounds expected in a 1000 day run.

Chapter 4 presents LZ's first search for WIMPs conducted during the first science run (SR1) with a 60 day exposure using a fiducial mass of 5.5 tonnes. The results from LZ's profile likelihood ratio (PLR) analysis are presented in which the data is shown to be consistent with a background-only hypothesis, achieving the highest sensitivity to spin-independent WIMP-nucleon scattering for masses greater than $9\,\mathrm{GeV/c^2}$. The most stringent limit is set for spin-independent scattering at $36\,\mathrm{GeV/c^2}$, rejecting cross sections above $\sigma_{\mathrm{SI}} = 9.2 \times 10^{-48}\,\mathrm{cm^2}$ at the 90% confidence level.

Accidental coincidence events are also introduced in chapter 4. Understanding the source of these events, and developing methods to reject them constitutes the majority of my original contributions to research presented in this thesis. These include, studying and understanding the sources of the isolated S1 and isolated S2 events contributing to the rate of accidental events, and developing and evaluating cuts to target these events. These cuts are discussed in chapter 4 along with a set of data quality cuts that remove periods of data taking during which the TPC saw elevated rates of activity; and event vetoes that utilise the Skin and Outer Detector to enhance the rejection of neutron and gamma ray backgrounds.

However, with the official cuts used in the SR1 analysis and the accidental rate seen during SR1, the planned 1000 day exposure would lead to 6.00 ± 1.5 events in the region of interest. This is well beyond the ~ 1 event requirement outlined in LZ's Technical Design Report [10].

Chapter 5 presents a solution to this in the form of an original analysis that uses Boosted Decision Trees (BDTs) to reject accidental events in the LZ data. In this chapter, the methodology developed to apply this machine learning (ML) technique to the LZ data is outlined. This includes: the selection of appropriate training data; steps to ensure bias mitigation in training of and classification by the BDT through the use of a 2D weighting; and the use of a threshold function to make classification with the BDT scores. By utilising the analysis developed, the expected background over a 1000 day run is reduced to 1.45 ± 0.8 events which meets the requirements for LZ to reach its experimental goals. This is achieved at the cost of a slight loss in signal efficiency, which is reduced from $92.3 \pm 0.2\%$ using the published SR1 analysis to $73.0 \pm 0.2\%$ using the BDT analysis.

Chapter 2

The Dark Matter Problem

This chapter introduces the theoretical motivations and astrophysical evidence for the existence of dark matter at the galaxy, galaxy cluster and cosmological scales. A collection of viable dark matter candidates that meet the constraints set by the astrophyiscal evidence are discussed with a focus on WIMPs and WIMP direct detection, including an overview of the current experimental landscape. The chapter concludes by describing the operating principles of a dual-phase liquid xenon TPCs.

2.1 Astrophysical Evidence

2.1.1 Galaxy Rotation Curves

An early yet compelling source of evidence for the existence of dark matter is the flatness of galaxy rotation curves at large radius. Based on standard Newtonian gravity the rotational velocity of spiral galaxies (e.g. the Milky Way) as a function of radius is expected to follow:

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

where G is the gravitational constant and $M(r)$ is the mass within radius r. If only the visible stars and gas were to provide all the gravitational mass, then one would expect the rotational velocity to fall as $\frac{1}{\sqrt{r}}$ beyond the galaxy's luminous centre where most of the visible matter is contained. Instead, a flat rotational curve is "observed", as shown in Figure 2.1. This indicates that $M(r) \propto r$ beyond the central

luminous region and implies a density distribution of the form $\rho \propto r^{-2}$. This mass is assumed to be distributed in a spherical non-luminous 'dark matter halo' around the galaxy. This was first conclusively observed by Vera Rubin and her collaborators in the 1970s [17]. The rotation curves also provide a way to estimate the local dark matter density, a crucial input for dark matter direct detection experiments as discussed later in this chapter.

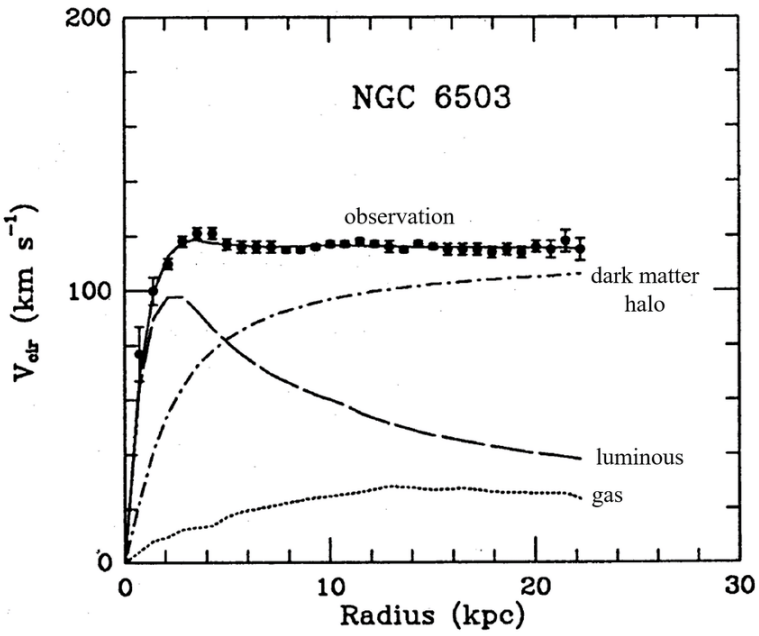


Figure 2.1: Measured rotation curve for NGC 6503 with contributions from the dark matter halo, disk and gas also shown [1].

2.1.2 Gravitational Lensing

Gravitational lensing experiments provide another source of evidence for dark matter. The bending of light from distant bright sources by massive objects in the foreground can be used to infer the gravitational potential (or mass) of those foreground objects. This phenomenon of graviational lensing is shown in the image in Figure 2.2. Discrepancies in gravitational mass of galaxies measured through lensing compared with the expected mass of the luminous part of the galaxy indicate the existence of additional non-luminous matter [18].

Figure 2.2: An image from the Hubble Space Telescope showing several blue, loop-shaped objects that have been duplicated by the strong gravitational lens of the galaxies (called 0024+1654) near the photograph's centre shown in yellow [2].

2.1.3 Galaxy Clusters

Two types of observations provide evidence for dark matter at the scale of the galactic clusters.

First is the discrepancy between the observed luminous mass and the mass inferred from the observed velocity of the objects in the system. This was reported in the 1930s by Fritz Zwicky who related the velocity dispersion in the Coma cluster to the observed mass and radius using the virial theorem. This suggested that the stellar mass alone could not account for the gravitational potential of the system [19]. This discrepancy also appears in later observations of X-ray emission from gas associated with galaxy clusters [20]. Again, using the virial theorem, the energy of the hot gas (T $\sim$ keV) is related to the gravitational potential of the galaxy cluster, and therefore the mass of the cluster. The mass discrepancies in these observations are dealt with by positing the existence of a dark matter halo that, along with the luminous stellar matter, contributes to the total mass of the galaxy clusters.

The second piece of evidence is provided by observations of galaxy cluster col-

lisions. The most cited of these being observations of the *Bullet Cluster*, which emerged as the result of a sub-cluster (the "bullet") colliding with the larger galaxy cluster 1E 0657-56 [21]. As shown in Figure 2.3, the total gravitational mass distribution in the bullet cluster is mapped through gravitational lensing (shown in blue) and the baryonic mass distribution is determined through X-ray emissions from the interstellar gas (shown in red). Comparing the location of the baryonic matter with the distribution of the gravitational mass as determined through lensing shows that the majority of the mass in the cluster is dark non-baryonic matter. Due to self interaction, the baryonic matter slows down, mixes and remains close to the interaction point. Whereas the dark matter in each cluster passes through thus indicating that it has a very small cross section both with itself and with baryonic matter [22].

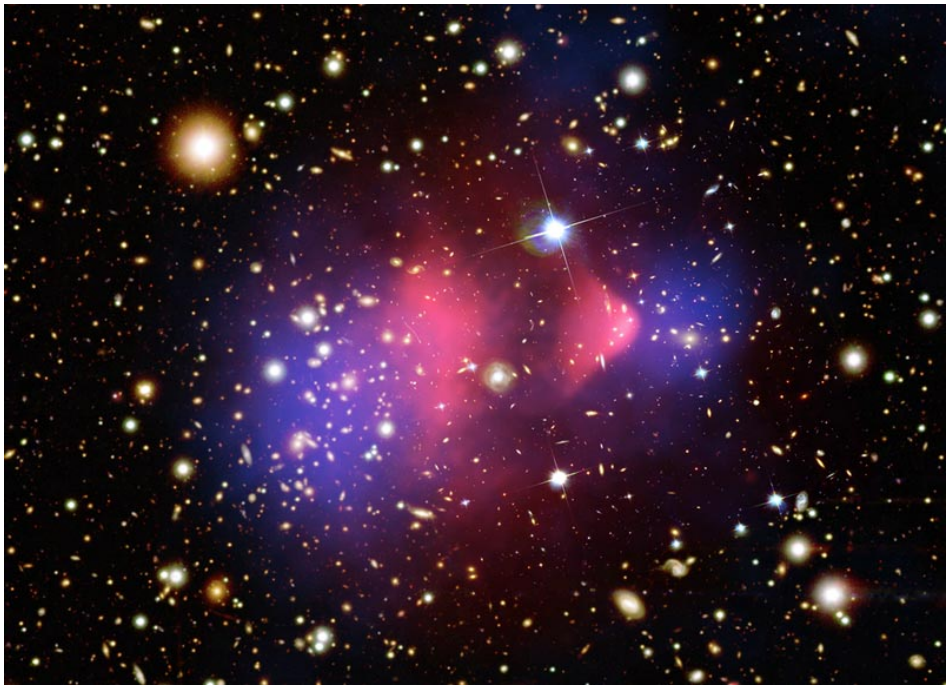


Figure 2.3: A composite image of the *Bullet Cluster* showing two colliding galaxies. The hot gas seen in X-rays by Chandra is shown in the two pink clumps and contains most of the "normal" matter in the two clusters. The blue areas in this image show where most of the mass in the two clusters is found through gravitational lensing. Optical images from Magellan and Hubble show the galaxies in orange and white [3]. Credits: X-ray: NASA/CXC/M.Markevitch et al.; Optical: NASA/STScI; Magellan/U.Arizona/D.Clowe et al.; Lensing Map: NASA/STScI; ESO WFI; Magellan/U.Arizona/D.Clowe et al. Available at Ref. [3].

2.1.4 Cosmological Evidence

Further evidence is found on the cosmological scale in the form of anisotropies in the cosmic microwave background (CMB). The CMB is an almost isotropic source of microwave photons matching a blackbody spectrum with temperature 2.72 K. It provides crucial evidence favouring the Big Bang model of an expanding universe over a Steady State universe, as it predicts a very hot and dense period in the evolution of the universe during which baryonic matter was completely ionised and the universe appeared as an opaque (due to Thompson scattering of photons) blackbody radiation emitting entity. As the universe cooled, approximately 380,000 years after the Big Bang, electrons began to recombine with free protons and form neutral hydrogen and helium atoms, thus decoupling the CMB photons from the opaque plasma. Since then, these photons have redshifted in an expanding universe and now appear to us as microwave photons. The CMB has been mapped in detail first by Wilkinson Microwave Anisotropy Probe (launched in 2001) [23] and later by the Planck satellite (launched in 2009) [5]. The latest CMB measurements from Planck are shown in Figure 2.4.

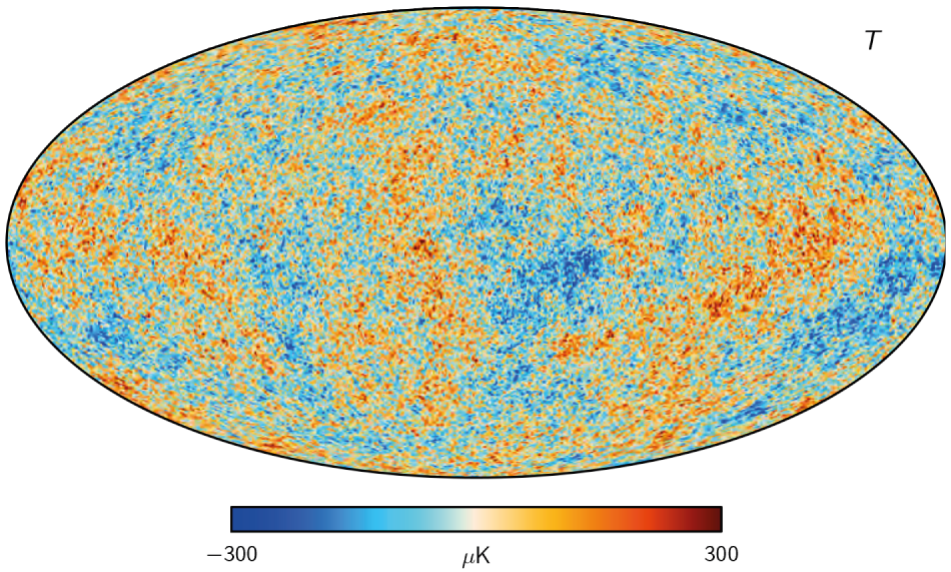


Figure 2.4: A full sky map of the anisotropies in the CMB temperature with respect to the mean temperature as observed by the Planck Space Telescope [4].

Despite being near perfectly isotropic, small anisotropies in temperature (as shown in Figure 2.4) are present within the CMB map. These anisotropies are

mapped at different angular scales and parameterised as the multipole moment ℓ of a spherical harmonic (see Figure 2.5). The CMB spectrum anisotropies are well described by the Lambda (dark energy) Cold Dark Matter model (ΛCDM), which has emerged as "Standard" moodel of cosmology and requires the existence of non-baryonic, cold (non relativistic) matter. Theoretical predictions of the power spectrum from ΛCDM depend on key cosmological parameters, crucially the matter density and the baryon density. Results from fits to the Planck data measure the baryon density as $\Omega_b = 0.0024 \pm 0.0001$ and the dark matter density as $\Omega_c = 0.120 \pm 0.001$ [5].

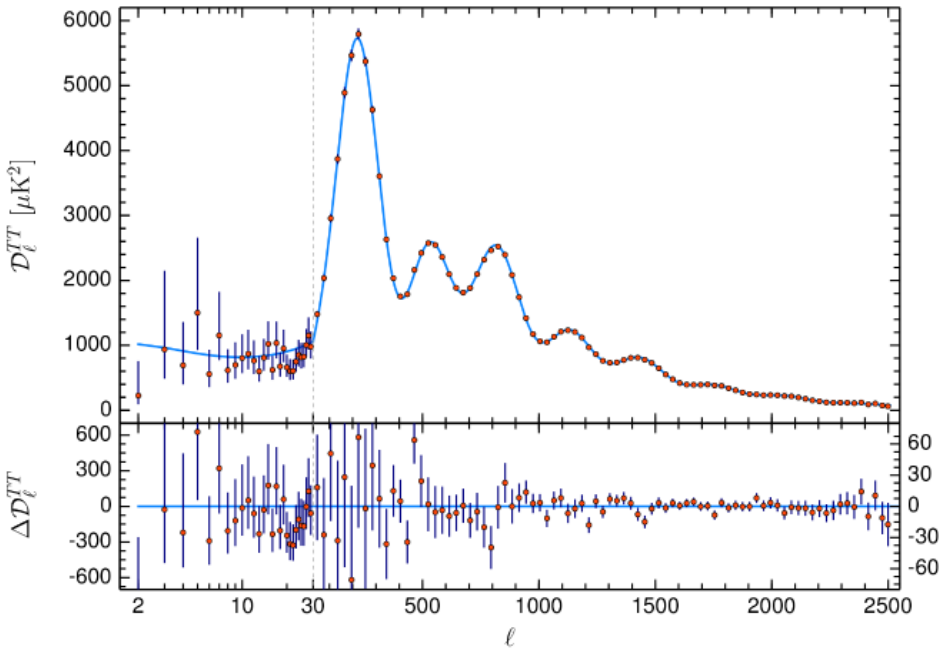


Figure 2.5: The top panel shows the Planck 2018 temperature power spectrum with respect to the spherical harmonic multipole parameter and a best fit assuming ΛCDM cosmology. The lower panel shows the residuals between the ΛCDM fit and the observed data [5].

The anisotropies are understood to be caused by baryon acoustic oscillations. These are density fluctuations that emerged during the early universe, before decoupling, as a result of the emergence of (gravitational) potential "wells" that baryonic and dark matter collapse into. As the photon-baryon fluid density increased in the collapse, the outward radiation pressure from the photons eventually overcame gravity leading to expansion. The expanding fluid then cooled until the gravita-

tion attraction was once again greater than the thermal pressure. These oscillations continued until recombination, after which the photons escaped leaving behind unperturbed dark matter centred at the over-density and surrounded by a shell of baryons which subsequently fell back inwards. This interaction between the oscillating baryons and unperturbed matter have implications for galaxy and large scale structure formations and have been detected in the clustering of galaxies [24].

2.2 Dark Matter Candidates

Based on the cosmological and astrophysical constraints discussed in the previous sections we can posit a set of key properties for dark matter. These are:

- A dark matter particle must be charge neutral (or millicharged) with the most stringent constraints stemming from the requirement that dark matter be fully decoupled from the baryon-photon plasma at recombination [25].
- It is stable or very long-lived (i.e. greater than the age of the universe) and must be produced through a mechanism that reproduces the appropriate value for the relic density [26].
- It is expected to be "cold" i.e. non-relativistic. This constraint comes from strong agreement between N-body simulations of the (ΛCDM) model and a wide range of observed structures [27, 28].
- It is massive and interacts with ordinary matter primarily through its gravitational influence.

2.2.1 WIMP Dark Matter

WIMPs are a collection of possible new, theoretically well motivated particles which could constitute dark matter. These are stable, massive particles that interact with ordinary matter via gravity and an additional force with an interaction rate that is at least as weak as the Weak force. In the ΛCDM model, it is assumed that during the early "hot soup" stage of the universe, all particles (including dark matter) have enough energy to remain in thermal equilibrium through annihilation and pair

production. However, as the universe expanded and therefore cooled, the reaction rates fell below the level required for thermal equilibrium. Particles remaining after this "freeze-out" are what make up the relic abundance observed today [29].

The evolution of the number density of dark matter particles n_χ is modelled using the Boltzmann equation [30]:

$$\frac{dn_\chi}{dt} + Hn_\chi = -\langle\sigma v\rangle(n_\chi^2 - n_{eq}^2) \tag{2.2}$$

where H is the Hubble constant, $\langle\sigma v\rangle$ is the thermally averaged cross section, and n_{eq}^2 is the equilibrium particle number density. For non-relativistic dark matter

$$n_\chi \sim (m_\chi T_f)^{3/2} e^{-m_\chi/T_f} \sim \frac{T_f^2}{M_{pl}\langle\sigma v\rangle} \tag{2.3}$$

where T_f is the freeze-out temperature and the planck mass $M_{pl} = \sqrt{\hbar c/G}$ [31]. Intuitively one can see that the annihilation rate is driven by the ratio $\frac{m_\chi}{T_f}$.

The relic dark matter abundance Ω_χ is then estimated as:

$$\Omega_\chi = \frac{m_\chi n_\chi^0}{\rho_c} \sim \frac{0.1}{h^2}\left(\frac{3\times10^{-26}\text{cm}^3/\text{s}}{\langle\sigma v\rangle}\right) \tag{2.4}$$

where n_χ^0 is the dark matter number density at the present time. One can then infer that a new particle with a cross section roughly similar to that of the electroweak scale reproduces the observed dark matter abundance $\Omega_\chi h^2 \sim 0.1$. This has been dubbed the "WIMP Miracle" and provides motivation for the WIMP paradigm. Furthermore, several WIMP candidates naturally arise in popular extensions to the Standard Model, such as supersymmetry [31].

Dark Matter Halo

The distribution or shape of the dark matter Halo surrounding spiral galaxies is generally assumed to have a cuspy (density increasing rapidly at small radii) two-

parameter profile given by:

$$\rho_{NFW} = \frac{\rho_c \delta_c}{\frac{r}{r_s}\left(1+\frac{r}{r_s}\right)^2} \tag{2.5}$$

where ρ_c is the critical density of the universe, r_s is a scale radius for isothermal density profiles, and δ_c is a dimensionless density parameter. This form of the density profile is known as the Navarro-Frenk-White (NFW) halo and was predicted through N-Body simulations of structure formation in an expanding universe [32, 33]. The form of this profile is seen to be universal for variety of halo masses, and spans four orders of magnitude from small galaxies to galaxy clusters.

More recent higher resolution simulations using the Einasto profile [34] have provided a better fit to the density profiles of dwarfs galaxies and galaxy clusters. The model is given by:

$$\rho = \rho_e \exp -d_n \left[\left(\frac{r}{r_e}\right)^{1/n} - 1\right] \tag{2.6}$$

where n is a free parameter that measures the "shape" of a galaxy's luminosity profile related to how centrally concentrated a galaxy's light profile is [34]. The term d_n is then a function of n such that ρ_e is the density at radius r_e which defines the radius containing half of the total mass. For a full discussion on dark empirical models for dark matter halos see Ref. [34].

Whilst both these profiles provide good fits to N-body simulations of galaxy structure formation, they still lead to the cuspy halo problem [35]. This is the discrepancy between observations of the dark matter density in the inner parts of galaxies derived from galaxy rotation curves and from N-Body simulations. The former indicate an approximately constant dark matter density, whereas the latter result in a steeply rising power-law-like behaviour. This core-cusp problem is one of the currently unresolved issues in small-scale cosmology [35].

2.2.2 Other Dark Matter Candidates

Neutrinos

Neutrinos have historically been leading candidates for dark matter, partly due to their indisputable existence, as well as the fact that they are weakly interacting and massive. However, Standard Model neutrinos are largely ruled out as an explanation for a significant portion of the dark matter in the universe since they are relativistic and do not match the appropriate relic density [36]. On the other, hand sterile neutrinos arising in extensions to the Standard Model could contribute to the warm dark matter in the universe, as discussed in [37].

Axions

The axion is a pseudo-Nambu-Goldstone boson proposed by Weinberg [38] and Wilczek [39] that arises via the Peccei–Quinn mechanism [40] as a solution to the strong CP problem which arises from the lack of observable Charge-Parity (CP) violation in the theory of strong interactions quantum chromodynamics (QCD). Astrophysical constraints leave only very light axions as viable dark matter candidates with an axion mass 20 μeV $\lesssim m_A \lesssim$ 1 meV [41, 42]. In contrast to WIMPs, axions do not naturally produce the correct relic density and thermal production cannot account for the observed bulk of dark matter. However several other, non-thermal production mechanisms have been proposed and are discussed in [43].

SuperWIMPs

Another class of particles that are a viable dark matter candidate are the super-Weakly Interacting Massive Particles (superWIMPs). In this framework, present dark matter is not required to interact via the weak force to still take advantage of the WIMP miracle and achieve the appropriate relic density. The postulated super-WIMP dark matter particles are the decay products of thermally produced WIMPs that freeze out as usual [44]. Therefore they inherit the appropriate relic density from WIMPs, and for $m_{SWIMP} \sim m_{WIMP}$ WIMPs decay into superWIMPs with the desired amount for them to account for a significant fractionn of dark matter. Signals for superWIMPs could be found in ultra-high energy cosmic rays, at colliders

and in indirect astrophysics searches. A popular candidate for superWIMPs is the gravitino, the spin 3/2 supersymmetric partner of the graviton. Another set of superWIMP candidates are axinos which are the supersymmetric partners of axions. These are discussed in Ref. [44].

2.3 WIMP Searches

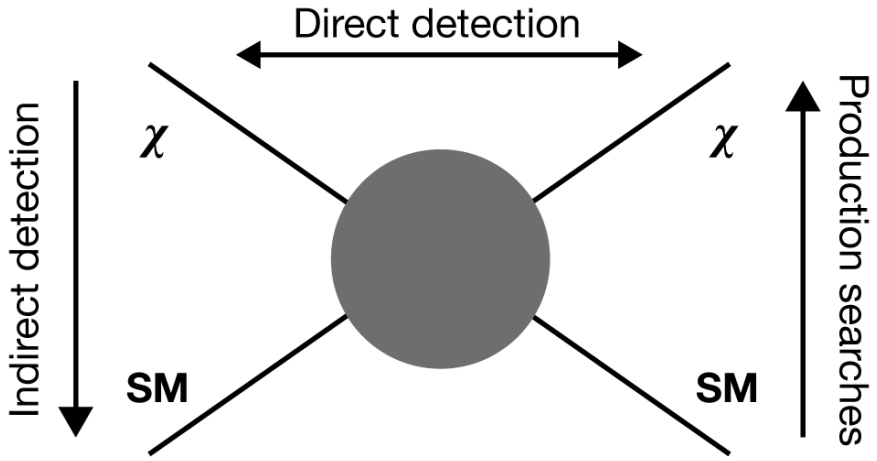


Figure 2.6: A cartoon Feynman Diagram illustrating the channels through which a dark matter search can be conducted.

The simplified Feynman diagram in Figure 2.6 illustrates three types of dark matter interactions and the corresponding three channels for dark matter searches. Firstly, by taking the arrow of time in the bottom-to-top direction we see a possible route to dark matter detection through its production in the collision of Standard Model particles at, for instance, the Large Hadron Collider (LHC) using a detector like ATLAS or CMS. Like neutrinos, dark matter will not interact with any of the subsystems in these detectors and events are instead reconstructed by looking for "missing" transverse momentum (E_T^{Miss}) in a collision. An example of such a dark matter producing process is shown in Figure 2.7, which involves the associated production of a Z-boson with a Higgs that eventually decays into invisible particles ($H \rightarrow$ inv) [6].

Secondly, dark matter particles can annihilate into Standard Model particles, as shown in Figure 2.6 with the arrow of time running from top to bottom. Astrophysical searches for such dark matter self-annihilation and decay processes through the

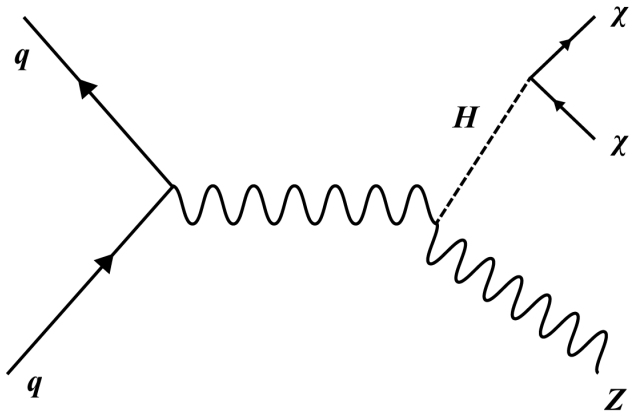


Figure 2.7: Feynman diagram showing the associated production of a Higgs boson and a Z boson, where the Higgs decays to dark matter particles [6].

observations of photons, charged cosmic rays and neutrinos constitute the indirect detection efforts. Experiments typically search for signals in regions of the universe with an expected overdensity of dark matter. Such regions include the galactic centre, nearby dwarf galaxies and galaxy clusters. A full review of indirect detection methods including an overview of recent results can be found in Ref. [45].

Lastly, with the arrow of time running in the horizontal direction in Figure 2.6, we posit a scattering interaction between a dark matter particle and a Standard Model particle. Searches for signals induced by such interactions using terrestrial detectors are categorised as *direct detection* experiments. Since the cross section for dark matter-SM scattering is very small direct detection experiments are searching for extremely rare events. As such, they require ultra-low background levels in the energy range of interest and are typically conducted in deep underground laboratories to shield against cosmic rays.

Direct detection searches for axion dark matter and axion like particles (ALP) are primarily based on their coupling to photons $g_{a\gamma}$, with low mass ($\lesssim 1$ μeV) searches also utilising the coupling to electrons (g_{ae}) and or nuclei (g_{aN}). The Axion Dark Matter Experiment (ADMX) collaboration uses the expected axion-photon coupling to search for cold dark matter from the local dark matter halo using a cold resonant microwave cavity immersed in a magnetic field [46]. Light axions and ALPs can also generate electron recoil (ER) signals via the axio-electric effect [47]

and searches for such interactions are currently ongoing [48–50]. A full discussion on Axion searches can be found in Ref. [51].

WIMP Direct detection which is the principle focus of the LZ experiment is discussed in detail in the next section and follows Refs. [7, 9, 31, 51–53].

2.3.1 WIMP Direct Detection

Direct detection experiments searching for WIMPs primarily focus on elastic scattering between a WIMP from the galactic DM halo and a target nucleus. Momentum transfer between a WIMP and nucleus leads to a nuclear recoil (NR) in the detector. The expected interaction of this NR is non-relativistic and can therefore be modelled as a two-body elastic scatter with a differential event rate per unit target mass is given by:

$$\frac{dR}{dE_R} = \frac{\rho_\chi}{m_\chi m_N} \int_{v_{min}}^{v_{esc}} v f(v) \frac{d\sigma}{dE_R} dv \tag{2.7}$$

where E_R is the recoil energy, ρ_χ is the local dark matter density, m_N and m_χ are the mass of the target nucleus and the incoming WIMP respectively, σ is the WIMP-nucleus scattering cross section, v is the WIMP velocity and $f(v)$ is the probability distribution function of v in the WIMP halo. The upper limit of the integral is the WIMP escape velocity $v_{esc} = 544$ km s^{-1}, i.e. the velocity above which WIMPs with would no longer be bound by the gravitational potential of the Milky Way. The lower bound, v_{min}, is the minimum velocity required to produce a measurable nuclear recoil and depends on the energy threshold (E_{thr}) of the detector to detect a nuclear recoil. Based on the simple kinematics of two-body scattering, v_{min}, is given by:

$$v_{min} = \sqrt{\frac{E_{thr} m_N}{2} \frac{(m_N + m_\chi)^2}{(m_N m_\chi)^2}} = \sqrt{\frac{E_{thr} m_N}{2} \frac{1}{\mu^2}} \tag{2.8}$$

where μ is the reduced mass. The number of events measured in a detector over live time T is then given by integrating equation 2.9 over the measurable energy range:

$$N = T \frac{\rho_\chi}{m_\chi m_N} \int_{E_{thr}}^{E_{max}} \int_{v_{min}}^{v_{esc}} \varepsilon(E_R) v f(v) \frac{d\sigma}{dE_R} dv\, dE_R \tag{2.9}$$

Where $\varepsilon(E_R)$ is the detector efficiency as a function of recoil energy, and the upper limit of the energy range integral is the maximum recoil energy $E_{max} = 2\mu^2 v_{esc}^2/m_N$ where the incoming particle fully back scatters. The differential event rate (2.9) of WIMP nuclear recoils is expected to be a featureless and smoothly decreasing exponential of the form:

$$\frac{dR}{dE_R} \propto \exp\left(\frac{-E_{nr}}{E_0}\frac{4m_\chi m_N}{(m_\chi + m_N)^2}\right) \tag{2.10}$$

where E_0 is the expectation value of the kinetic energy distribution of the incident WIMP. Intuitively, the exponential form tells us that an experiment's low energy threshold is a significant driver for its overall WIMP sensitivity.

WIMP-nucleus scattering occurs through two possible interactions: a spin-dependent axial-vector interaction and a spin-independent scalar interaction. The differential cross section in equation 2.9 can be expressed as the sum of a spin-independent (SI) and a spin-dependent (SD) component [51]:

$$\frac{d\sigma(E_{nr})}{dE_{nr}} = \frac{m_N}{2v^2\mu^2}\left[\sigma_{SI}F_{SI}^2(E_{nr}) + \sigma_{SD}F_{SD}^2(E_{nr})\right] \tag{2.11}$$

where the nuclear form factors F_{SI} and F_{SD} account for loss of coherence at high momentum transfers, where the nucleus the structure of nucleons within the nucleus must be taken into account. The spin independent cross section σ_{SI} is given by:

$$\sigma_{SI} = \sigma_n \frac{\mu^2}{\mu_n^2}\frac{(f_p Z + f_n(A-Z))^2}{f_n^2} \tag{2.12}$$

where Z and A are the number of protons and the atomic mass of the target nuclei respectively, μ and μ_n are reduced masses of the WIMP-nucleus system and the WIMP-nucleon system respectively, σ_n is the WIMP-nucleon cross section, and f_p and f_n are the coupling constants to protons and neutrons. By assuming $f_p = f_n$ the above expression can be simplified as:

$$\sigma_{SI} = \sigma_n \frac{\mu^2}{\mu_n^2}A^2 \tag{2.13}$$

which shows an A^2 dependence, thus favouring heavy target nuclei to search for spin-independent interactions. For spin-dependent interactions the spin structure of the nucleus must be taken into account and the cross section is described by:

$$\frac{d\sigma_{SD}}{d|\mathbf{q}|^2} = \frac{8G_F^2}{\pi v^2}\left[a_p\langle S_p\rangle + a_n\langle S_p n\rangle\right]\frac{J+1}{J}\frac{S(|\mathbf{q}|)}{S(0)} \tag{2.14}$$

where $\mathbf{q}$ is the momentum transfer, $\langle S_p\rangle$ and $\langle S_N\rangle$ are the expectation values of the total spin operators for protons and neutrons in the target nucleus [54]. $S(|\mathbf{q}|)$ is the spin-structure function and J is the total nuclear spin of the target. The spin-structure functions contain information about how spin is distributed between the nucleons. These have been modelled either using detailed numerical calculation or using the "odd-group" model, which assumes all the nuclear spin is carried by the "odd" group, i.e. either by all the protons or all the neutrons, whichever is most unpaired. This leads to only one of $\langle S_p\rangle$ and $\langle S_N\rangle$ being non-zero [31].

2.3.2 Detector Target Material

With the above nuclear recoil rate equations (2.9 & 2.10), one can compare the expected event rate for different detector targets materials, as shown in Figure 2.8 for the SI case. The A^2 dependence of the SI interaction rate (see equation 2.13) results in detectors using heavier targets like Xe or Ge expecting to see a higher number of scalar WIMP-nucleus interactions.

Another key consideration is ease with which radioactive impurities can be removed from the target material. These can be in the form of radioactive isotopes that produce an intrinsic radioactive background (discussed in more detail in Sec 3.5.2) which would scale with detector size.

2.3.3 Current State of the Field

There have been several experimental searches spanning decades for WIMPs using different target materials. Together these have probed a large range of WIMP masses and cross sections with sensitivity increasing over time as a result of detector masses and exposures have increasing. Figure 2.9 shows the evolution of the

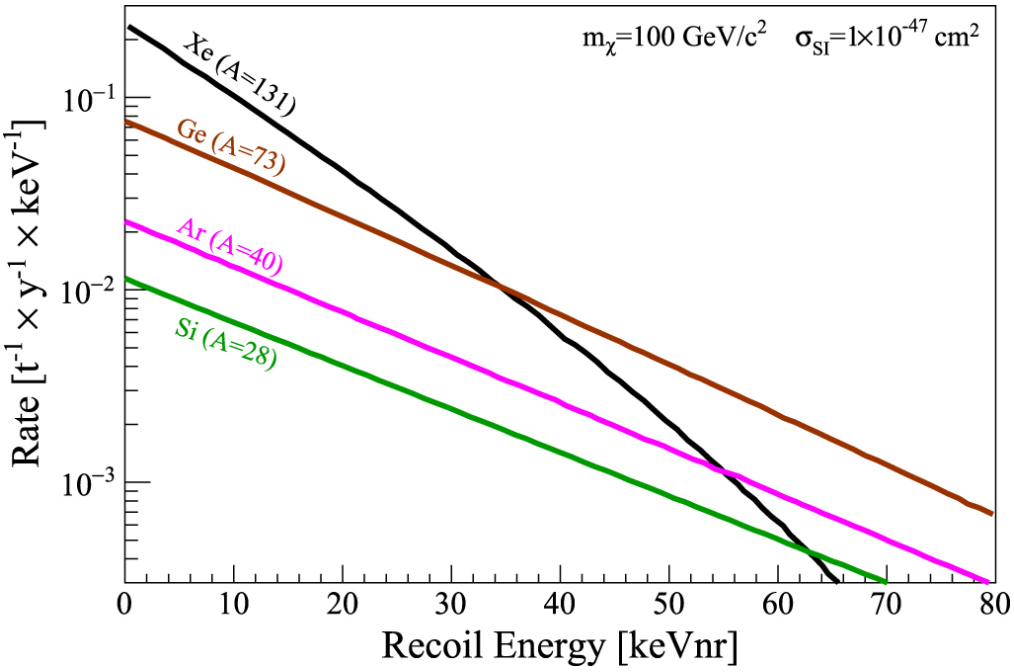


Figure 2.8: Nuclear recoil spectra for a $m_\chi = 100$ GeV$/c^2$ incident WIMP for typical nuclear recoil energies [7].

limit on the WIMP-nucleon spin-independent cross section for a 60 GeV WIMP.

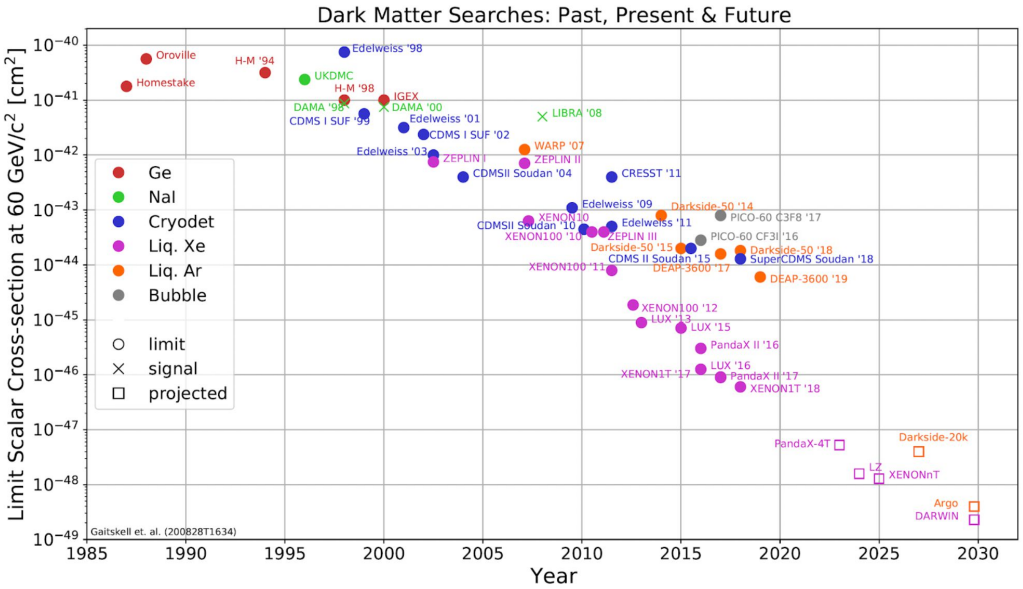


Figure 2.9: The evolution of the limits set on the spin-independent WIMP-nucleon cross section cross section for a 60 GeV WIMP. Figure courtesy of the LZ collaboration.

In recent years the most competitive limits have been set by experiments using xenon as the target material in dual-phase time projection chambers (TPCs). In

addition to the Xenon atom's high atomic mass contributing to high rate at low energy (as shown in Figure 2.8), the technology has scaled rapidly going from $\sim$ 10 kg of target Xenon in the late 2000s to $\sim$ 7000 kg in 2022. Xenon is not reactive which means it can be purified to a high degree and electronegative impurities can be removed without significant difficulty. This is crucial to the efficient functioning of TPCs, as discussed below. Moreover, Xenon is known to have only two naturally occurring long-lived isotopes - ^{136}Xe and ^{124}Xe - with well characterised decay spectra leading to well understood intrinsic backgrounds.

2.4 Signal Generation in Dual Phase LXe TPCs

This section describes the signal generation and detection in dual-phase liquid xenon (LXe) TPCs and follows the review in Ref. [55].

An illustration of a particle interaction in a liquid-noble TPC is shown in Figure 2.10. The bulk of the detector volume consists of the xenon target in liquid phase with a thin gaseous region at the top. This is viewed from the both the top and bottom with arrays of photosensors, usually photo-multiplier tubes (PMTs). Incoming particles interact with the target nuclei causing electron or nuclear recoils producing primary scintillation light (S1) that is measured by the photosensors, and ionisation electrons that are drifted to the top of the detector using an applied vertical electric field. These electrons are then extracted out of the liquid region and accelerated through the gas phase which produces delayed scintillation (S2) through electroluminescence [55]. S1 and S2 light production is discussed in more detail in sections 2.4.2 and 2.4.3 respectively.

With the S1 and S2 signal the 3D position of the interaction in the liquid xenon can be reconstructed. The XY position is obtained using the localised S2 light pattern in the top array. The vertical Z position is calculated using the time between the S1 and S2 signals, i.e. the time taken to drift electrons from the interaction site to the gas phase. Position reconstruction in LZ is discussed in the next chapter. Finally, the S1 and S2 signal can be used to reconstruct the energy of the recoil, this is discussed in section 2.4.4 below.

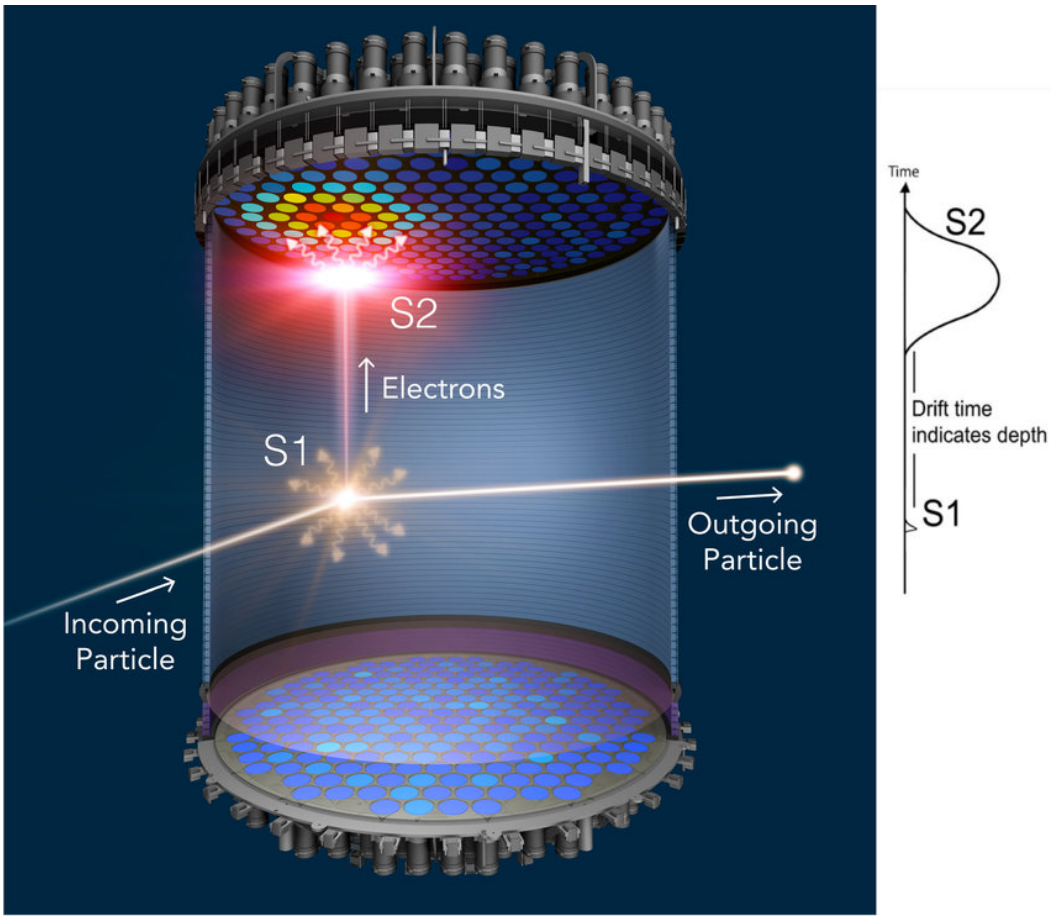


Figure 2.10: An illustration of a particle interaction of the dual-phase xenon TPC. An electric field is applied along the cylindrical part of the TPC to drift electrons towards the gas region. Incoming particle that interact in the active LXe produce prompt scintillation (S1), and a delayed electroluminescence signal (S2) produced by electrons extracted into the gas. Figure courtesy of the LZ collaboration.

2.4.1 Energy Deposition

The energy transferred in an interaction between an incoming particle and atoms in the target volume is split into three channels. For electrons, one can write:

$$E_0 = N_i E_i + N_{ex} E_{ex} + N_i \varepsilon \tag{2.15}$$

where N_i and N_{ex} are the mean number of ionised and excited atoms respectively. The first channel of energy transfer is ionisation with E_i representing the mean energy to ionise a a xenon atom. The second is excitation with E_{ex} the mean excitation energy of xenon atoms. And lastly, ε is the mean kinetic energy of ionised electrons

immediately after the last excitation or ionisation [55, 56]. These electrons transfer energy to the medium in the form of heat as they do not have enough energy to ionise or excite atoms. For LXe ε is taken to be 4.6 eV [55].

Liquefied noble gases exhibit a band structure of electronic states. This allows the energy quantities in Eq. 2.15 to be conveniently written in terms of the band gap, E_g, giving an energy transfer equation for electron interactions in liquid phase as:

$$\frac{E_0}{E_g} = N_i \frac{E_i}{E_g} + N_{ex} \frac{E_{ex}}{E_g} + N_i \frac{\varepsilon}{E_g} \tag{2.16}$$

This can be rewritten in terms of the energy transferred per ionisation, $W = E_o/N_i$, which gives:

$$\frac{W}{E_g} = \frac{E_i}{E_g} + \frac{E_{ex}}{E_g}\frac{N_{ex}}{N_i} + \frac{\varepsilon}{E_g} \tag{2.17}$$

The above formulation must be modified for nuclear recoils, in which a large fraction of the transferred energy is spent on the target atom recoiling and causing elastic collisions with other atoms rather than towards ionisation or excitation. This energy is therefore transferred as heat and lost. A first principles theory of energy loss of such recoiling ions in NR interactions has not been developed [55], however an approximate model known as Lindhard–Scharff–Schiøtt (LSS) theory has been developed and is reviewed in Ref. [57].

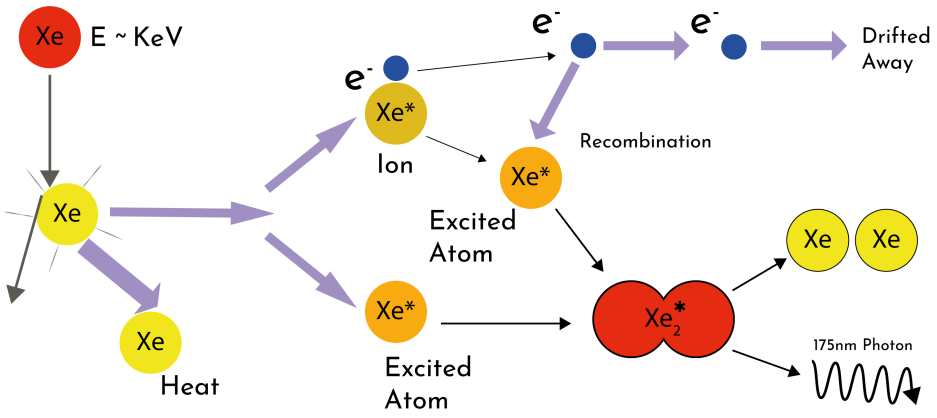


Figure 2.11: A simplified illustration showing the processes that create scintillation light and ionisation electrons from recoils in the liquid Xenon. High resolution recreation of the same figure in [8].

Nevertheless, equation 2.17 can still be instructive in developing an intuition for recoil signal generation in dual-phase TPCs. Figure 2.11 shows the process by which a recoil creates measurable S1 and S2 light. It shows that ionised atoms lead to scintillation light through a different route compared to excitation, with some of the ionisation electrons being drifted away. This leads to differences in the measured signal that depend on the ratio of excitations to ionisation, i.e. the second term $\frac{N_{ex}}{N_i}$ in equation 2.17. For electron recoils in liquid xenon this ratio is estimated as ~ 0.13 whereas for nuclear recoils $\frac{N_{ex}}{N_i} \sim 1$ [55, 58]. This difference allows for discrimination between ER and NR signals in dual-phase TPCs.

2.4.2 Primary Scintillation (S1) in LXe

Keeping the illustration in Figure 2.11 in mind, we see that a recoil produces atomic excitation and ionisation. Both of these processes lead to the production of S1 light. Excitation of the Xe atoms leads to the formation of strongly bound diatomic molecules in an excited state (excimer). The excimers are both vibrationally and electronically excited and subsequently undergo relaxation to the ground state. Vibrational relaxation is mostly non-radiative though can occasionally occur through the emission of infrared photons. Electronic de-excitation occurs through vacuum ultraviolet (VUV) emission. This process can be written as:

$$X + \mathrm{Xe} \rightarrow \mathrm{Xe}^* + X \qquad \text{Recoil and excitation} \qquad (2.18)$$

$$\mathrm{Xe}^* + \mathrm{Xe} \rightarrow \mathrm{Xe}_2^{*,\mathrm{v}} \qquad \text{Excimer Formation} \qquad (2.19)$$

$$\mathrm{Xe}_2^{*,\mathrm{v}} \rightarrow \mathrm{Xe}_2^* + X \qquad \text{Relaxation} \qquad (2.20)$$

$$\mathrm{Xe}_2^* \rightarrow \mathrm{Xe} + \mathrm{Xe} + \gamma \qquad \text{VUV Emission} \qquad (2.21)$$

where Xe* indicates electronic excitation and $\mathrm{Xe}_2^{*,\mathrm{v}}$ indicates both electronic and vibrational excitations.

Secondly, ionisation also leads to the emission of primary scintillation light through the process of recombination of the ionised electrons with positively charged diatomic Xe_2^+ molecules.

$X + \mathrm{Xe} \rightarrow \mathrm{Xe}^{+} + e^{-} + X$	Ionisation	(2.22)
$\mathrm{Xe}^{+} + \mathrm{Xe} + \mathrm{Xe} \rightarrow \mathrm{Xe}_2^{+} + \mathrm{Xe}$	Diatomic molecule formation	(2.23)
$e^{-} + \mathrm{Xe}_2^{+} \rightarrow \mathrm{Xe}^{**} + \mathrm{Xe}$	Excimer Formation	(2.24)
$\mathrm{Xe}^{**} + \mathrm{Xe} \rightarrow \mathrm{Xe}^{*} + \mathrm{Xe}$	Relaxation	(2.25)
$\mathrm{Xe}^{*} + \mathrm{Xe} \rightarrow \mathrm{Xe}_2^{*,\mathrm{v}}$	Excimer Formation	(2.26)
$\mathrm{Xe}_2^{*,\mathrm{v}} + \mathrm{Xe} \rightarrow \mathrm{Xe}_2^{*} + \mathrm{Xe}$	Relaxation	(2.27)
$\mathrm{Xe}_2^{*} \rightarrow \mathrm{Xe} + \mathrm{Xe} + \gamma$	VUV Emission	(2.28)

In both channels, VUV photons are ultimately emitted from one of the two lowest excited states of the Xe_2 molecule, the triplet and singlet state, as they de-excite to the ground state. The recombination process introduces a significant time delay in Xe as it is slower compared to the de-excitation times. This leads to the S1 pulse shape that can be parameterised by an exponentially decaying function, with three fundamental time constants: the singlet lifetime, the triplet lifetime, and the recombination time, as well as the ratio of singlet states to triplet states [59]. However, NR events have a higher ionisation density ($N_i \sim N_{ex}$ as discussed in the previous section) and rapid recombination, therefore the only parameter necessary to model the time evolution is the ratio of the singlet and triplet states. The singlet and triplet lifetimes are 3.1 ± 0.7 ns and 24 ± 1 ns respectively [59]. The recombination time is modelled using the Noble Element Simulation Technique (NEST) [59] for a given drift field, and is inversely proportional to the ionisation density i.e. the more spread out the charge distribution that longer it takes for them to recombine.

The probability of recombination depends on the strength of the applied electric field that drifts the ionised electrons towards the gas phase. The S1 pulse shape differences can also provide some discrimination between electron and nuclear recoil events, however for low energy collisions with only a few photons detected such analyses can be prohibitively challenging.

2.4.3 Secondary Scintillation (S2)

The free electrons that do not recombine are drifted towards the gas phase of the detector with an applied electric field and subsequently extracted into the gas phase and accelerated through it using a much higher field. This acceleration of electrons in the gas phase produces electroluminescence light (S2 light) that is then detected by the PMT arrays with observed number of photons being proportional to the number of free electrons.

A high field extraction region is required to transport electrons from the liquid to the gas as it is energetically favourable for them to remain in the liquid. This is achieved by creating an electric field between an electrode grid that sits just below the liquid surface and one that sits a short distance ($\mathcal{O}(1\text{ cm})$) above it as shown in Figure 2.12.

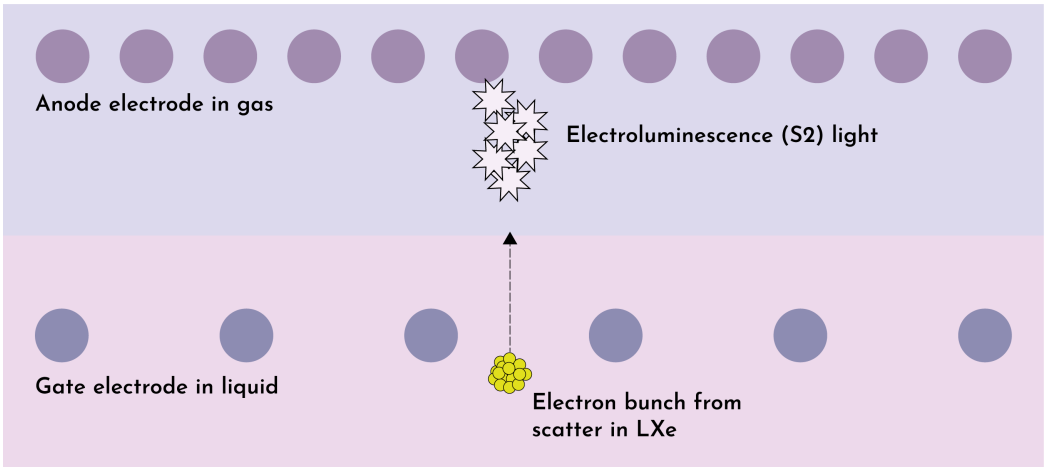


Figure 2.12: An illustration showing the process by which the electrons that are drifted away from the interaction site to the top of the detector are extracted into the gas phase by a strong electric field between the anode electrode in the gas and the gate electrode in the liquid. High resolution recreation of the same figure in [9].

The electron cloud will diffuse whilst being transported through the detector to the gas phase. The diffusion coefficient has a longitudinal D_L and a transverse term D_T. The electron diffusion can be used to measure drift time in low energy S2-only analyses where, for instance, the S1 has not been detected. This is done by using the relationship between drift time and the S2 pulse width, measured using calibration data.

The S2 signal is orders of magnitude larger than the S1 signal with events near the top of the detector producing a square pulse, and events closer to the bottom producing a more Gaussian shape due to diffusion of the electron cloud.

To minimise the loss of electrons during transport to the liquid surface the liquid xenon must be purified to remove electronegative impurities, such as oxygen and water that can capture electrons. Loss of these electrons is corrected for in the analysis by measuring the electron lifetime using calibration sources [10].

2.4.4 Energy Reconstruction

The energy of an interaction can be reconstructed with the S1 and S2 signal as follows:

$$E = W(N_i + N_{ex}) \cdot 1/f_n \tag{2.29}$$

where W is the energy required to produce a scintillation photon and for LXe is measured to be 13.7 eV [11], f_n accounts for the fractions of energy transferred to measurable electronic excitations, the rest is lost to heat [58].

Equation 2.29 can be written in terms of the number of measured electron (n_e) and the number of measured photons (n_γ), and subsequently in terms of S1 and S2. With r as the probability of recombination one can write:

$$n_e = N_i(1-r) \tag{2.30}$$

$$n_\gamma = N_{ex} + N_i r \tag{2.31}$$

Which can then be substituted into Eq. 2.29 to give:

$$E = W(n_e + n_\gamma) \cdot 1/f_n = W\left(\frac{S1}{g1} + \frac{S2}{g2}\right) \cdot 1/f_n \tag{2.32}$$

where g1 and g2 are the S1 light and charge collection efficiencies respectively which can be measured using calibration data. This equation then allows one to get a reconstructed recoil energy from a measured S1 and S2 in a Dual Phase LXe TPC.

Chapter 3

The LZ Experiment

An overview of the LZ Detector is shown in Figure 3.1. The experiment is located 4850 feet underground at the Sanford Underground Research (SURF) Laboratory at Homestake mine in Lead, South Dakota. At the heart of the LZ experiment is the Xenon detector comprising of the LXe Time Projection Chamber (TPC) and its surrounding Xe Skin detector. The Xenon detector is housed in an inner cryostat vessel (ICV) which is suspended inside an outer cryostat vessel (OCV), both of which are fabricated from low radioactivity titanium [60]. The OCV is surrounded on all sides by a gadolinium-loaded liquid scintillator (GdLS) Outer Detector (OD). The OD is viewed with 8" PMTs that stand in the Davis Campus water tank, which shields from neutrons and cavern gammas [11]. The LZ TPC was assembled and integrated into the ICV above ground in a radon reduced cleanroom at SURF. This chapter discusses the design and operation of these components, the LZ calibration programme and the dominant backgrounds affecting the WIMP search.

3.1 LZ Time Projection Chamber

An illustration of the LZ TPC is shown in Figure 3.2. It is comprised of Xe in dual phase (liquid-gas), two PMT arrays viewing the target material from the top and the bottom, the electrode grids, and the enclosing PTFE covered field cage. The bulk of Xe in the LZ TPC is in the liquid phase with a thin layer (8mm) of gaseous Xe at the top. As discussed in the previous chapter, interactions in the liquid phase lead to prompt scintillation light (S1 signal) and ionisation electrons are extracted to the

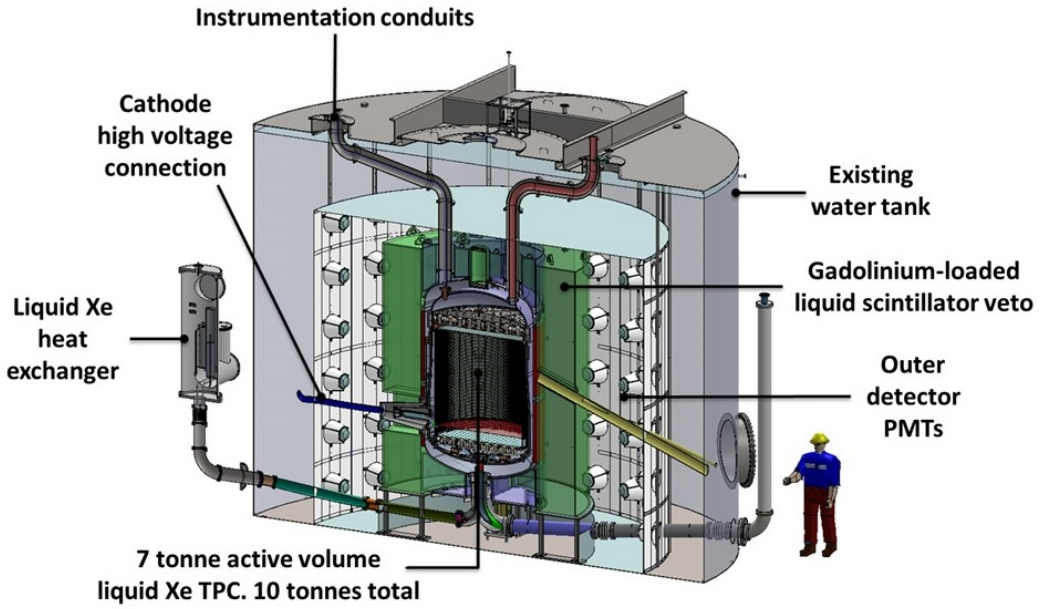


Figure 3.1: A cross section of the LZ detector highlighting its various subsystems [10].

gas phase where they produce electroluminescence light i.e. a delayed S2 signal. The TPC is designed to optimise the detection of light and charge produced from interactions in the 7 tonnes of active Xe and to suppress backgrounds through self shielding. Its key features are discussed below.

3.1.1 PMT Arrays (Detection of Light)

Primary xenon scintillation (S1 light) and electroluminescence photons (S2 light) are emitted in the vacuum ultraviolet (VUV) range (ultraviolet light < 200 nm). The S1 light collection efficiency is maximised primarily through the use of high quantum efficiency (QE) PMTs (discussed below) optimised for the detection of VUV photons, and the use of materials with high VUV reflectivity (e.g. PTFE walls) to minimise photon extinction. The S2 signal on the other hand consists of enough photons even at the lowest energies due to the gain from electroluminescence (measured at 47.1 ± 1.1 photons detected per electron [12]), and the main design driver of the top array is to optimise the spatial resolution for accurate position reconstruction.

The LZ PMT arrays are shown in Figure 3.3. Most S1 light produced in the bulk of the detector is collected by the bottom array. This is primarily because

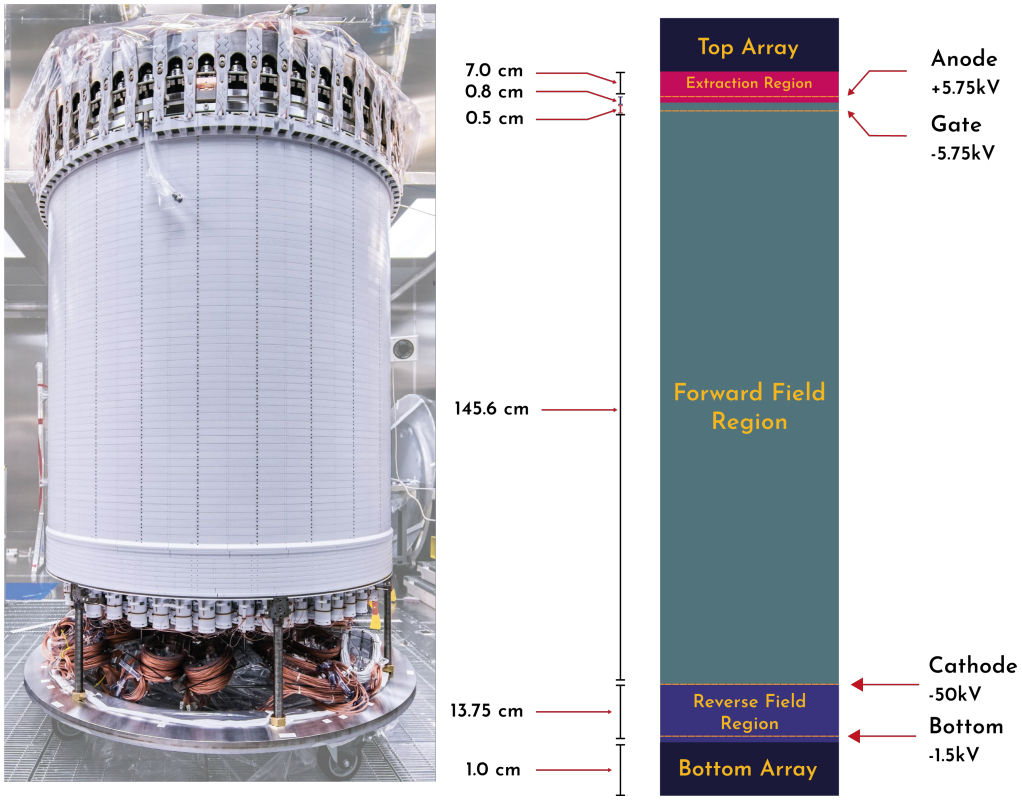
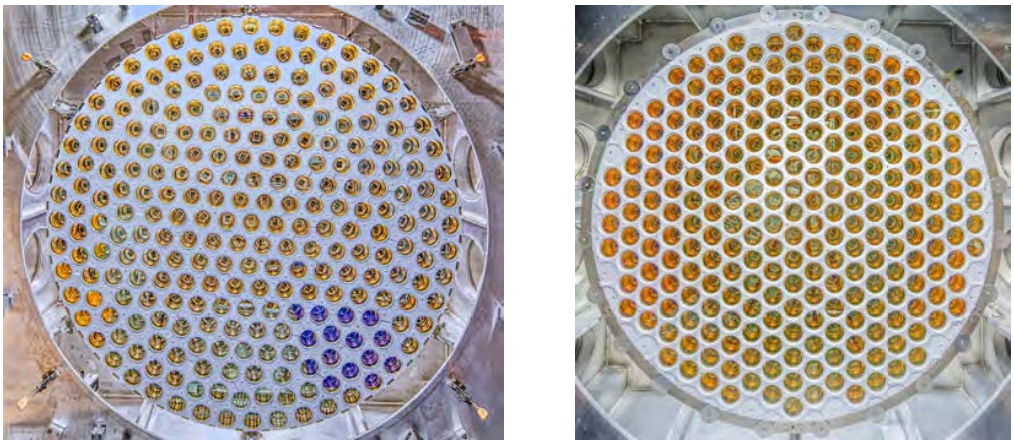


Figure 3.2: An overview of the LZ TPC. On the left is a photograph of the finished LZ TPC as built in the radon reduced cleanroom in the surface assembly lab at SURF. On the right is an almost to scale graphic illustrating the different field regions of the TPC and the distances between them. Photograph taken by Matthew Kapust, Sanford Underground Research Facility.



(a) **(b)**

Figure 3.3: The LZ PMT arrays. (a) An image of the TPC Top PMT array which shows the hexagonal structure in the centre moving to a circular arrangement on the outside; (b) An image of the TPC Bottom PMTs arranged in a closely packed hexagonal array. Photos taken by Matthew Kapust, Sanford Underground Research.

of reflection at the liquid surface, and refractive index matching between the PMT window and LXe in the bottom array leading to near perfect photon transmission at the boundary between the PMT window and LXe. The bottom array is designed to maximise photocathode coverage and hence consists of 241 3" PMTs close-packed in a hexagonal array.

On the other hand the top array plays the critical role of reconstructing the x, y position of an interaction from the S2 signal, especially of background interactions near the wall. These "wall events" are associated with the decays of radon progeny plated out on the TPC walls that, due to energy and charge loss at the wall, appear as a broad distribution of low-energy signals and overlap in part with the NR signal region in S1-S2 space. To minimise inaccuracy in the reconstruction of the x, y position of wall events the top array is constructed with a "hybrid" layout consisting of two nearly circular rows of PMTs at the perimeter and transitions to a hexagonal pattern in the centre. The spaces between PMTs on both arrays are fully covered with highly reflective PTFE.

Figure 3.4: A photograph of the Hamamatsu R11410-22 3" PMTs used in the LZ TPC taken by Matthew Kapust, Sanford Underground Research.

The 3" PMTs (model R11410-22) used in both the top and bottom array to view

the active region of the LZ TPC were specifically designed with Hamamatsu for operation at low temperature, high QE in the VUV range, high radiopurity, and low rates of spurious light emission (dark counts). These are shown in Figure 3.4. The average cold QE for the LZ TPC PMTs was measured to be 30.9% [61]. For a signal to be generated in a PMT, an incident photon must be transmitted through the PMT window and strike the photocathode producing free electrons through the photoelectric effect. These photoelectrons are then accelerated onto the first dynode where they are multiplied by means of secondary electron emission. This cycle is repeated at each successive dynode until the electrons are collected by the anode which outputs the electron current to an external circuit [62]. The nominal operating gain of the TPC PMTs is 3.5×10^6 measured at the end of the signal cables [11], and the typical operating voltage is 1750V [62].

3.1.2 Electric Field Regions (Charge Detection)

To detect S2 light, ionisation electrons produced through recoils in the LXe are first transported to the top of the detector, extracted out of the liquid and subsequently accelerated into the gaseous xenon. This is achieved through three electric field region shown in Figure 3.2). To produce a uniform electric field between the cathode and the gate electrodes a set of 57 equally-spaced field rings is embedded within the PTFE and connected by pairs of 2 GΩ high-voltage resistors.

Starting at the bottom of the TPC, the three electric field regions are the *Reverse Field Region* (RFR), the *Forward Field Region* (FFR), and the *Extraction Region*. These are created using four custom-made high voltage electrode grids known as the bottom, cathode, gate and anode grid. Making up the bulk of the TPC volume, the FFR exists between the cathode and the gate grid. In this region, ionisation electrons that are produced in particle interactions are drifted towards the anode under a nominal electric field of $300\,\mathrm{V\,cm^{-1}}$. which is created by operating the cathode at $-50\,\mathrm{kV}$. The extraction region, where electrons are accelerated through the gaseous Xenon to produce S2 light, is the region between the gate grid and the anode grid and has a much higher electric field between $6\,\mathrm{kV\,cm^{-1}}$ to $8\,\mathrm{kV\,cm^{-1}}$. Lastly, the RFR plays a protective function and is responsible for shielding the bottom array

PMTs from the external −50 kV field produced by the cathode. Interactions in the RFR only produce S1 light since ionisation electrons are drifted downwards and therefore do not undergo electroluminescence.

Veto Detectors

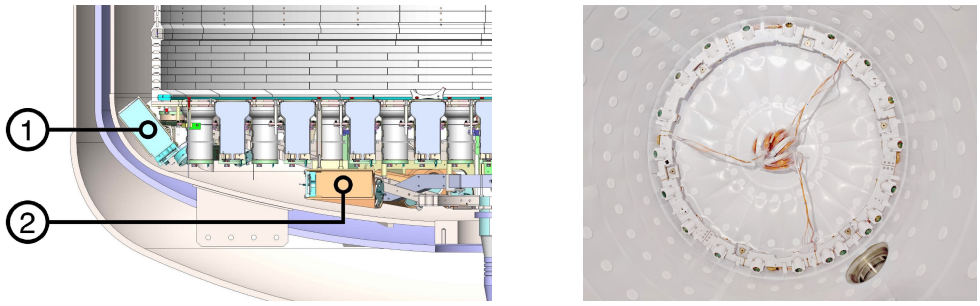
The LZ TPC is surrounded by two optically separated veto detectors that further enhance discrimination of neutron and gamma ray backgrounds. These are the LXe Skin detector that surrounds TPC and the OD that surrounds the OCV.

3.2 Skin Detector

The Xe Skin is a scintillation-only coincidence detector containing approximately 2 tonnes of LXe divided into two regions: a cylindrical side skin region and a dome region underneath the TPC. The side skin is viewed from above by 93 1-inch Hamamatsu R8520-406 PMTs located below the liquid surface, and from below by a further 20 2-inch Hamamatsu R8778 PMTs viewing upward (see Figure 3.5). The dome region of the skin is instrumented with an additional 18 2-inch R8778 PMTs, 12 of which look radially inward and 6 outward. To maximise light collection all PMTs and truss structures are dressed in PTFE, and both the ICV and the bottom dome is lined with PTFE. The high density of LXe in the skin leads to incoming gamma rays, that are likely to interact with the Xe atoms in the TPC, also interacting (or being absorbed) and producing scintillation light in the skin detector, thus making it possible to veto such backgrounds with high efficiency.

3.3 Outer Detector

As shown in Figure 3, the LZ OD surrounds the OCV with nearly 4π coverage and consists of ten segmented acrylic tanks. Four side tanks, three on the top and one on the bottom. The tanks hold 21.5 tonnes of linear alkylbenzene (LAB) liquid scintillator that is doped at 0.1% by mass with gadolinium. The high neutron capture cross section of Gd significantly increases the capture rate thereby reducing the mean neutron capture time. Neutrons are predominantly captured on ^{155}Gd and ^{157}Gd and subsequently release 7.9 MeV and 8.5 MeV in a cascade of, on average,



(a) (b)

Figure 3.5: The CAD drawing on the left (a) shows the location of the 2" bottom side skin PMTs (1) and the lower dome PMTs (2). The photograph on the right (b) shows the highly reflective PTFE panelling attached to the side of the ICV and the lower side skin PMT ring at the bottom of the vessel [11].

4.7 gammas which subsequently produce scintillation light in the LAB that is detected by the 120 8-inch PMTs that view the OD. Approximately 10% of neutrons are captured on H, emitting a single 2.2 MeV gamma. To minimise photon extinction in the OD, a Tyvek curtain covers the PMTs from behind, above and below, and the OCV is also encased in a layer of Tyvek [11].

The position-averaged neutron tagging efficiency of the OD has been measured at 89.3 $\pm$ 3% using AmLi calibration sources planed at nine locations close to the TPC. The OD and Skin vetoes are expected to reduce the total NR backgrounds over a 1000 live days from 12.31 to 1.24 NR counts with the OD providing most of the vetoing power [63].

3.4 Calibrations

Properly understanding the detector response of the LZ TPC and various subsystems requires a rigorous in-situ calibrations programme. The calibrations are used to understand the energy and position reconstruction, understand time offsets between the skin, OD and TPC, and to characterise ER and NR bands, i.e region between 10% and 90% quantile lines in S1 vs S2 space. The LXe self-shielding limits calibration in the central part of the TPC volume using external sources, hence prompting the use of both internal and external calibration sources. An overview of the calibration sources used in LZ is shown in Table 3.1.

External sources are deployed to the vacuum region between the ICV and the

Table 3.1: An overview of the calibration sources used for LZ detailing the half-life of the isotope, the type of decay, the energy range and purpose of the source. The first section in the table below lists the gaseous sources, the second section lists several sealed sources lowered into the cryostat side vacuum, and the third section lists gamma sources and the fourth section contains the two DD sources [10].

Isotope	**Type**	**Energy [KeV]**	$\tau_{1/2}$	**Purpose**
^{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
^{57}Co	γ	122	0.74 y	Xe Skin
^{22}Na	γ	511, 1275	2.61 y	TPC and OD sync
^{133}Ba	γ	356	10.5 y	ER response
^{54}Mn	γ	835	312 d	High-energy ER response
^{60}Co	γ	1173, 1333	5.27 y	High-energy ER response
^{228}Th	γ	2615	1.91 y	High-energy 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
^{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 light and charge yields
DD	n	$272 \rightarrow 400$	—	NR light and charge yields
DD	n	2450	—	NR light and charge yields

OCV via three dedicated conduits. The AmLi, AmBe and ^{252}Cf sources are used to characterise the NR band. In addition, a selection of photoneutron sources are used to calibrate the low energy nuclear recoil response between around 1 keV to 4.6 keV. Understanding this region is especially important for other low energy NR signals of interest such as ^{8}B.

The Adelphi Technologies DD109 deuterium-deuterium fusion neutron generator which produces mono-energetic 2.5 MeV and 272 keV neutrons is used to create localised NR calibration events. This technique was first developed for LUX and allows in-situ NR calibration down to low energies.

LZ uses a number of "internal" calibration sources in gaseous form that are injected into the circulation path to mix with the LXe and reach the central active volume. These sources include three short lived sources (^{83m}Kr, ^{131m}Xe, ^{220}Rn) and two long lived sources (^{3}H, and ^{14}C). The long lived tritium is be deployed as tritiated methane CH_3T gas so that it can be efficiently removed by the getter. The use of tritated methane for calibrations was pioneered by LUX and due to its low Q value of 18.6 keV it is very well suited to calibrate the ER band down to threshold [64].

3.5 Dominant Backgrounds

A summary of the expected backgrounds in LZ within a 5.6 tonne fiducial volume and a 1000 day run after analysis cuts are applied is shown in Table 3.2. These backgrounds are discussed in below.

3.5.1 Material Radioactivity

Radioactivity of materials used in the detector can contribute to ER and NR backgrounds. It has therefore been crucial to minimise decays around the active volume as well the neutron production from spontaneous fission of naturally occurring ^{238}U. As such, all materials used in the construction of LZ underwent an extensive screening campaign for ^{40}K, ^{137}Cs, ^{60}Co, ^{238}U, ^{235}U to ensure a maximum contribution of less than 0.4 nuclear recoil events and 1×10^{-6} events/(keV $\cdot$ kg $\cdot$ year) electron recoil events in a 1000 live days [11].

Table 3.2: Summary of estimated backgrounds from all significant sources in LZ during a 1000 day WIMP search exposure. Counts are shown for a region of interest relevant to a 40 GeV/c^2 WIMP and after single scatter, skin and OD veto, and a 5.6 tonne fiducial volume cuts are applied. The corresponding ER energy range is 1.5 keV to 6.5 keV and a 6 keV to 30 keV range for NRs.

Background Source	ER (cts)	NR (cts)
Detector Components	9	0.07
Surface Contamination	40	0.39
Dust (intrinsic activity, 500 ng cm^{-2})	0.2	0.05
Plate-out (PTFE panels, 50 nBq cm^{-2})	-	0.05
^{210}Bi mobility (0.1 µBq kg^{-1})	40	-
Ion misreconstruction (50 nBq cm^{-2})	-	0.16
^{210}Pb (in bulk PTFE, 10 mBq kg^{-1})	-	0.12
Laboratory and Cosmogenics	5	0.06
Laboratory Rock Walls	4.6	0.00
Muon Induced Neutrons	-	0.06
Cosmogenic Activation	0.2	-
Xenon Contaminants	819	0
^{222}Rn (3.26 µBq kg^{-1})	681	
^{220}Rn (0.09 µBq kg^{-1})	111	
^{nat}Kr (0.015 ppt)	24.5	
^{nat}Ar (0.45 ppb)	2.5	
Physics	258	0.51
^{136}Xe $2\nu\beta\beta$	67	-
Solar Neutrinos: pp + ^{7}Be + ^{13}N + ^{83}Kr	191	-
Diffuse Supernova Neutrinos	-	0.05
Atmospheric Neutrinos	-	0.46
Total	1,131	1.03
Total (with 99.5% ER discrimination, 50 % NR efficiency	5.66	0.51
Sum of ER and NR in LZ for 1000 days, 5.6 tonne FV, with all analysis cuts	6.49	

Events from external material radioactivity can effectively be cut by defining a fiducial volume and utilising the self-shielding property of liquid xenon. However, this cannot help mitigate against radioisotopes that are dispersed in the xenon. Such contaminants include radon emanation from materials and from dust on the detector surfaces which leads to naked beta emission (emission of a beta particle without an associated gamma) of ^{214}Pb in the ^{222}Rn decay chain. Such ER interactions originating from radon emanation are the dominant source of background for LZ [63]. To mitigate against this the materials used in xenon-wetted components in LZ underwent a radon emanation measurement programme before being selected. LZ determines the ^{222}Rn activity to be 3.26 μBq/kg for the active volume [12].

3.5.2 Intrinsic Backgrounds

Since natural xenon contains trace amounts of ^{85}Kr and ^{39}Ar these intrinsic backgrounds which disperse throughout the liquid would create ER events throughout the detector via beta emission. Requirements for mitigating against such intrinsic contamination of the LXe were met through a dedicated Xe purification programme using chromatography [65]. Through *in situ* mass spectroscopy measurements the mass concentrations were determined to be 144 parts per quadrillion (ppq) gram/gram of ^{nat}Kr/Xe and 890 parts per trillion (ppt) gram/gram of ^{nat}Ar/Xe, with a systematic uncertainty on the sampling rate of 15% [66].

3.5.3 Cosmogenic Backgrounds

Rare event experiments such as LZ are subject to several cosmogenic backgrounds, such as neutrons produced from muon-induced electromagnetic and hadronic cascades. The flux of these particles is suppressed through underground operation, attenuation in the water tank that surrounds the OD and through vetoing using the OD. This results in the neutron flux becoming a negligible background contributor in LZ.

Cosmogenic activation, that is activation of isotopes due to the cosmic ray flux, can lead to an overall increase in the background rate. Specifically the activation in the purified xenon during transport which leads to contamination by ^{127}Xe ($T_{1/2}$

= 36.4 d) [67, 68], ^{37}Ar ($T_{1/2}$ = 35.0 d) [69], and activation in the 2.5 tonnes of titanium used in LZ which leads to the production of ^{46}Sc ($T_{1/2}$ = 83.8 d). The equilibrium decay rate for ^{127}Xe was measured to be 2.7 ± 0.5 mBq/kg by the LUX experiment [70]. The ^{46}Sc is decay rate is estimated using GEANT4 [71] and ACTIVIA [72] simulations to be 4.8 mBq/kg of titanium after 6 months of activation at sea level and surface assembly of the TPC within the cryostat at SURF followed by an 8 month cooling down period after transportation underground [11].

3.5.4 Physics Backgrounds

There are three known physics backgrounds in LZ that generate single-scatter events uniformly in the detector with no corresponding veto signal. These are: neutrino-electron scattering (ER), $2\nu\beta\beta$ ^{136}Xe decay (ER) and neutrino-nucleus scattering (NR). ER neutrino backgrounds primarily originate in the form of solar pp neutrinos with small contributions from ^{7}Be and CNO chains [63]. NR neutrino backgrounds are produced by ^{8}B and hep solar neutrinos, neutrinos from diffuse supernovae, and atmospheric neutrinos that can undergo coherent elastic scattering, a standard model process recently measured for the first time by the COHERENT collaboration [73]. Lastly, the expected $2\nu\beta\beta$ ^{136}Xe decay rate is based on measurements by EXO-200 [74] and KamLAND-Zen [75].

3.5.5 Non-standard Backgrounds

A number of rare non-standard backgrounds are also considered. Firstly, multiple scatter of gamma rays where one vertex occurs in a region of the detector from where charge cannot be collected (e.g. the RFR) thus leading to no S2 light from that interaction vertex. These events are dubbed 'gamma - X' and can leak into the NR band. From simulation, LZ expects 0.1 of these events in the fiducial volume.

And secondly there exist a broad category of "accidental coincidence" events that occur if there is a random coincidence between a process leading to an S1 and no S2 and another uncorrelated process creating an S2 with no associated S1. If such a coincidence happens within a time window corresponding to a physical drift length it would appear as a real event in S1-S2 space. Such accidental topologies

are discussed in detail in chapter 3 and 4.

Chapter 4

SR1 WIMP Search and Accidentals

This chapter reports results from LZ's first search for Weakly Interacting Massive conducted during the first science run (SR1) with a 60 day exposure using a fiducial mass of 5.5 tonnes. The results were first published on the 8th of July 2022 on arxiv [12] and have been submitted to Physical Review Letters. Using a profile-likelihood ratio (PLR) analysis, the data have been shown to be consistent with a background-only hypothesis, achieving the highest sensitivity to spin-independent WIMP-nucleon scattering for masses greater than $9\,\mathrm{GeV/c^2}$. The most stringent limit is set for spin-independent scattering at $36\,\mathrm{GeV/c^2}$, rejecting cross sections above $\sigma_{\mathrm{SI}} = 9.2 \times 10^{-48}\,\mathrm{cm}^2$ at the 90% confidence level.

4.1 Detector Conditions

The data used in SR1 result were collected from the 23rd December 2021 to the 11th May 2022. The search data collected during this period totals 89 live days after removing periods for detector maintenance and calibration activity, as well as 3% loss due to data acquisition (DAQ) dead time and a 7% loss to periods removed due to anomalous trigger rates. Additionally, live time is lost as a result of data quality cuts applied to remove periods of data taking during which the TPC experiences elevated rates of activity, these cuts are described in Sec. 4.6. After these omissions a final live time of 60 ± 1d remains during which an interaction could be recorded.

The detector was operated with a drift field of 193 V/cm ± 4% established by the cathode and gate electrodes, and an extraction field of 7.3 V/cm at a radial position

$r = 0$. The maximum drift time in the LZ TPC under these conditions was 951 µs.

The temperature and pressure of LXe were stable to within 0.2%, at 174.1 K and 1.791 bar(a). The liquid level was stable within 10 µm measured through precision capacitance sensors. The 10 tonne of xenon used in LZ was continuously purified at 3.3 tonne/day through a hot getter system and through an inline radon reduction system at a flow rate of 0.0165 tonne/day [76]. As a result, the observed electron lifetime during SR1 was between 5000 µs and 8000 µs, much longer than the maximum drift time in the TPC.

4.2 LZ Data

The LZ DAQ system records events by operating a digital filter sensitive to S2 signals in the TPC. Data is recorded for a 2 ms window before a trigger and 2.5 ms after, creating a 4.5 ms event window. Data is recorded with a low and a high gain-amplification channel for all TPC, Skin and OD PMTs. The raw waveform amplitudes are then normalised by the PMT and amplifier gains. The waveforms and PMT hitmaps from interactions in the detector can be visualised using the official LZ offline event viewer, as shown in Figure 4.1 with an example event from CH_3T electron recoil calibration data.

The waveforms from each channel are then summed for the TPC, Skin and OD. The integrated waveform area represents the number of photons detected (phd), a measure that takes into account the double photoelectron - a phenomenon by which the VUV photons produced by the xenon scintillation lead to the emission of two 'photoelectrons' from the PMT photocathodes [77, 78].

Figure 4.2 shows an S1 pulse from a CH_3T calibration event. Figure 4.2a shows the waveforms recorded by the individual channels and Figure 4.2b shows the summed waveform.

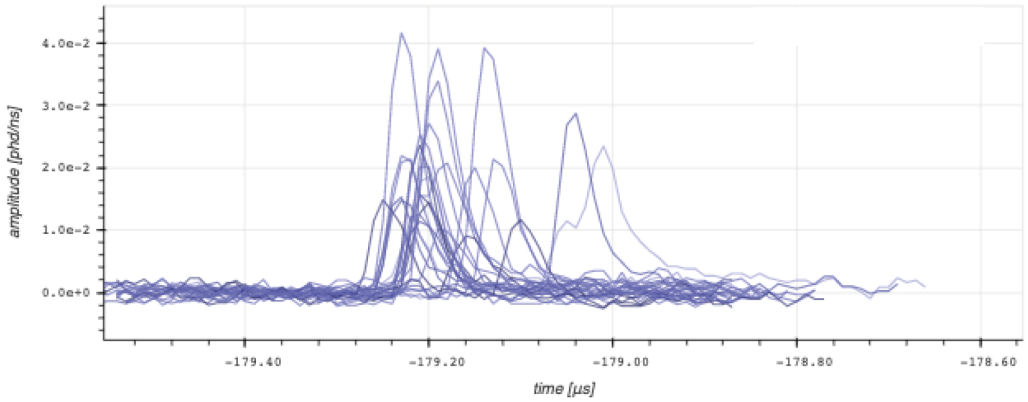
A pulse finding algorithm which determines the pulse boundaries is then applied on this summed waveform. From these separated pulses, S1 and S2 signals are then identified based on their hit pattern and pulse shape. S1 pulses are required to have a 3-fold coincidence requirement i.e. a signal over noise threshold in at

Figure 4.1: A CH_3T calibration event in the TPC visualised using the LZ Offline Event Viewer. The PMT hitmap in the TPC is shown on the right along with the summed waveform in the TPC, Skin Detector and OD shown in the three panels on the left.

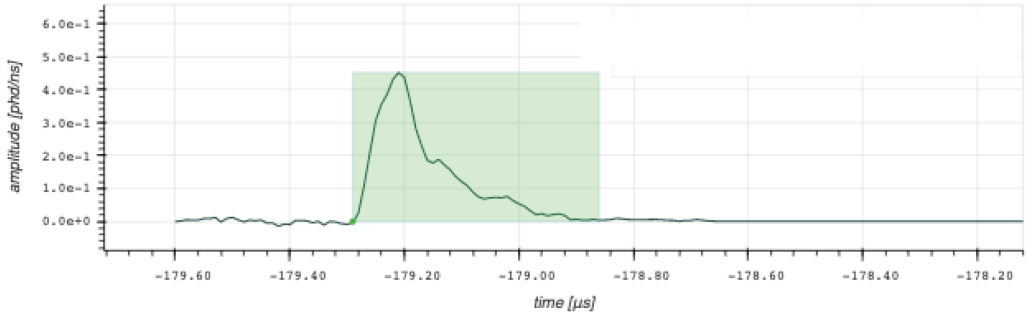
least 3 PMTs. S2 Pulses are easily identified due to their comparatively larger size compared with other pulses in the event, and their concentration in the top array. The nominal S2 threshold for LZ is 300 phd, which corresponds to approximately 6 extracted electrons.

Single scatters are then identified by using the time ordering of the most prominent S1 and S2 pulses. An event with a single S1 preceding a single S2 is classified as a single scatter, whereas an event with a single S1 and multiple S2 pulses is classified as a multi-scatter event. The single S1 requirement for multi-scatter events is due to the merging of the S1 signal from the two interactions since the time of arrival of scintillation light is shorter than the timing resolution of the PMTs.

The S2 hit pattern is used to determine the x, y position of the interaction using the MERCURY algorithm [79]. MERCURY was tuned using a uniformly distributed



(a)



(b)

Figure 4.2: Visualisation of an S1 pulse in the TPC using the Offline Event Viewer. (a) shows the waveforms from the individual channels. (b) shows the summed waveform of the S1 pulse. The green box in (b) marks out the pulse boundaries as determined by the pulse finding algorithm.

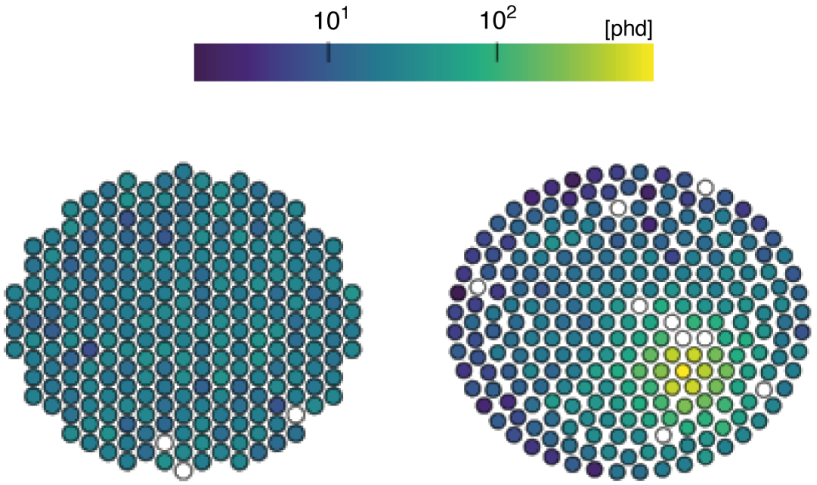


Figure 4.3: A PMT hit map created by an S2 pulse from a CH_3T calibration event. This hit map is used to reconstruct the x, y position of an interaction in the TPC.

radioactive sources in the TPC and has a 1σ resolution of 4 mm for S2 signals of 3000 phd. An example of an S2 hit pattern is shown in Figure 4.3. Due to asymmetric light collection at the TPC edge, the resolution is worse by approximately a factor of two for interactions near the wall. The z position of an interaction is inferred from the drift time, and has a a 1σ resolution of 0.7 mm for events near the cathode electrode, i.e. near the bottom of the TPC.

4.2.1 Reduced Quantities

The S1 and S2 pulses are parameterised and analysed using a number of *reduced quantities* (RQs). These include: the full width at half maximum (FWHM) of the summed waveform; the top bottom asymmetry (TBA) which measures the distribution of light between the top and bottom PMT arrays; the integrated area of the pulse (pulse area), the number of PMTs that coincidentally record light (coincidence); the Area Fraction Time X (AFTX) which are a set of RQs that measure the time at which the integrated area is equal to X percent of the total area, for instance AFT5 is the time at which the integrated area is 5% of the total area. Analysis cuts for background rejection are developed by identifying populations of events using the various RQs. These cuts are discussed in the sections below.

4.3 Calibrations

The detector response model is developed using a number of calibration sources described in this section. Variations in the S1 and S2 signals based on the position of the interaction in the detector are modelled using internal injection sources ^{83m}Kr (32.1 keV and 9.4 keV) and ^{131m}Xe (164 keV) that produce monoenergetic electron recoils. The number of detected S1 photons is corrected to that of an equivalent event at the geometric centre of the detector, and this value is called S1c. The S1 corrections are driven by the variations in the position dependent light collection efficiency and PMT quantum efficiency. The corrected S2 signal (S2c) is normalised to the radial centre and top of the detector, i.e. a drift time of 0. These corrections are motivated primarily by non-operational PMTs and extraction field non-uniformity caused by electrostatic deflection of the gate and anode electrodes. The corrected S1

and S2 values are used in energy reconstruction, which is calibrated in the WIMP region of interest (ROI) using the well understood monoenergetic peaks of ^{83m}Kr and ^{131m}Xe ER sources. At slightly higher energy the monoenergetic transition of cosmogenically activated ^{129m}Xe (236 keV) is used for calibration. The light and charge collection parameters were measured to be $g_1 = 0.115 \pm 0.002$ phd/photon and $g_2 = 46.4 \pm 1.5$ phd/electron. With single electron electroluminescence measured to produce 57.6 ± 1.9 phd/electron we can see an electron extraction efficiency of $80.5 \pm 3.7\%$. The g_1 and g_2 values are measured using the dispersed tritium (CH_3T) source and the DD neutron generator, and validated by comparing ER and NR band predictions to values calculated using NEST [59].

The ER band is defined using the tritiated methane CH_3T beta source, and the NR band using DD neutron sources. These calibration data points are shown in $\log_{10}$S2c-S1c space in Figure 4.4 along with the derived ER and NR bands.

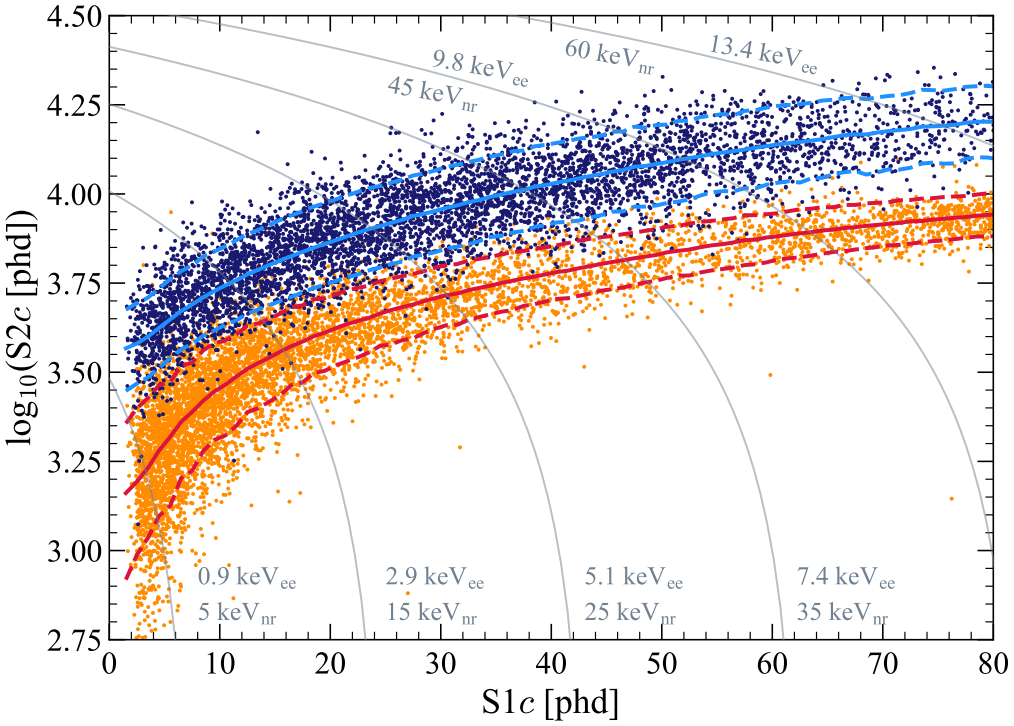


Figure 4.4: Calibration events in $\log_{10}$S2c-S1c from the tritiated methane CH_3T source (dark blue points, 5343 events) and the DD neutron source (orange points, 6324 events). Solid blue (red) lines indicate the median of the ER (NR) simulated distributions, and the dotted lines indicate the ER and NR bands. Thin grey lines show contours of constant electron-equivalent energy (keV_{ee}) and nuclear recoil energy (keV_{nr}). [12]

4.4 Background Model

Source	Expected Events	Best Fit
β decays + Det. ER	218 ± 36	222 ± 16
ν ER	27.3 ± 1.6	27.3 ± 1.6
^{127}Xe	9.2 ± 0.8	9.3 ± 0.8
^{124}Xe	5.0 ± 1.4	5.2 ± 1.4
^{136}Xe	15.2 ± 2.4	15.3 ± 2.4
^{8}B CEνNS	0.15 ± 0.01	0.15 ± 0.01
Accidentals	1.2 ± 0.3	1.2 ± 0.3
Subtotal	276 ± 36	281 ± 16
^{37}Ar	[0, 291]	$52.1^{+9.6}_{-8.9}$
Detector neutrons	$0.0^{+0.2}$	$0.0^{+0.2}$
$30\,\mathrm{GeV/c^2}$ WIMP	–	$0.0^{+0.6}$
Total	–	333 ± 17

Table 4.1: Number of expected events from various sources for the $60\,\mathrm{d} \times 5.5$ tonnes exposure, before and after the combined fit of the background model plus a $30\,\mathrm{GeV/c^2}$ WIMP signal to the selected data. ^{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.

This section describes the background model used in SR1. It consists of nine components that are listed in Table 4.1 including the expected number of events. The backgrounds are grouped together according to their spectra in the ROI or the uncertainty on their rate.

The dominant ER background in SR1 originates from the β decay of impurities dispersed in the xenon. These are: ^{214}Pb from the ^{222}Rn decay chain, ^{212}Pb from the ^{220}Rn decay chain, and ^{85}Kr. The β backgrounds in LZ have a broad energy spectra that is nearly flat across the ROI. The concentrations of the Pb backgrounds are estimated by fitting to energy peaks outside the ROI, and the ^{85}Kr rate estimate is informed by in-situ mass spectroscopy measurements of 144 ± 22 parts per quadrillion gram/gram of $^{\mathrm{nat}}$Kr. There is an additional small ($< 1\%$) contribution to the β background from γ-rays emitted from detector components and the cavern walls.

Another source of ER backgrounds comes from solar neutrinos which have a nearly flat spectrum in the ROI and a rate that can be predicted very precisely following Refs. [80–84]. As such in the background model solar neutrinos are treated separately to ERs from β decay.

Decays from the naturally occurring isotopes of Xe contribute to the ER backgrounds through double electron capture of ^{124}Xe and double β decay of ^{136}Xe. The estimates for these are driven by the known isotopic abundances, lifetimes and decay schemes [85–87].

During SR1 the Xenon contained short-lived cosmogenically produced isotopes that decayed during the run. Most notable of these were ^{127}Xe and ^{37}Ar [67,68,88]. De-excitations of ^{127}Xe following L or M shell electron captures fall within the ROI if the γ-rays from the de-excitations of the daughter ^{127}I escape the TPC. The overall rate of ^{127}Xe electron captures is constrained by measuring the K-shell atomic de-excitations that occur outside the ROI. The Skin is effective at tagging the ^{127}I nuclear de-excitation γ-ray thereby reducing the background rate by a factor of 5. The number of ^{37}Ar events is estimated by calculating the above ground exposure of the xenon to cosmic rays and correcting for decays that occurred prior to the start of the search. A flat constraint of zero to three times the predicted estimate of 97 is imposed due to the large uncertainties on the prediction.

The NR background is made up of radiogenic neutrons and coherent elastic neutrino-nucleus scattering (CEνNS) from ^{8}B solar neutrinos. The rate of ^{8}B neutrinos is calculated using Refs. [80–84] and is small due to the $S2 > 600$ phd requirement in the SR1 analysis. The radiogenic neutron rate is constrained by measuring the number of single scatter events within the FV that are also tagged in the OD and then applying the measured neutron tagging efficiency ($88.5 \pm 0.7\%$). A likelihood fit of the NR componenet in the OD-tagged data is consistent with observing zero events, as predicted by simulations and radioassay of the detector materials. A data driven constraint of 0.0 ± 0.2 events is therefore applied to the search, which agrees with simulations based on detector material radioassay [11].

Lastly, accidental coincidence events are also observed in the LZ data. These

occur when an isolated S1 pulse coincides with an isolated S2 pulse within a time window that corresponds to one drift length. Improving the rejection of these event forms a large part of the study presented in this thesis, as such these events are discussed in more detail in the next section.

4.5 Accidental Backgrounds

Understanding the origin of accidental events and applying effective suppression techniques is crucial to conducting a successful WIMP search. This section covers the known sources of isolated S1 and isolated S2 pulses, and the approach to determining the distribution and rate of accidental accidental events during SR1. Accidentals are removed through a combination of cuts on pulse-level RQs and on event level cuts that utilise correlations between pulse RQs and the drift time. These cuts are described in the next section.

4.5.1 Isolated S1 pulses

An isolated S1 pulse is a pulse that meets the 3-fold coincidence requirement and is therefore classified as an S1 but has no measured S2 associated with it. These are either pulses that are misclassified as an S1 or caused by scatters in regions of the TPC from where charge is either lost (e.g. near the wall) or cannot be collected (e.g. in the RFR). Three main sources of isolated S1 like pulses that are not produced due to scintillation were identified during SR1.

Firstly, PMT dark counts or single photoelectrons (SPEs) measured in PMTs can pile up leading to 3-fold coincidence pulses that appear as candidate S1 signal. Dark counts are typically caused by spontaneous thermionic emission of single electrons from the photocathode or early dynode stages [89] and has the same response as a photoelectron.

Secondly, Cherenkov light produced in the PMT quartz window through the β decay of ^{40}K contaminants present in the glass. Such pulses can meet the 3-fold coincidence requirement and appear as candidate S1 signals if, for instance, some of the light escapes the PMT of origin and reaches other PMTs or the Cherenkov pulse coincides with PMT dark counts. Most of the photons from a Cherenkov pulse

are concentrated in the PMT of origin and hence have a *high single channel* (HSC) contribution, i.e. most of the light is in a single PMT. Nominally such events can be removed relatively easily with a cut on light concentration (see HSC cut described in section 4.6.3) however in cases where the Cherenkov light is emitted from a dead PMT, no light is measured in the PMT of origin making such cases harder to identify. Accidental events from such isolated S1 pulses are removed using event level cuts that exploit the expected drift time dependence of some S1 and S2 pulse RQ, such as the top bottom asymmetry of the S1 pulse and the width of the S2 pulse (see section 4.6 for details).

As part of work undertaken during this PhD, Cherenkov events as a source isolated S1 signals were studied extensively through GEANT4 simulation. The machine learning techniques developed to target accidental events that are presented in the next section were initially developed on simulations to target such Cherenkov pulses and that earlier work has informed the current understanding of this isolated S1 source.

Thirdly, fluorescence in the PTFE [90] following large pulses can contribute to the isolated S1 rate. For instance, the tail of a large S2 pulse shows an elevated rate of SPEs that are possibly due to fluorescence in the PTFE. If such an S2 occurs in the time between event triggers (i.e. during dead time) the tails with an elevated SPE rate can leak into another event contributing to a 3-fold isolated S1 pulse through SPE pileup or coincidence with an other pulse (e.g. Cherenkov).

4.5.2 Isolated S2 Pulses

In events near the liquid surface with very short drift times the S1 pulse can merge with the S2 pulse creating a pulse that looks like an isolated S2. Similarly, recoils in the gas region at the top of the detector cause merged scintillation and electroluminescence signals which appear as isolated S2 pulses with an early peak or short rise time. In some interactions, the S1 pulse is not detected due to the typically low S1 light collection efficiency achieved in dual-phase xenon TPCs. This can result in an isolated S2 pulse (e.g. few keV scale interactions) in the data. Additionally, as was observed by LUX and XENON1T the electrode grids can contribute to the S2-

only rate through radiogenic emissions [91]. Three radiogenic sources have been identified: firstly, ^{210}Pb plated out on the wires during the production of the grids, secondly ^{222}Rn daughters that could drift onto the grid wires during operation, and thirdly intrinsic contamination due to presence of trace ^{238}U and ^{232}Th in the grid wires.

4.5.3 Rate Calculation

The before cuts distribution of accidentals in $\log_{10}$S2c-S1c space is determined by "stitching" together waveforms from the isolated S1 and S2 events to give composite single scatter accidental events. This data is referred to as the *ChopStitch accidental* data as it was constructed using LZ's *salt*[1] generator which is called ChopStitch. The distribution of ChopStitch events in the ROI before analysis cuts is shown in Figure 4.5. This data set is used to evaluate the performance of the analysis cuts against accidental events.

The total number of before cuts accidental events in the ROI is estimated by measuring the number of events with an unphysical drift time (UDT) (> 951 µs), i.e. a drift time that corresponds to a distance that is larger than the length of the TPC. Since accidentals are coincidentally paired isolated S1 and isolated S2 pulses, we observe such UDT accidental events at the same rate as the physical drift time accidental events. No correlations between isolated S1 and isolated S2 events were found in the SR1 data. An example event display visualisation of a UDT event is shown in Figure 4.6.

The ChopStitch accidentals distribution in $\log_{10}$S2c-S1c is then normalised to the rate of UDT events in the ROI after OD and Skin veto cuts, fiducial volume cuts and live time cuts are applied. The final accidentals distribution and rate used in the SR1 background model is determined by applying all the cuts targeting accidental events (as detailed in the next section) on this normalised distribution.

[1] Salting is the process by which fake events (salt) are added to the data used to develop analysis cuts to mitigate against experimental bias. It is used as an alternative to blinding [80].

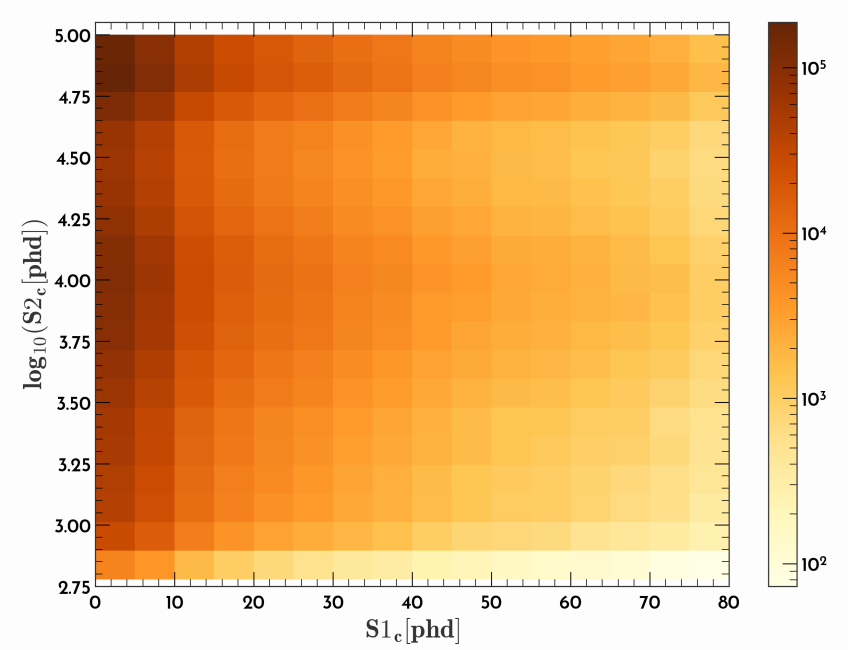
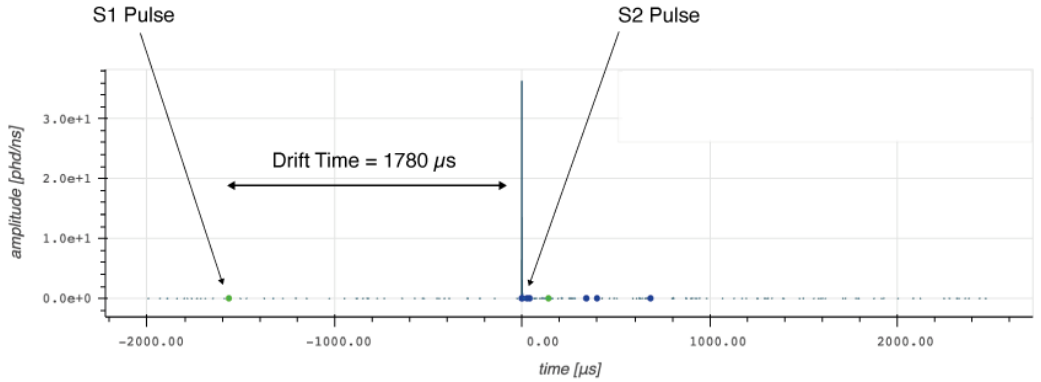
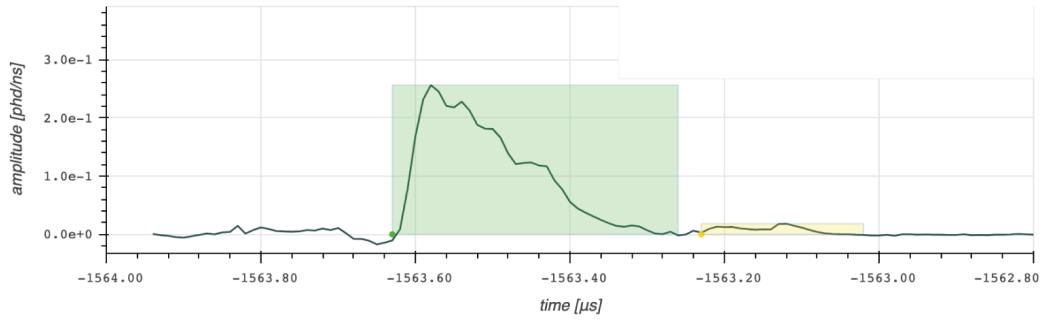


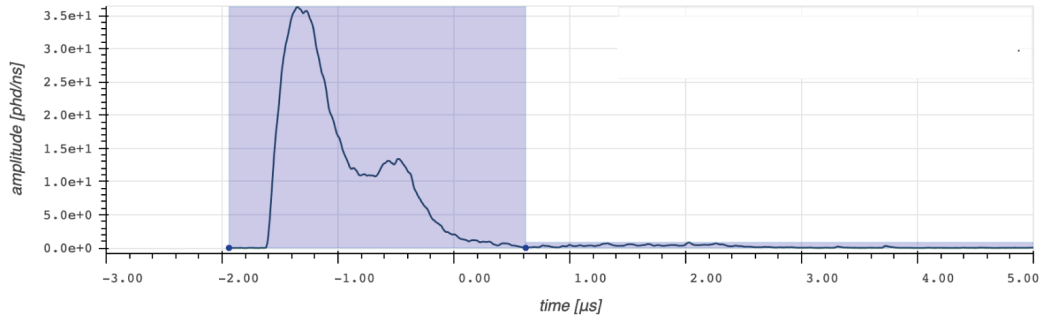
Figure 4.5: The distribution of the ChopStitch accidentals in the ROI in $\log_{10}$S2c-S1c space.



(a)



(b)



(c)

Figure 4.6: An accidental coincidence event with an unphysical drift time of 1780 µs visualised using LZ's official offline event display. (a) shows the full 4 µs event window with the position of the S1 and S2 pulse highlighted; (b) shows the S1 pulse with the pulse boundaries defined by the green box, and (c) shows the S2 pulse with the pulse boundaries defined by the purple box.

4.6 SR1 Core Cuts

The WIMP search ROI during SR1 is defined as $3 < S1c[\text{phd}] < 80$, and $600 < S2c[\text{phd}] < 10^5$. Poorly reconstructed events, spurious pulses and other poor quality data are removed through a set data quality cuts. In addition, a set of analysis cuts were developed to target accidentals that remove more than 99.5% of the ChopStitch accidentals leading to an expected 1.2 events remaining in the ROI after cuts. This section describes the live time, physics and analysis cuts used in the SR1 analysis.

4.6.1 Live Time & Data Quality Cuts

The cuts described in this section target events recorded during periods of data taking during which the TPC experienced elevated rates. Such cuts lead to a loss of live time. Figure 4.7 shows the events in the ROI before and after the live time cuts are applied. The live time cuts are:

1. **Hot Spot Exclusion**

   This cut removes periods of data taking during which thedetector saw elevated emissions from "hot spots" on the electrode grids. These flare ups are periodic and last for $\sim$ 10s. Grid hot spots during SR1 are discussed in Ref. [91].

2. **E-train Veto**

   Elevated rates of single electron emission are observed following a large S2. Events with such "e-trains" were removed during SR1 using the *e-train veto* which removes periods of data taking during which the electron emission rate is above background levels, thus resulting in a loss of live time.

3. **Muon Veto**

   The muons produced in the atmosphere reach the Davis Cavern at a rate of $5.31 \pm 0.17 \times 10^{-9}$ s^{-1}cm^{-2} leading to an expectation of roughly 13.4 ± 0.4 muons per day [92]. As these muons pass through the TPC they leave long trails of ionisation producing localised regions of high electron rates lasting up to a few seconds. The Muon Veto uses the coincident muon signal left in the TPC, OD and Skin detector to veto muon candidates events, and impose a hold-off to remove periods of the aforementioned high electron rates.

4. **Elevated S1-Rate Exclusion**

   The *S1 Elevated Rate Exclusion* cut removes periods of data taking during which the rate of pulses passing the 3-fold coincidence requirement, and therefore being classed as an S1, is higher than usual.

5. **Bad Buffer Exclusion**

   A data quality cut that removes events in which the presence of DAQ buffers lead to an incomplete waveform being recorded.

6. **Burst Noise Exclusion**

   Occasionally the PMTs in LZ experience a burst of noise that is particularly problematic in the OD PMTs. The burst noise exclusion cut removes periods when the OD PMTs are experiencing such noise.

7. **Elevated SPE Rate**

   This cut aims to remove periods of data taking during which there is an elevated rate of single photoelectrons detected in the TPC PMTs. As mentioned in 4.5.1, such SPEs can pile up and create isolated S1-like pulses.

8. **Excess Area**

   Finally the excess area cut removes events with a larger than expected amount of light detected in an event before, between and after an S1 and an S2. This can be caused due to an S2 occurring in the detector during DAQ deadtime or a trigger hold-off leading to an e-train or muon train that leaks into another event.

4.6.2 Physics Cuts

The set of cuts listed in this section do not affect live time and events are removed either because of a coincident signal in the veto detectors (Skin and OD) or because they occurred outside of the fiducial volume.

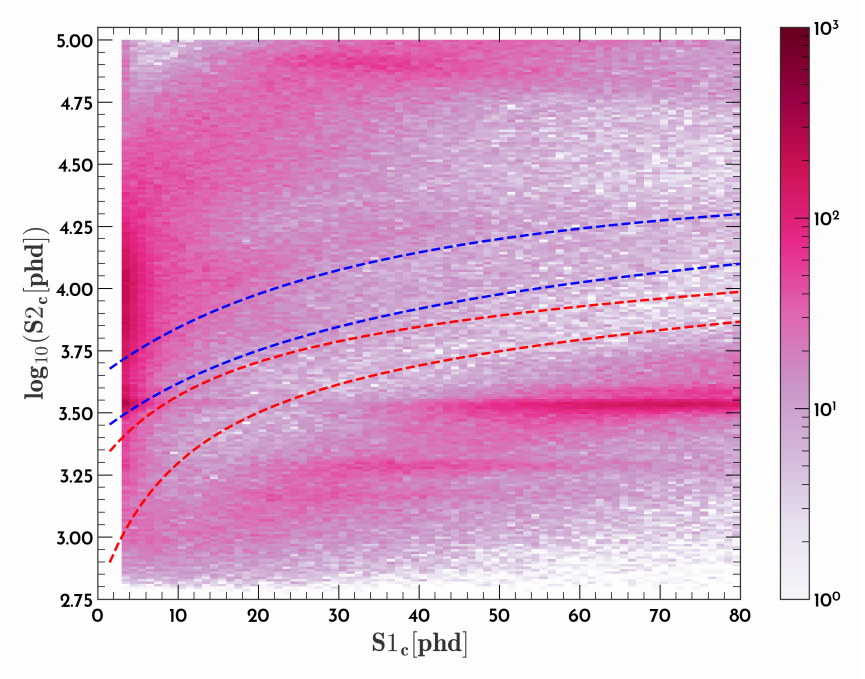
Fiducial Volume

Events outside the central fiducial volume (FV) are removed by making a cut in radial position and drift time. This cut utilises the self-shielding capability of Liquid

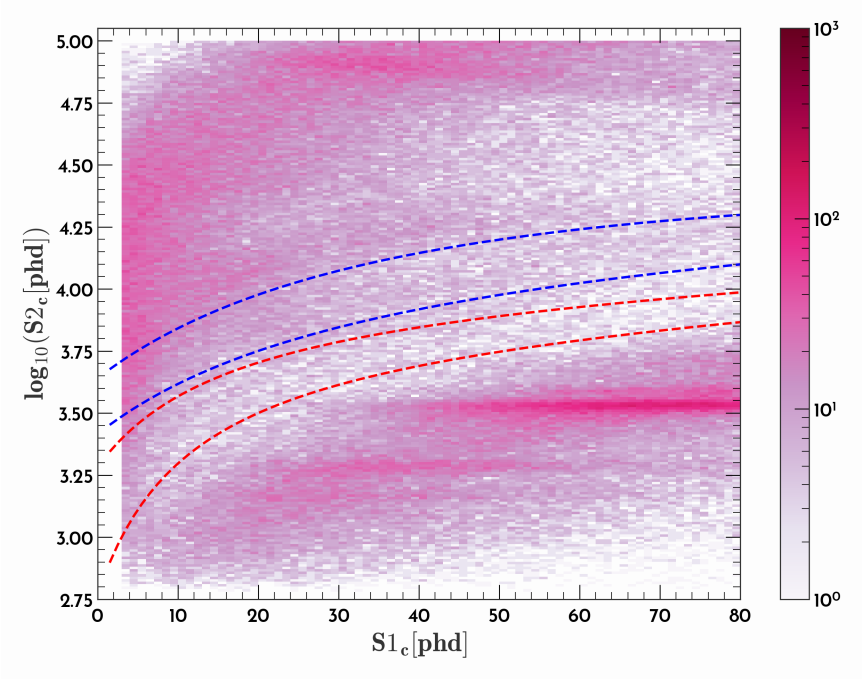
Xenon to remove a significant amount of background events including backgrounds from the wall. In addition, events within 6 cm of the (x,y) position of two ladders of TPC field-cage resistors embedded in the TPC wall are also removed. This leaves a 5.5. tonne liquid xenon target. Figure 4.8a shows the distribution of events in reconstructed r^2 and z, in which one can see a higher density of events in the outer regions of the detector as well as a quieter, self-shielded inner region. Figure 4.8b shows the events remaining in the ROI before and after the FV cut is applied. Applying the FV cut reduces the $\sim$150,000 events that pass the live time and data quality cuts to 655.

Veto Detector Cuts

The instrumented Skin and OD (described in Sec. 3.1.2) are key to LZ's background removal strategy. The Skin and OD veto cuts remove events with a coincident interaction in either the Skin or the OD in order to target interactions from γ-rays and neutrons. Events from cavern γ-rays are removed by vetoing events with a prompt signal in the OD (Skin) within $\pm$0.3 µs ($\pm$0.3 µs) of a TPC S1. Neutrons are removed using a delayed coincidence veto whereby events with an OD pulse of greater than $\sim$200 keV within 1200 µs are tagged as neutrons. Figure 4.9 shows the 539 events inside the FV in $\log_{10}$S2c-S1c space that pass the live time cuts and the OD and Skin vetoes.

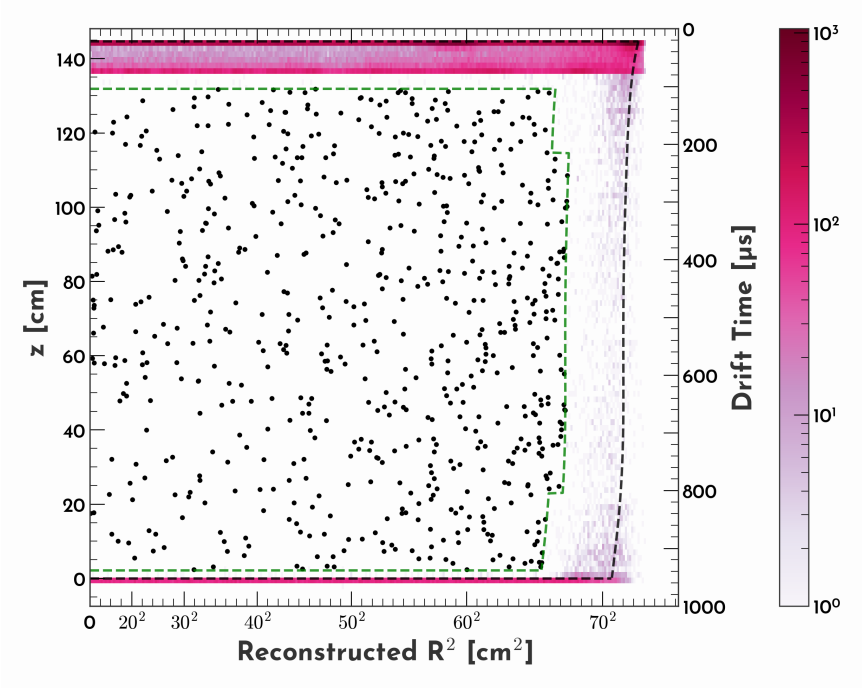


(a)

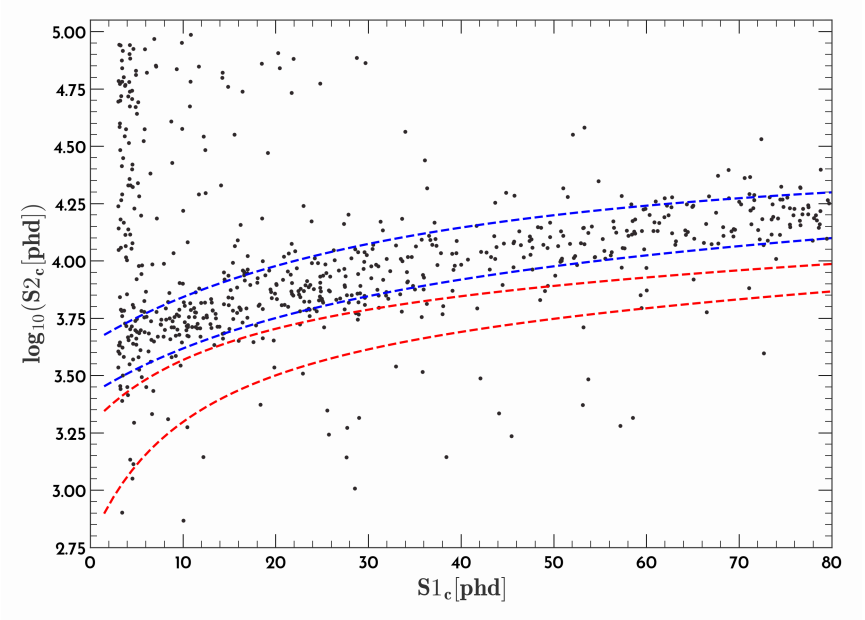


(b)

Figure 4.7: Distribution of single scatters in the TPC in $\log_{10}$S2c-S1c space before and after the live time cuts are applied. The red and blue dashed lines show the ER and NR band respectively. (a) shows $\sim$ 300,000 single scatters in the ROI with no cuts applied; (b) shows $\sim$ 150,000 events after the live time cuts are applied.



(a)



(b)

Figure 4.8: Distributions of single scatters in the TPC showing the impact of the fiducial volume cut. (a) shows the positional distribution in reconstructed r^2 and z of the $\sim$ 150,000 events in the ROI that pass the live time and data quality cuts. The dashed black line shows the extent of the active TPC and the dashed green line shows the FV. The events that are outside the FV are shown with the pink histogram and events inside are shown with the black markers; (b) shows the 655 events inside the FV in $\log_{10}$ S2c-S1c space. The ER and NR band are shown with the dashed blue and red lines.

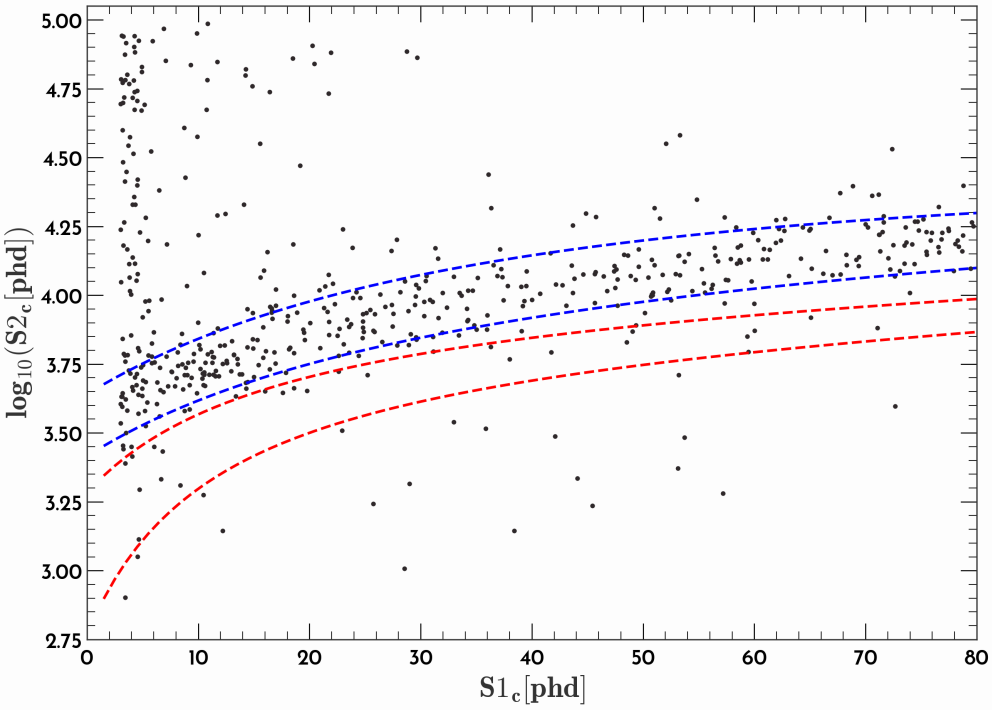


Figure 4.9: Shows the 539 events inside the FV in $\log_{10}$S2c-S1c space that pass the live time and data quality cuts, and the OD and Skin vetoes. The ER and NR band are shown with the dashed blue and red lines. The accidental cuts have not yet been applied to these events.

4.6.3 Cuts to Remove Accidental Events

The cuts presented in this section were developed to target accidental coincidence events based on both pulse shape information such as TBA, pulse width etc, and event-level information such as the relationship between the S2 width and drift time. Using the cuts below a 99.5% accidental rejection efficiency was achieved leaving, 1.2 ± 0.3 events in the ROI.

S1 Cuts

- **High Single Channel Cut**

  This cut targets pulses with a high single channel contribution, i.e. pulses with a larger than expected concentration of light in a single PMT. An example of such pulses (as discussed in the previous section) is Cherenkov light emitted in the PMT glass where most of the light is concentrated in the PMT of origin. The principle RQ to target such pulses is the *max channel area* which measures number of photons detected in the PMT that records most light.

- **Multiple Prominent Pulses Cut**

  This cut removes events with multiple prominent S1 pulses that are classified as single scatters. Such events can't certainly be treated as single scatters, nor are they necessarily just made up of multi-scatter events as the additional pulses could be from non TPC interactions such as a Cherenkov pulse in the PMT glass that is misclassified as an S1.

- **Stinger**

  A class of pulses referred to as stingers was identified in which an S1 like pulse immediately follows an electroluminescence pulse. This S1-like pulse is caused by electrons that drift past the anode plane into a low field region and then turn around and terminate on an anode wire.

- **S1 Top Bottom Asymmetry vs Drift Time**

  The Top Bottom Asymmetry (TBA) of a pulse is an RQ that measures the ratio of light in a pulse between the top and bottom array. A TBA of

-1 indicates that all the light is measured in the bottom array, whereas a TBA of 1 indicates all the photons arrive in top array PMTs. The TBA of S1 pulses is dependent on the position of the interaction in the TPC through the fact that S1 light collection efficiency depends on the z position is different for the top and bottom array.

- **Pulse Shape Cuts**

  These are a collection of cuts based on pulse shape differences between scintillation S1 signals and other sources of 3-fold coincidence pulses classified as S1. Cuts are made on the following 5 RQs: the FWHM of the S1 pulse; the root mean squared width; the fraction of the total pulse area detected within the first 50 ns and 100 ns (prompt fraction 50 ns and prompt fraction 100 ns); and the AFT75 - AFT 25[ns].

- **S1 Photon Timing Cut**

  This cut utilises observed differences between the distribution of the channel peak positions for scintillation S1 signals, as seen in calibration data compared with isolated S1 pulses from random trigger data. The cut is applied on the root mean square of the channel peak times.

S2 Cuts

- **S2 Width vs Drift Time**

  Due to the linear relationship between the diffusion of the drifting electron cloud and the z position of the interaction, the width of the measured S2 pulse is correlated with the drift time of an event. This cut utilises this relationship to reject accidental coincidence events for which no such relationship exists.

- **S2 Pulse Shape cuts**

  Similar to the S1 pulse shape cuts, accidental events with anomalous S2 pulses are rejected using cuts applied on S2 pulse shape RQs. These include: a cut is applied on the root mean squared (RMS) width of the S2 pulse, a cut on the S2 *rise time* - measured by the difference between

the AFT 5 and the AFT 1, and a cut on the FWHM of the S2 pulse.

- **S2 XY Reconstruction**

  This cut removes events where two S2 pulses coincidentally merge resulting in an ambiguous PMT hit pattern which leads to their position not being properly reconstructed. Such events could be multi-scatters or due to the coincidence of a candidate event with an isolated S2 pulse. Events with such S2 signals are removed using the goodness of fit output from the MERCURY algorithm.

- **Corrected S2 TBA (Above Anode Cut)**

  This cut removes S2-only pulses from events occurring in the gas region of the TPC using the differences in the TBA of an S2 originating from interactions in the gas compared with S2 signals from the LXe.

The above accidental cuts are summarised in table 4.2 which shows the number of expected accidental events in the 60 day exposure used in SR1. After applying data quality and veto cuts we expect 350 accidental events in the TPC which is reduced to 1.3 events once all the cuts are applied. The cumulative survival fraction as well as the number of expected events after each cut is also given. It should be noted however that this values given are for the specific cut order in the table whereby analysis begins with the fiducial volume cut and ends with the corrected S2 TBA cut. The expected number of accidental events is calculated by applying the cuts to the ChopStitch accidental events (introduced in Sec. 4.5.3) normalised to the number of UDT events.

SR1 Accidentals Cut	Expected Accidental Events	Cumulative Survival
-	350	100.00%
Fiducial	222.6	63.60%
S1 HSC	177.55	50.73%
S1 prominence	177.34	50.67%
Stinger	101.81	29.09%
S1 TBA vs DT	78.12	22.32%
S1 pulse shape	69.7	19.91%
S1 photon timing	66.55	19.01%
S2 Pulse Shape	6.36	1.82%
S2 width vs DT	3.14	0.90%
S2 XY Reconstruction	2.36	0.67%
Corrected S2 TBA	1.3	0.37%

Table 4.2: Number of events remaining after each stage of the SR1 accidentals cuts. A cumulative survival rate of each cut applied in the above order is also given.

4.7 WIMP Search Result

A total of 335 events remain in the ROI after all cuts are applied. Figure 4.10 shows the distribution of these events in $\log_{10}$ S2c-S1c along with contours representing a 30 GeV/c^2 WIMP, a flat NR distribution and the background model. The signal model assumes a spin-independent Maxwell Boltzmann velocity distribution, parameterised as in Ref [80] with inputs from Refs. [93–98]. The WIMP model has an approximately exponentially decreasing energy spectrum with a shape that depends on the WIMP mass [52]. The distribution of the surviving events in reconstructed r^2 and z is shown in Figure 4.11.

The statistical inference of the WIMP scattering cross section and mass is performed with an extended profile likelihood statistic in the $\log_{10}$ S2c-S1c observable space, with a two-sided construction of the 90% confidence bounds [80]. The background and signal component shapes are modelled in the observable space using the GEANT4-based package BACCARAT [99, 100] and a custom simulation of the LZ detector response using the tuned NEST model. Figure 4.12 shows the background model and data as a function of reconstructed energy.

Above the smallest tested WIMP mass of $9\,\mathrm{GeV/c^2}$, the best-fit number of WIMP events is zero, and thus the data are consistent with the background only hypothesis with a p-value of 0.96. Figure 4.13 shows the 90% confidence limit on the spin independent WIMP-nucleon cross section signa1 as a function of mass. The minimum of the limit curve is at $m_\chi = 36\,\mathrm{GeV/c^2}$ with a limit of f $\sigma_{\mathrm{SI}} = 9.2 \times 10^{-48}\,\mathrm{cm^2}$. The LZ experiment has achieved the highest sensitivity to spin-independent WIMP-nucleon scattering for masses greater than $9\,\mathrm{GeV/c^2}$. For WIMP masses between $13\,\mathrm{GeV/c^2}$ and $36\,\mathrm{GeV/c^2}$, background fluctuations produce a limit that is constrained to a cross section such that the power of the alternative hypothesis is $\pi_{crit} = 0.16$ [12, 101].

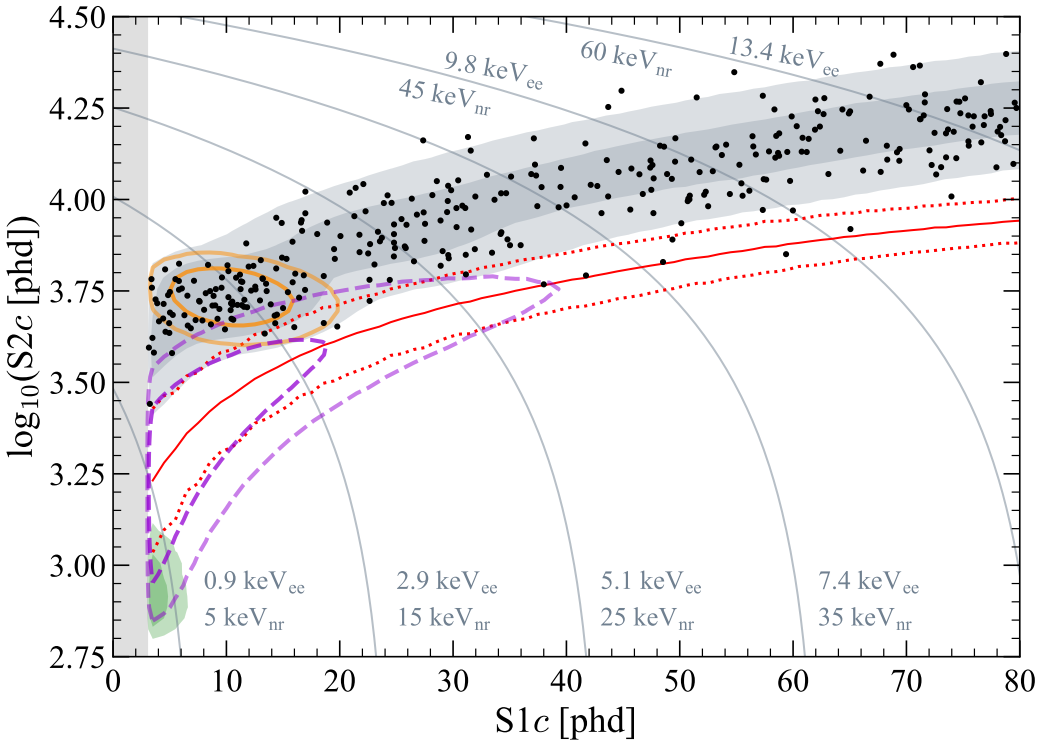


Figure 4.10: WIMP-search data (black points) after all cuts in $\log_{10}$S2c-S1c space. Contours enclose 1σ and 2σ of the following models: the best-fit background model (shaded grey regions), the ^{37}Ar component (orange ellipses), a $30\,\mathrm{GeV/c^2}$ WIMP (purple dashed lines), and ^{8}B solar neutrinos (shaded green regions). The red solid line indicates the NR median, and the red dashed lines NR band. Thin grey lines indicate contours of constant energy [12].

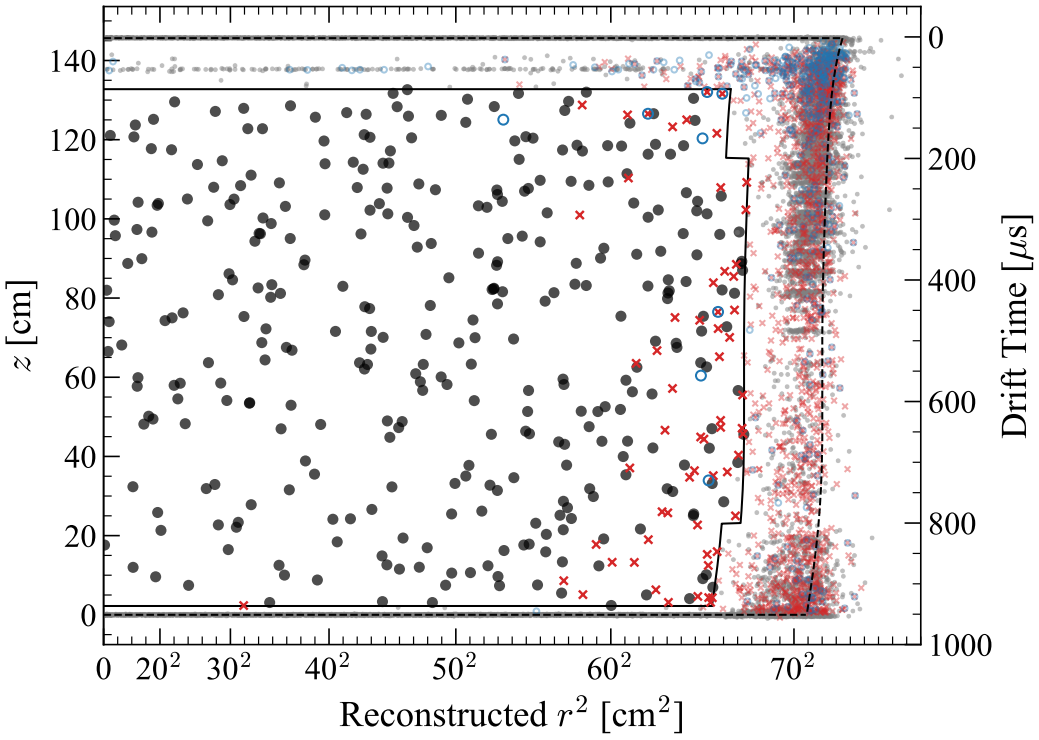


Figure 4.11: Data in reconstructed r^2 and z after all analysis cuts. Black (grey) points show the data inside (outside) the FV. Red crosses and blue circles show events vetoed by a prompt LXe Skin or OD signal, respectively. The solid line shows the FV definition, and the dashed line shows the extent of the active TPC. Field non-uniformities cause the reconstructed r position of the active volume boundary to vary as a function of z. Events with drift time of approximately 50 µs are from recoils in the gas which produce S1 and S2 pulses with a fixed time separation [12].

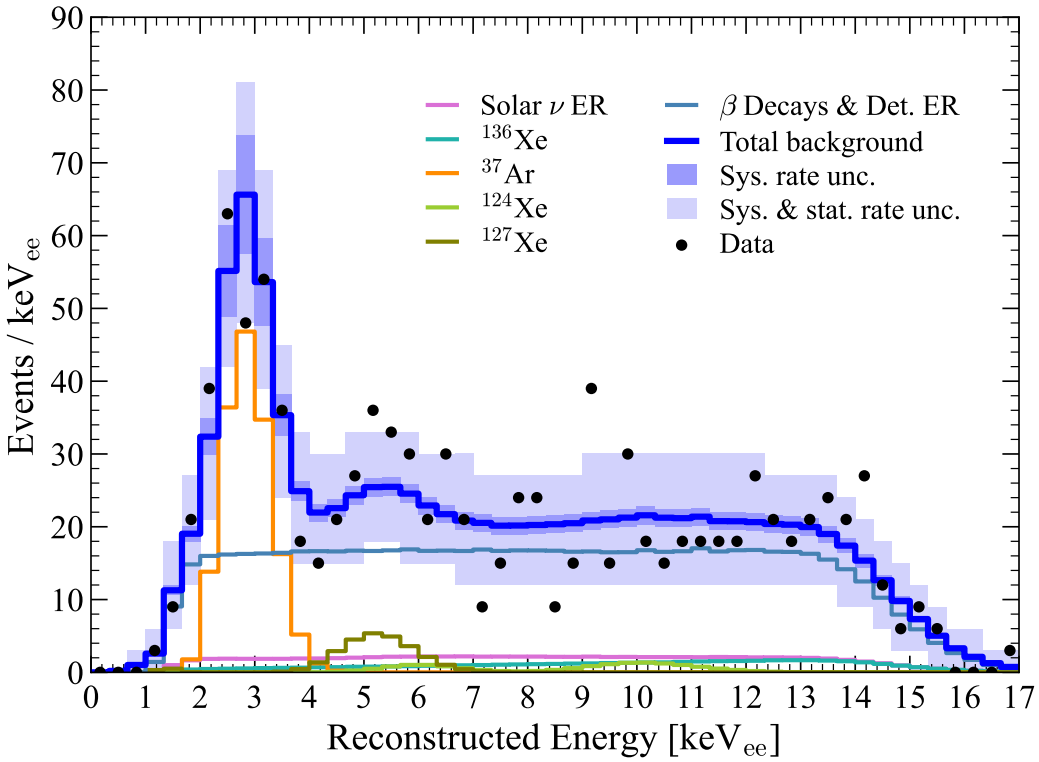


Figure 4.12: Reconstructed energy spectrum of the best fit model. Data points are shown in black. The blue line shows total summed background. The darker blue band shows the model uncertainty and the lighter blue band the combined model and statistical uncertainty. Background components are shown in colours as given in the legend. Background components from ^{8}B solar neutrinos and accidentals are included in the fit but are too small to be visible in the plot. [12].

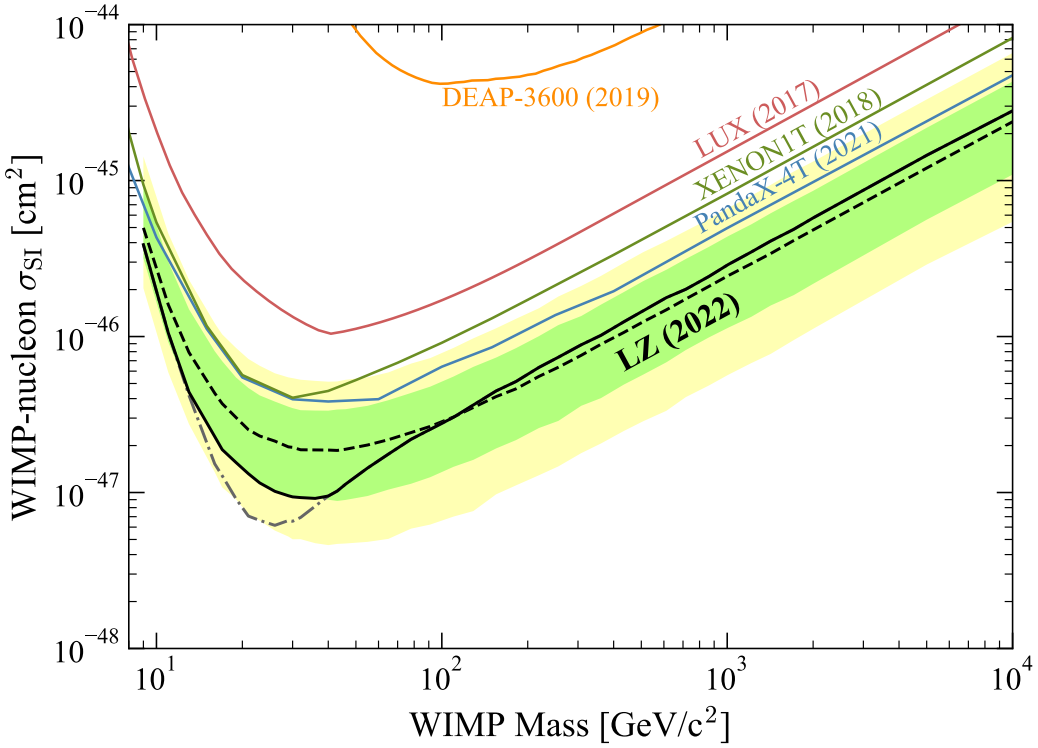


Figure 4.13: The 90 % confidence limit (black line) for the spin-independent WIMP cross section vs. WIMP mass. The green and yellow bands are the 1σ and 2σ sensitivity bands. The dotted line shows the median of the sensitivity projection [12]. Also shown are the PandaX-4T [13], XENON1T [14], LUX [15], and DEAP-3600 [16] limits.

Chapter 5

Using XGBoost to Reject Accidental Coincidence Events

This chapter outlines a methodology developed to reject accidental events in the WIMP-search data using a Boosted Decision Tree based analysis using the XGBoost model [102]. Using this analysis the number of accidentals surviving is reduced to 0.54 $\pm$ 0.3 events in the full ROI and 0.09 $\pm$ 0.05 inside the NR band compared with 1.2 $\pm$ 0.3 and 0.36 $\pm$ 0.1, respectively in the SR1 analysis. The four-fold reduction of accidentals populating the NR band is achieved at the cost of a slight loss in signal efficiency, which reduced from 92.3 $\pm$ 0.2% for NR events using the published SR1 analysis to 73.0 $\pm$ 0.2% using the analysis presented in this chapter. Applying this BDT analysis to the SR1 WIMP-search data leaves 311 events in the ROI compared with 335 remaining using the SR1 analysis. One of the events removed is identified as an accidental and the rest are true single scatters that are sacrificed in the interest of signal purity.

Although the BDT outperforms the standard sequential cuts approach taken in SR1, another benefit to the approach developed here is that it allows analysers to combine a set of analysis cuts into a single cut on a BDT score. Tuning a single cut to increase the signal purity is much more feasible than tuning nine individual cuts. The performance gain and flexibility offered by this approach is advantageous given that the the tolerance towards accidental events depends on the particular physics analysis being undertaken. For instance, a future search for low-energy ^{8}B solar

neutrinos utilising a lower coincidence requirement may require a higher accidental rejection rate at the expense of a lower but tolerable signal efficiency.

5.1 Boosted Decision Trees

The use of Boosted Decision Trees (BDTs) for event classification has been a mainstay in particle physics for some time now, having been first used by the D0 experiment in the search that led to the first evidence of single top quarks [103]. More recent uses include the search and discovery of the $VH \to b\bar{b}$ decay by the ATLAS experiment [104], as well as in rare event experiments such as LUX [105] and PandaX-II [106]. BDTs provide a fast, scalable and relatively simple to train multivariate analysis approach that has regularly matched the performance or outperformed deep neural networks in machine learning competitions (e.g. Kaggle challenges) [102]. This section describes the working of BDTs in general and of the XGBoost algorithm in particular.

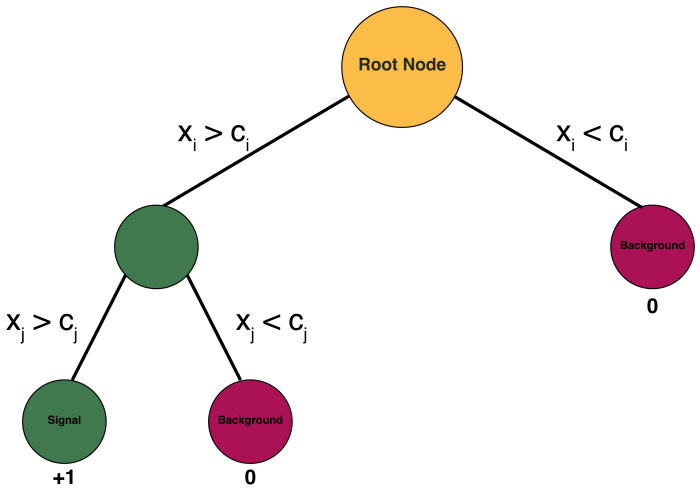


Figure 5.1: A single decision tree with 2 nodes corresponding to a binary cut on two features x_i and x_j with cut values c_i and c_j.

A single decision tree (as shown in Figure 5.1) is a relatively intuitive way to categorise data. However, it is susceptible to statistical fluctuations in the training and test data. This is overcome by using tree ensemble models which make predictions

using a weighted sum of the outputs of K individual decision tree models.

$$\hat{y}_i = \sum_{k=1}^{K} w_k f_k(x_i) \tag{5.1}$$

where $\hat{y}_i$ are the model predictions, f_k is the k^{th} decision tree and w_i is the weight associated with the output of i^{th} tree.

The set of functions that form the tree ensemble are learnt by minimising a learning objective or loss function. For classification tasks this is usually the binary cross entropy function, also known as the Shannon Entropy, defined as:

$$L = \sum_{i}^{m} l(\hat{y}_i, y_i) = \sum_{i}^{m} (y_i \log(\frac{1}{\hat{y}_i}) + (1 - y_i) \log(\frac{1}{1 - \hat{y}_i}) \tag{5.2}$$

where the sum is over all events in the training data m, y_i is the truth value, and $\hat{y}_i$ is the model prediction. This loss function represents the expectation value of the *suprisal* or the "information content" given by $\log(\frac{1}{p(x)})$, where $p(x)$ is the probability of a random event x. The suprisal tends to zero as $p(x)$ tends to one, or in other words, the more sure one is of an outcome x the lower the suprisal. As a result, the loss function is minimised when the BDT output $\hat{y}_i$ is close to either 1 or 0 for signal and background events, respectively.

Since tree ensemble methods contain functions rather than scalars they cannot be optimised using traditional optimisation techniques in Euclidean space. Instead, such models are trained using methods like gradient tree boosting (e.g. XGBoost [102]) or adaptive boosting (e.g. AdaBoost [107]). Intuitively, gradient boosting can be understood as a procedure whereby at each iteration stage k a new tree $f_k(x_i)$ is added in an attempt to rectify the errors of the sum of all the previous trees. The Extreme Gradient Boosting (XGBoost) algorithm which is used in this study, improves on the general approach by adding a regularisation term to the loss function that penalised model complexity and mitigates against overfitting [102]. The

loss function used in XGBoost is defined below as:

$$L = \sum_{i}^{m} l(\hat{y}_i, y_i) + \sum_{k} \Omega(f_k) \;\; ; \;\; \Omega(f_k) = \gamma T + \frac{1}{2}\lambda \|w\|^2 \tag{5.3}$$

where γ and λ are regularisation constants, T is the number of leaves in each tree and w is the sum of the output scores of a single tree. In addition, the XGBoost algorithm uses a second order Taylor approximation of the loss function in E.q 5.3 which allows for fast optimisation [108]. For a detailed discussion on the XGBoost algorithm consult Ref. [102].

5.1.1 BDTs for Background Rejection

In the context of particle physics analyses, a typical task is discriminating between a signal and a background event. For such binary classification problems, the signal and background events are assigned a label of 1 and 0, respectively, and the BDT learns to transform a set of input parameters to a BDT score between 0 and 1. This is illustrated in Figure 5.2 using toy distributions. Using the labelled data, one can know the distribution of the BDT scores assigned to events from both the background and signal data and set a cut threshold above which events are accepted. Each threshold has an associated signal efficiency and accidental leakage as illustrated using the toy distributions in Figure 5.2 using two threshold values, t_1 (black dashed line) and t_2 (pink dashed line), with t_1 resulting in a higher signal efficiency than t_2 but at the cost of an increase in background leakage. The thresholds are set based on the signal efficiency and background leakage associated with that threshold and the requirements of the analysis.

5.1.2 Tunable Hyperparameters

The analysis presented here was using the python xgboost (version 1.7.5) package [109] which is a wrapper library around the underlying C++ implementation. XGBoost comes with a number of tunable hyperparameters (listed in Ref. [110]) that can be used to configure the model. The hyperparameters tuned in this study are described below, and the values used in the analysis are listed in Table 5.2.

- **Shrinkage (eta)**

  The eta or shrinkage is the weight applied to the output of a new tree added at each iteration. Setting a low shrinkage is a step towards mitigating against overfitting and making the boosting process more conservative. It is analogous to the learning rate or step size in a gradient descent model, and a lower eta requires a higher number of boosting rounds, i.e. a higher number of trees.

- **Max Depth**

  The maximum depth of a tree controls the number of input feature an individual tree makes cuts on to produce an output. This parameter increases the models complexity and a high value can lead to overfitting. The XGBoost default max depth is 6.

- **Number of Trees**

  This parameter sets the number of boosting rounds applied in which a new decision tree function is added to the model. Models with a large number of trees take longer to train, have a higher complexity and could overfit to the training data.

- **Column Subsampling Fraction**

  At each round of boosting the new tree is given a random subsample of the total features to make cuts on. This parameter sets the subsampling fraction. For example, a subsample of 0.5 gives a new tree half of the total available features selected through random sampling. This step mitigates against overfitting.

- **Row Subsampling Fraction**

  Similar to column subsampling, row subsampling is another measure undertaken to avoid overfitting. This parameter sets the fraction of events that form the subsample used for tree construction at each boosting round.

- **L1 & L2 Regularisation**

  These two parameters represent the regularisation constants in the loss func-

tion (Eq. 5.3). The first parameter γ penalises new trees that use a large number of leaves, and the second parameter λ adds a penalty to the loss function based on the sum of the output leaf values.

5.1.3 Train-Test Splitting and *k*-fold Cross Validation

To evaluate the performance of a machine learning model on unseen data, the total available events can be split into training and test data. The training data is used to iteratively train the model whereas the test data is used only to evaluate the performance. Since the test data is not used to calculate the loss function and update model parameters it does not directly influence model training. However, this approach does not allow us to utilise the entire data set when evaluating the signal efficiency and background leakage.

Another approach to machine learning training and evaluation is k-fold cross validation. In this procedure k models $\{m_0, m_1, m_2, ..., m_k\}$ are instantiated and the available data is split into k subsets $\{d_0, d_1, d_2, ..., d_k\}$. Each model m_i is trained on all data sets except d_i and is used to score subset d_i. After training all k models, each of the k subsets are scored resulting in all available data being assigned a score by a BDT trained on $k-1$ subsets of the data. With this, the signal efficiency and accidental leakage of a BDT trained on $k-1$ subsets of the data can be evaluated using all available events through k experiments rather than with a single experiment evaluated on a single subset (of size 1/k of all the available data).

The choice of k is largely determined by the total number of signal and background events available and set to ensure that there is enough training and validation data available in each fold. This study uses a 5-fold cross validation procedure to evaluate BDT performance, leading to 5 models each trained on four fifths of the total available data and used to score a quarter of the remaining unseen data.

5.2 BDT Based Approach to Rejecting Accidentals

5.2.1 Total Available Data

Three data sets were used in the BDT analysis, one for the background accidental events and two for signal ER and NR events. The background data set consisted

of Chopstitch accidental events which (as discussed in Sec 4.5.3 of the previous chapter) model the accidentals by "stitching" together isolated S1 and S2 events. Within the ROI ($3 <$ S1c <80 and $600 <$ S2c $<10^5$) this data set contains 4.6 million events, as shown in Figure 5.3. Since the data set is constructed by sampling from the isolated S1 and isolated S2 distributions observed in the SR1 data, we expect it to be representative of the types and rates of accidental coincidence events expected in the WIMP search data. The position distribution of the accidental events is uniform since the isolated S2 events are uniform in X, Y and the z position is determined from the time between the S2 and the randomly coincident isolated S1 pulse.

The signal class is populated by ER events obtained from CH_3T calibrations and NR events that are modelled following a similar methodology to the to Chopstitch accidentals. These Chopstich NR events are constructed using S2 pulses taken from the CH_3T data are sampled uniformly in X, Y and Z, and S1 pulses taken from the AmLi data that are chosen such that the Z position of the AmLi event corresponds to the Z position derived from the S2 width of the S2 it is being paired with. This leads to an available data set of positionally uniform NR like events that can be used to train the BDT in the NR region.

Hence both ER and NR events used in the training data have a uniform position distribution in the detector. The tritium is a dispersed source and leads to ER events that are uniformly distributed in the TPC, and the NR events are positionally uniform by construction.

Not all the available data is used in training the BDT. Only events that lie in the region in $\log_{10}$ S2c-S1c space where single scatters are expected are used, i.e. events in the ER and NR band. This selection procedure is described in the following sections.

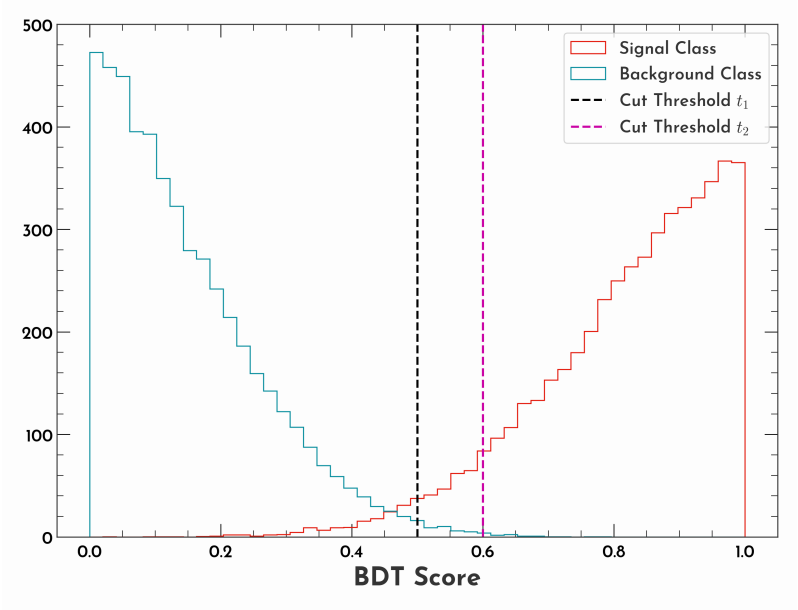


Figure 5.2: An illustration showing the output BDT score of signal and background events using toy distributions. The distribution of the BDT score of signal events is shown in red and the background is shown in blue. Two acceptance thresholds t_1 (black dashed line) and t_2 (pink dashed line) with different signal efficiency and accidental leakage are also shown.

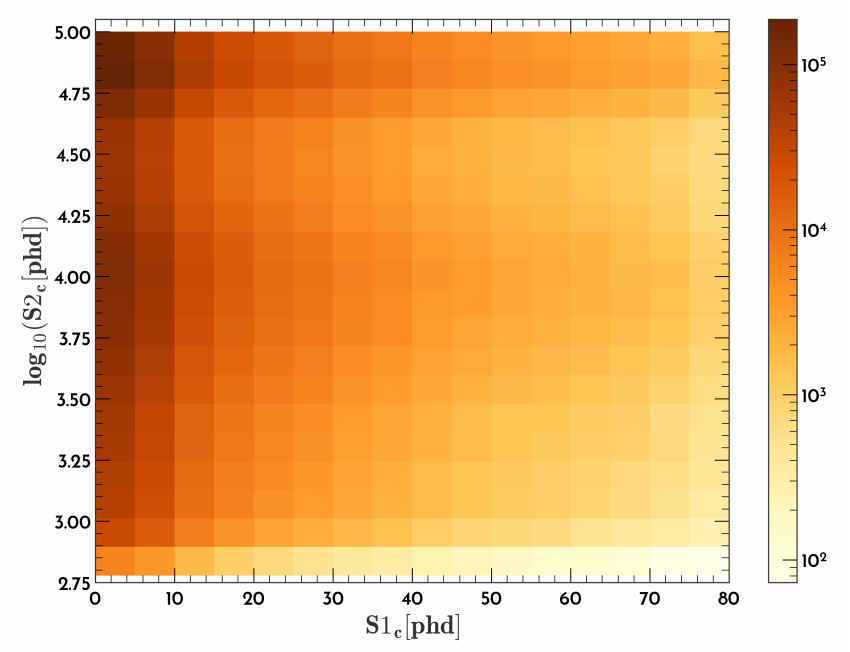


Figure 5.3: The distribution of the Chopstitch accidentals in the ROI in $\log_{10}$ S2c-S1c space.

5.2.2 Oveview of the BDT Analysis

The illustration in figure 5.4 provides an overview of the BDT based analysis procedure to test whether or not an event in the LZ data is produced by an accidental coincidence. Firstly, the livetime cuts discussed in Sec. 4.6.1 are applied to remove periods of data taking during which the TPC experienced elevated event rates. The second stage involves applying the physics cuts which include vetoeing events with

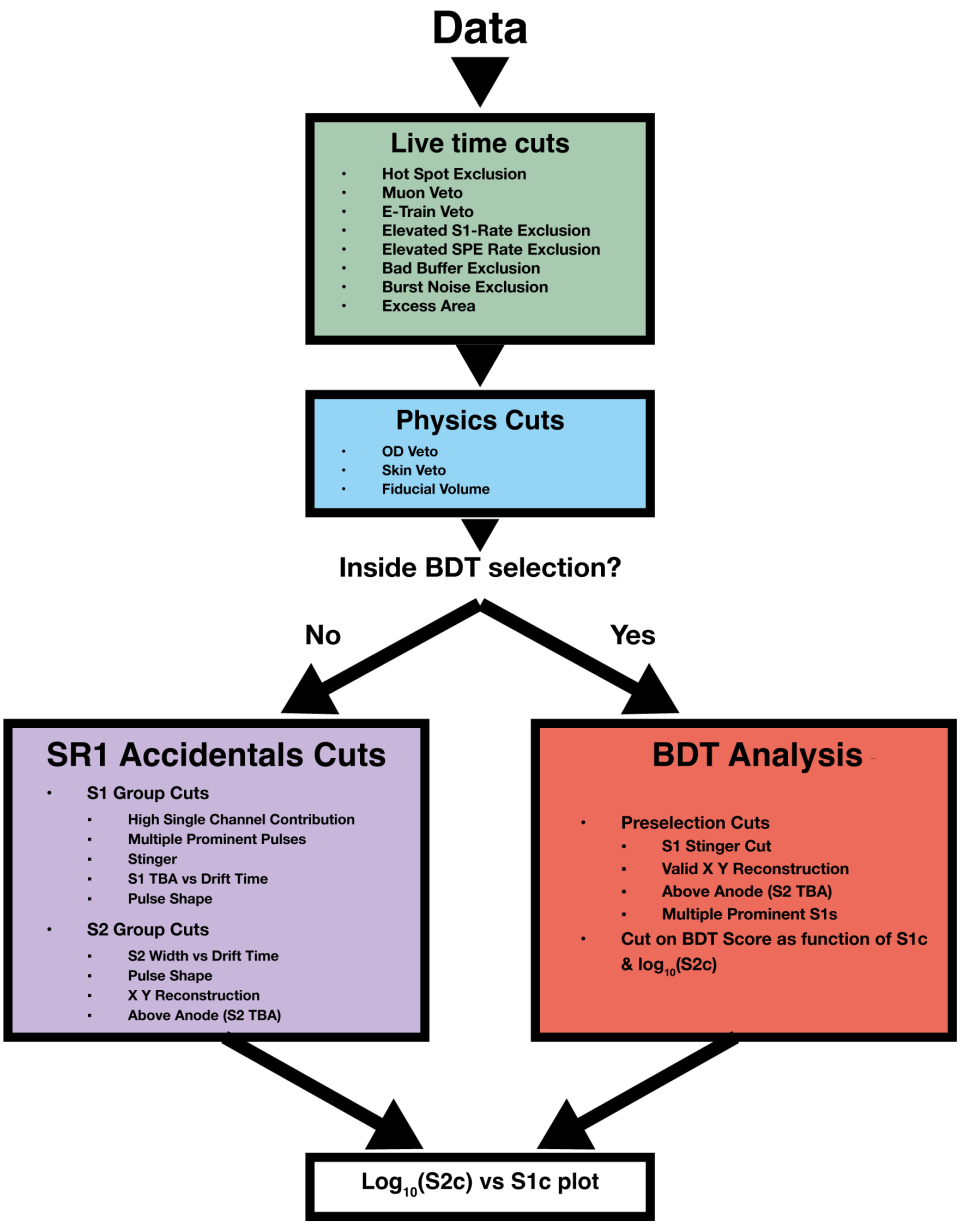


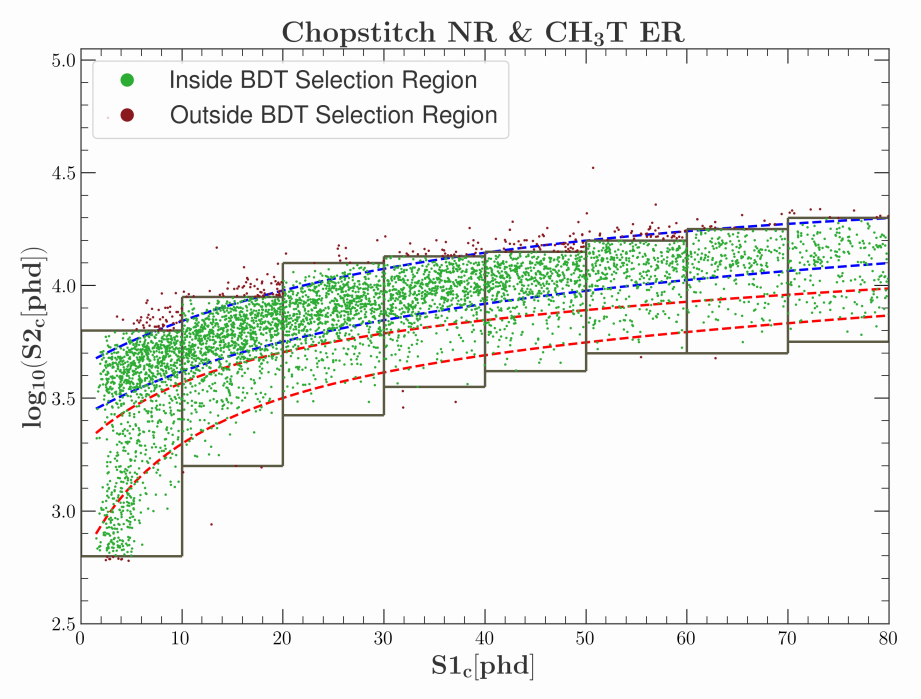
Figure 5.4: An illustration summarising all analysis steps applied to the SR1 WIMP-search data with accidentals rejected using the BDT procedure outlined in this chapter.

a coincident signal in the Skin or OD detectors, and applying the fiducial volume cut. Following from these cuts, events that fall within a *BDT selection region* (defined in Sec 5.2.3) are subjected to a set of preselection cuts (outlined in Sec 5.2.4). The remaining events are then assigned a BDT score using the k-fold procedure described in Sec 5.1.3. Events with a score below the acceptance threshold are then rejected. The events outside of the BDT selection region are accepted or rejected using the SR1 cuts. Therefore, all events that pass the physics cuts are classified, either using the BDT or the SR1 cuts.

5.2.3 BDT selection region

The task for the BDT is to learn to score events in a way that maximises the separation between real scatters in the TPC and accidental interactions. Accidentals are not real interactions and therefore there is no physical relationship between the amount of ionisation light and charge produced to create a relationship between the S1 and S2. Thus these events can occur anywhere in the entire ROI depending on the distribution of isolated S1 and isolated S2 events. However, in order to simplify the problem for the BDT, the classification task is confined to the region in $\log_{10}$S2c-S1c space where one expects to see ER or NR single scatter events. Events outside this region are classified using the standard SR1 cuts.

Figure 5.5a shows the combined single scatter training data (CH_3T and Chopstitch NR) in $\log_{10}$S2c-S1c space along with the ER and NR bands within which 90% of ER and NR events are expected to fall. Using these events the BDT region is defined as shown by the grey grid in Figure 5.5a and the signal data with green markers are used in the BDT training and evaluation. The Chopstitch accidental events that are used in the training are selected using the same criteria and are shown in Figure 4.5. This selection grid is later used to define bins in $\log_{10}$S2c-S1c space which are used to add sample weights such that the distribution of both signal and background is flat in $\log_{10}$S2c-S1c. This is discussed in section 5.2.5. With continued data taking increasing the number of events available for training, the coverage of the selection grid can be expanded to include all of the remaining Chopstitch NR and CH_3T calibration events currently outside the box in Figure 5.5a. Note, events



(a)

Figure 5.5: A plot in $\log_{10}$S2c-S1cspace illustrating the definition of the BDT selection region outlined by the grey grid. (a) shows CH_3T and Chopstitch NR with the green markers showing the signal events that are used to train the BDT and the maroon markers showing events outside the BDT selection region that are not used by the BDT.

outside this region are still reg

5.2.4 Preselection Cuts and Training Data

A subset of the SR1 cuts are used as preselection cuts and applied to the Chopstitch accidental events. These are the fiducial volume cut, the stinger cut, the above anode cut and the multiple prominent S1 cut. The fiducial volume cut is applied since this cut selects events that are in the quieter inner region of the detector where the WIMP search will be conducted, and where removal of accidentals is most critical. The stinger cut is applied as this requires information from the pulse preceding the S1 and adding this to the BDT would require additional features and complexity. Moreover the current implementation is highly effective at removing stinger events, therefore this cut is used in the preselection. Similarly the multiple prominent S1 cut also requires additional pulse information and was therefore used in the preselection. And lastly, the above anode cut was used in preselection as it is well studied

and targets a specific population of gas events. Moreover it was found that using this cut in the preselection led to improved overall performance.

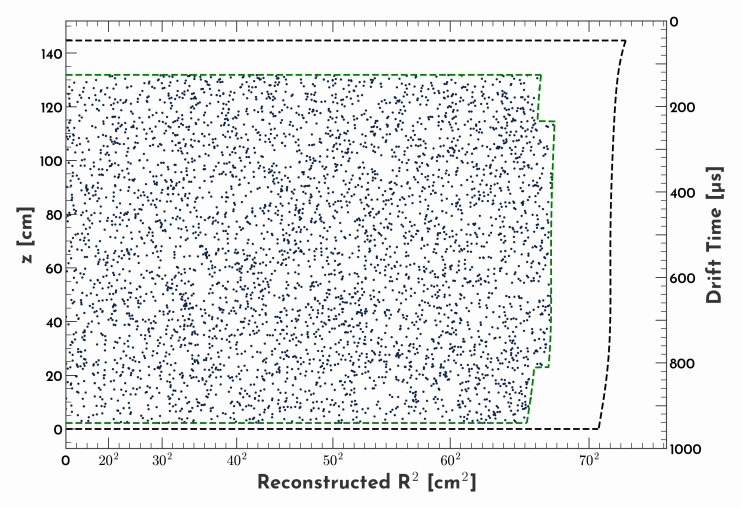
Out of the 3.8 million Chopstitch accidental events in the entire ROI, 186,000 events lie within the BDT selection region and pass the above preselection cuts. These form the background events used in the training data along with signal events consisting of 4,291 CH_3T events and 964 Chopstitch NR events. Figure 5.6 shows the positional distribution of these events and demonstrates that the training data has uniform coverage in the TPC and is therefore unlikely to contribute to a positional bias.

The same cannot however be said about the distributions of the training data in $\log_{10}$S2c-S1c space, which if left unchanged during training can lead to biases in the classification process. Furthermore the imbalance in the number of signal events compared with background is also likely to cause biases. These potential issues are mitigated against through a 2D reweighting procedure described below.

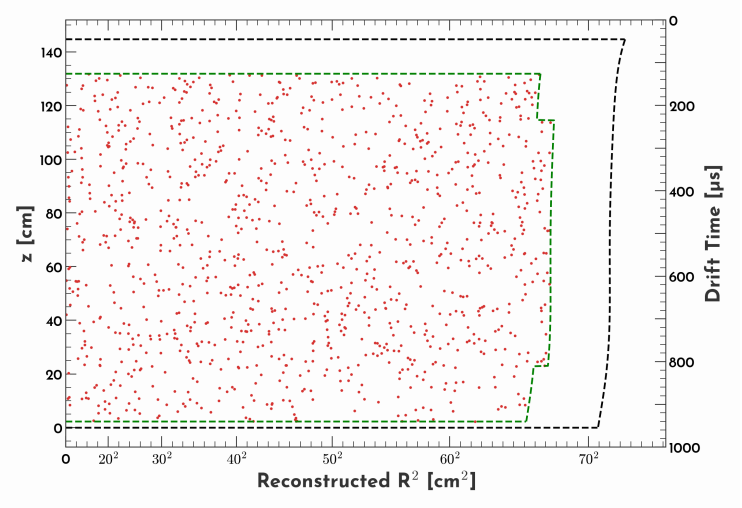
5.2.5 Sample Weights and Bias Mitigation

The objective of the BDT is to classify events based on selected input parameters in a way that is agnostic to the size of the S1 or S2 pulses in the event. Given that the accidental distribution peaks near low S1c, if the BDT were to be given information on the S1 size it could maximise performance by assigning a much lower BDT score to events in the low energy (e.g. $S1c < 20$) region thereby leading to sub-optimal signal efficiencies in this region and higher accidental leakage in the higher energy regions. Even though neither S1c nor S2c are parsed to the BDT, this information is correlated with a number of the input parameters that are given to the BDT and therefore affects the distributions of these parameters. This is illustrated in Figure 5.7 which shows the distribution of the S1 full width at half maximum (FWHM) divided by S1c for a set of events with a non-uniform and uniform distribution in $\log_{10}$S2c-S1c space. Given the differences in the shapes of the uniform and non-uniform distributions of the input feature, it is clear that the distribution of the events in $\log_{10}$S2c-S1c would influence the training of the BDT.

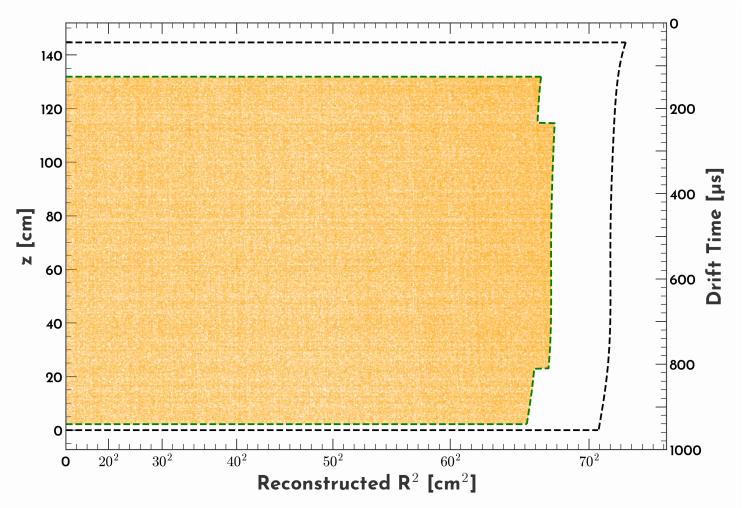
The weighting procedure is visualised in Figure 5.8, which shows the weighted



(a)



(b)



(c)

Figure 5.6: The positional distribution in reconstructed R^2 and z for the signal and background events used as training data for the BDT. (a) shows the CH_3T, (b) shows the Chopstitch NR, and (c) shows the Chopstitch accidentals. The dashed black line shows the extent of the active volume and the dashed green line shows the fiducial volume definition.

and unweighted distributions of the signal and accidental events, respectively. The sample weights to achieve a flat distribution in $\log_{10}$S2c-S1c are calculated by first binning the data in 10 phd bins in S1c. The data in each S1c bin is then binned further in 4 equally spaced bins in S2c with the minimum and maximum $\log_{10}$S2c defined by the BDT selection region. A weight is then assigned to events in each bin such that the weighted sum of events in each bin is constant across bins. This results in a weighted distribution that is flat in $\log_{10}$S2c-S1c space within the chosen binning, with all bin heights being equal. The signal (CH_3T and Chopstitch NR) and background (Chopstitch accidentals) data are binned and weighted separately to ensure that the weighted number of signal and background events passed to the BDT are equal as a step towards mitigating effects resulting from the imbalance in signal and background events.

5.2.6 Input Parameters

The input parameters used in the BDT are listed in Table 5.1 and are a combination of pulse and event level RQs. For each input RQ, the corresponding SR1 cut (as discussed in Sec. 4.6) is listed. All bar one the input parameters used in the BDT are also used in the official SR1 accidentals cuts. These RQs were chosen because they are physically motivated and were validated in the SR1 analysis through extensive ”hand scanning“ of isolated S1 and isolated S2 events using the offline event display. The addition of the S2 pulse shape parameter not used in the SR1 analysis (S2 AFT25 - AFT5) was found to improve the performance of the BDT, possibly because it compliments the RQ associated with the S2 rise time (AFT5 - AFT1) in targeting accidental events with isolated S2 pulses originating from interactions in the gas region (as discussed in Sec. 4.5.2).

In conventional machine learning terms this process of identifying the right RQs and validating them through handscanning can be described as feature engineering grounded in domain knowledge. As opposed to a purely data driven approach which in the extreme may involve using all available RQs as input parameters, this process deploys the BDT as less of a ”Black Box” classifier and more of an optimiser of existing physically motivated analysis cuts.

Input RQ	Associated SR1 Cut
$\frac{\text{S2 AFT95} - \text{AFT5 [ns]}}{\text{Drift Time [µs]}}$	S2 Width vs Drift Time
$\frac{\text{S2 RMS Width [ns]}}{\log_{10}(\text{S2c [phd]})}$	S2 RMS
$\frac{\text{S2 FWHM [ns]}}{\log_{10}(\text{S2c [phd]})}$	S2 FWHM
$\frac{\text{S2 AFT5} - \text{AFT1}}{\log_{10}(\text{S2c [phd]})}$	S2 Rise Time
$\frac{\text{S1 TBA}}{\text{Drift Time [ms]}}$	S1 TBA vs Drift Time
$\frac{\text{S1 Max Channel Area [phd]}}{\text{S1 [phd]}}$	S1 High Single Channel
$\frac{\text{S1 FWHM [ns]}}{\text{S1c [phd]}}$	S1 FWHM
$\frac{\text{S1 Area at 50ns}}{\text{S1c[phd]}}$	S1 Prompt Fraction 50ns
S1 Channel Peak Time RMS	S1 Photon Timing
$\frac{\text{S2 AFT25} - \text{AFT5}}{\log_{10}(\text{S2c [phd]})}$	Not used in SR1

Table 5.1: List of the RQs used as input parameters to train the BDT model and the associated SR1 cut that uses the same RQs.

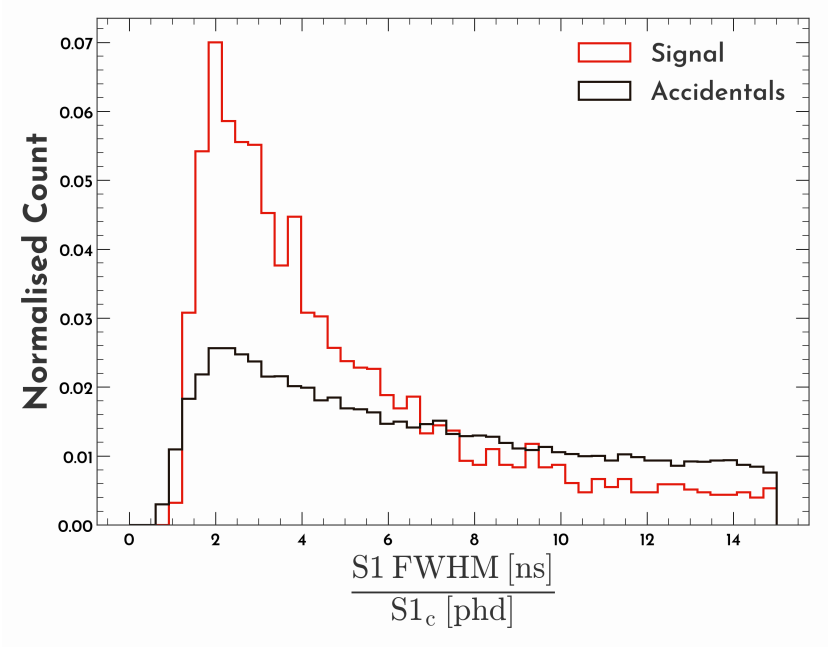
Figure 5.9 and Figure 5.10 shows the distributions of these RQs for single scatters and accidental events. Quantities that are correlated with S1c or S2c are given as a fraction of the pulse size. The distinctions in the distributions are the result of differences in the underlying physics that produce the isolated S1 and isolated S2 events. However some of these difference are not easily seen in the 1D distributions. For instance, the relationship between S2 width and drift time for real single scatter events is governed by the diffusion of the electron cloud as it travels through the detector. This can be seen in Figure 5.11 which plots the S2 Width as measured by

AFT95 - AFT5 against the drift time for both accidental events and ER events from CH_3T data. The correlation between the S2 width and drift time can be seen here, as well as the uniformity in drift time distribution for accidental events with narrow S2 widths. The isolated S2 pulses contributing to events with a narrow S2 occur at the top of the detector, either in the gas or near the liquid surface. This is a region of the detector with lower S1 light collection efficiency leading to events where the S2 is measured but the S1 is lost. Such events are targeted using the S2 width over drift time RQ and more importantly RQs with information about the rise time of the S2 pulse (AFT5 - AFT1, AFT25 - AFT5, and FWHM).

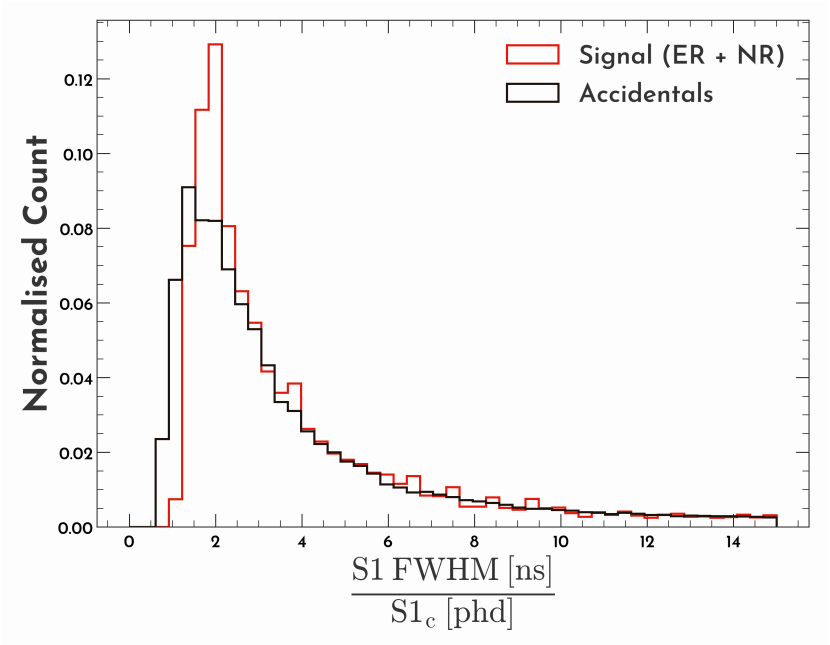
Similarly there is a population of events with wide S2 pulses that have S2 widths characteristic of events that are close to the bottom of the detector where Radon daughters from the cathode and gate grid contribute to the isolated S2 rate.

As such, whilst the 1D visualisations give an indication of the separation power of a variable the input parameters are selected through experimentation and included based on the effect on the final performance as measured by the area under the receiver operating characteristic (ROC) curve (see Sec 5.3.2). In Figure 5.9a it appears that the FWHM of the S1 pulse does not have much separation power and would therefore not be a useful variable to include in the analysis, however it was found that removing this parameter led to a reduction in performance.

The performance of the BDT including an analysis of feature importance is discussed in Sec 5.3.1.



(a)



(b)

Figure 5.7: Figure illustrating the differences in the distributions of a pulse RQ based on the shape of the distribution in $\log_{10}$S2c-S1c space. (a) shows the distribution of the full width at half maximum of the S1 divided by S1c for events with a non-uniform distribution. (b) shows the same for events with sample weights assigned such that the weighted distribution in $\log_{10}$S2c-S1c space is flat.

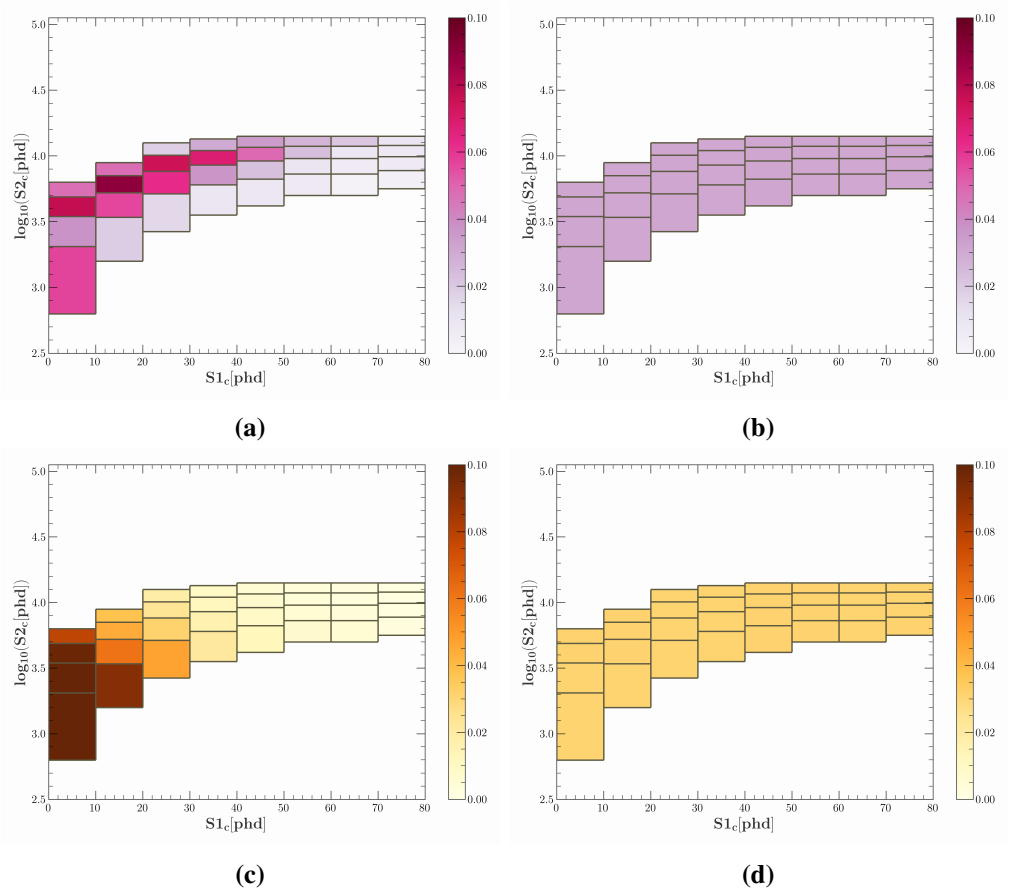


Figure 5.8: The weighted and unweighted distributions of the training data in $\log_{10}$ S2c-S1c space. Figures (a) and (c) show the unweighted distributions of the signal and accidental events, respectively. Figures (b) and (d) show the resulting flat distributions for signal and accidental events, respectively created using the weights assigned to the events.

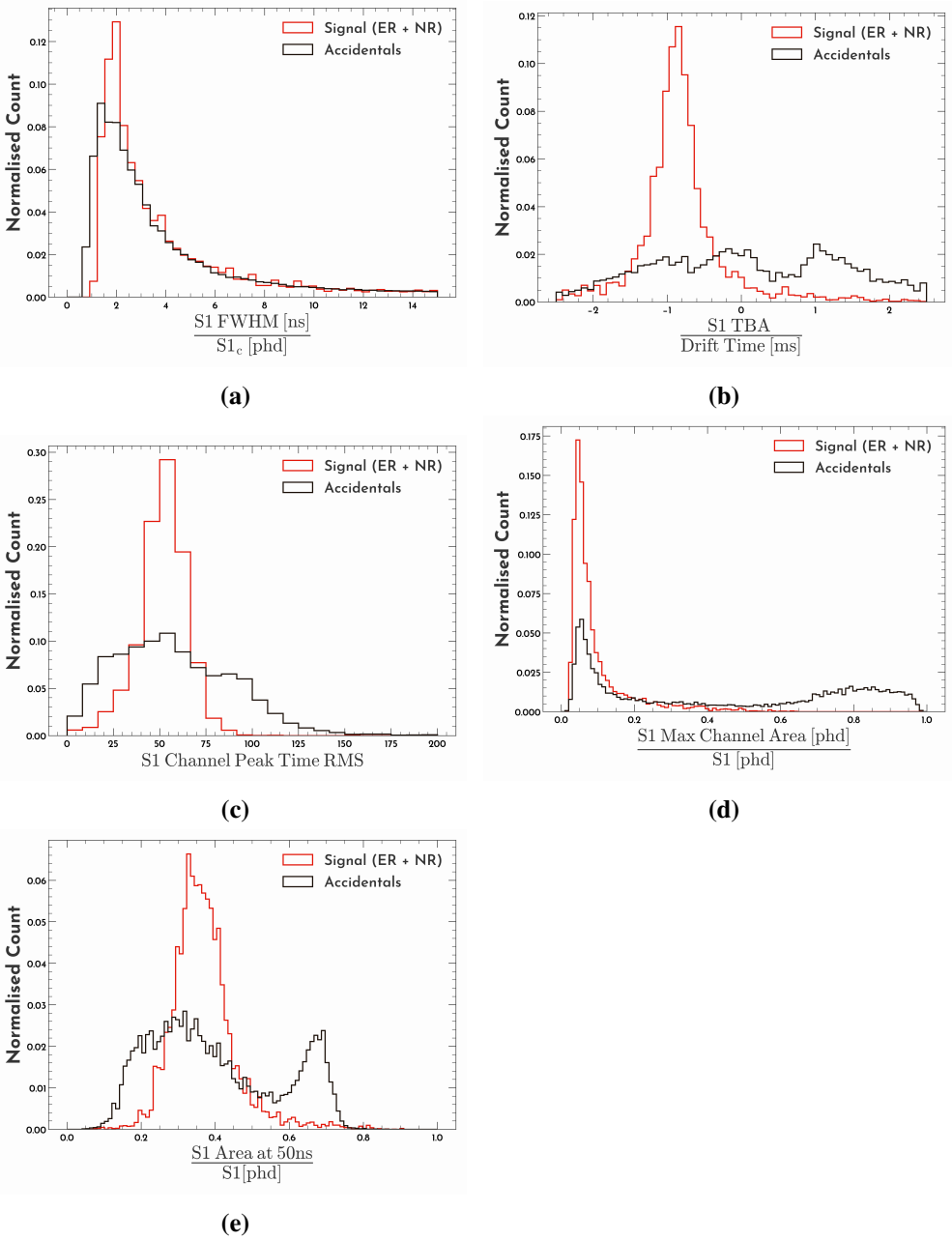


Figure 5.9: Distributions of the input parameters derived from the S1 pulse from both signal and Chopstitch accidental events. The distributions are constructed using the sample weights assigned such that the events have a flat distribution in $\log_{10}$ S2c-S1c space.

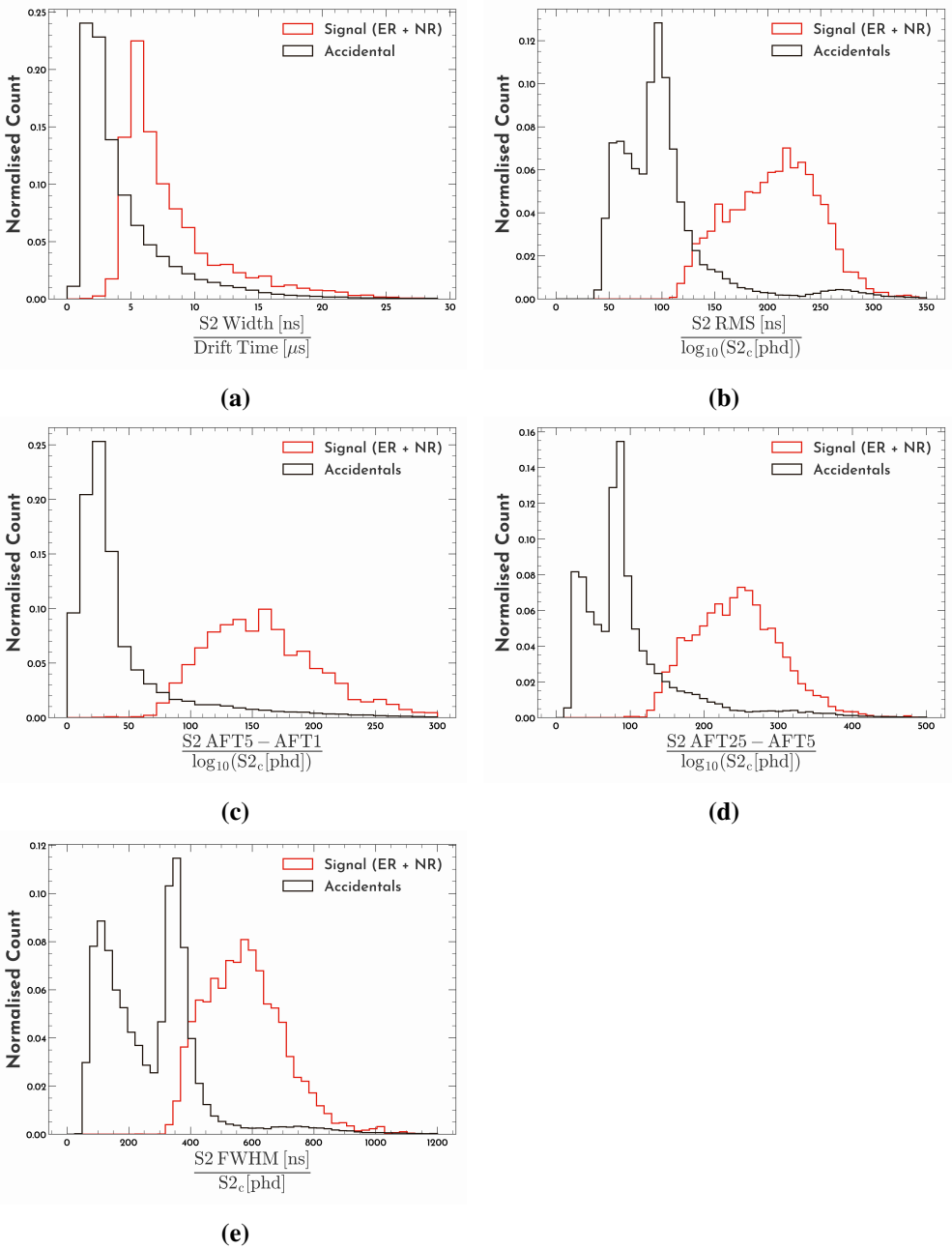


Figure 5.10: Distributions of the input parameters derived from the S2 pulse from both signal and Chopstitch accidental events. The distributions are constructed using the sample weights assigned such that the events have a flat distribution in $\log_{10}$ S2c-S1c space.

XGBoost Hyperparameter	Value Used
Number of Trees	500
Max Tree Depth	7
Shrinkage	0.01
L1 Regularisation	0.4
L2 Regularisation	0.4
Row Subsample	0.8
Column Subsample	0.8

Table 5.2: List of the hyperparameters and their values used when training the BDT models.

5.3 BDT Training

With the RQs listed in Table 5.1, the model is then built using a 5-fold cross validation procedure to with 1 or 0 as the target variable for signal and accidental events, respectively. The model therefore consists of five XGBoost BDTs with each one trained on four fifths of the total available data and used to score the remaining fifth of the events. The hyperparameters used in this study are listed in Table 5.2 and were set using a grid search over the number of trees and max tree depth. The rest of the hyperparameters were heuristically set to conventionally suitable values. Although further hyperparameter tuning can be done to improve the performance, the results presented below show that SR1 analysis can be improved upon even with this relatively simple BDT set up. Furthermore, it was found that the BDT performance was more sensitive to changes in the input features and data set size than the hyper parameters suggesting that higher gains in performance are likely to be found by engineering new features and adding more events to the signal class rather than through extensive hyperparameter tuning.

Figure 5.12 shows the probability distribution of the BDT score for the signal and accidental data. Accidental events are predominantly assigned a score closer to 0 and single scatters are given scores closer to 1. The overlap in the two distributions contributes to the accidental leakage and loss in signal efficiency.

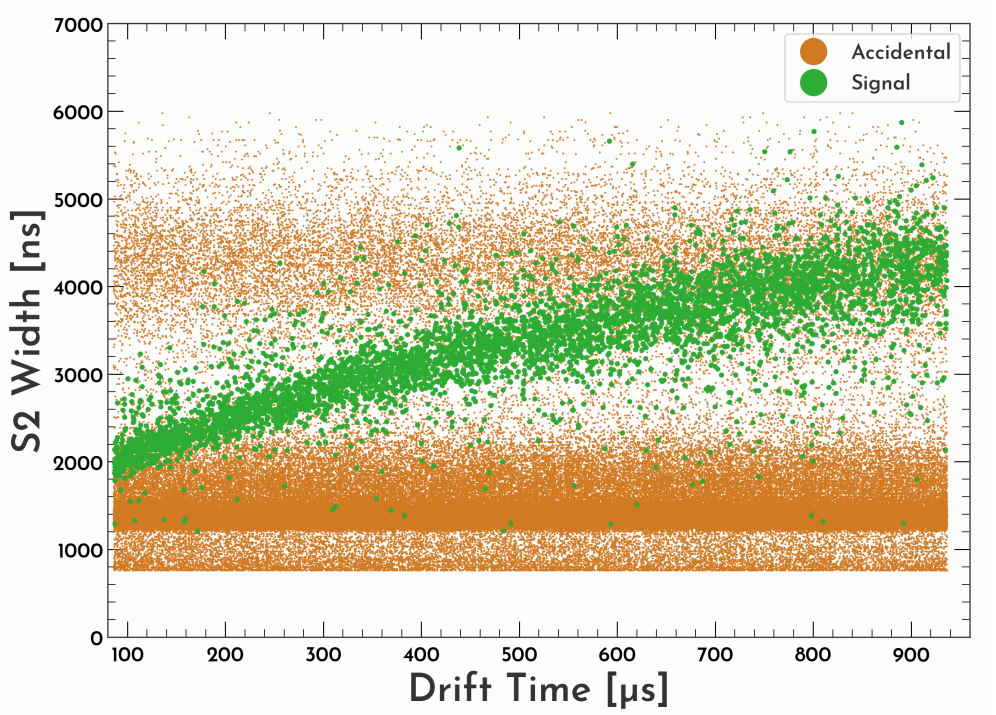


Figure 5.11: S2 Width defined as AFT95 - AFT5 plotted against drift time for accidental and signal (CH_3T + Chopstitch NR) events. The signal events show a correlation between the two parameters due to the diffusion of the electron cloud as it is transported through the detector. No such correlation exists for accidental coincidence events.

5.3.1 Feature Importance

Three metrics were used to assess the importance of the input features used by the BDTs. These were the the information gain, the weight and the coverage. Figure 5.13 shows the average feature importance for the 5 BDT models measured using the three metrics as well as the standard deviation in the importance score between models. The weight measures the frequency with which a feature is used to make splits across all trees. The information gain is the reduction in the loss function that is achieved by splitting on a given variable averaged across all trees. The coverage measures the number of events that are split using a given variable across all trees.

Figure 5.13b shows that the rise time of the S2 has a significantly higher gain and is an incredibly powerful parameter to reject accidentals. The separating power can also be seen in Figure 5.10c which shows that high proportion of accidental events have sharply rising S2 pulses. This is due to the large contribution of isolated S2

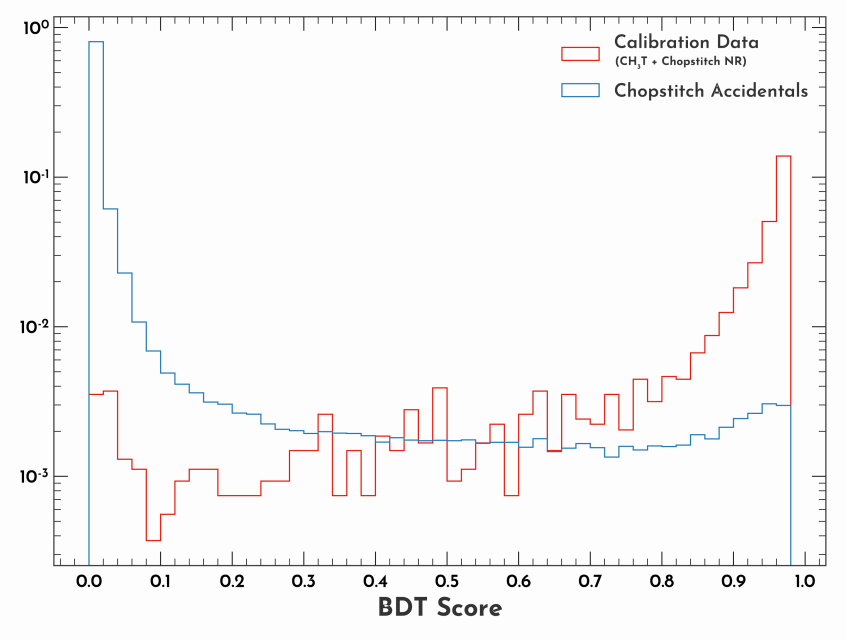


Figure 5.12: The probability distributions of the BDT scores assigned to the labelled training data. The distribution of the accidental events is shown in blue and the single scatter calibration data is shown in red. Note the log scale on the y-axis.

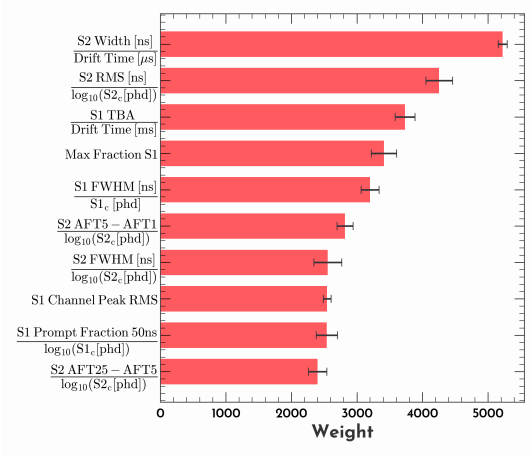
signals near the liquid surface and gas region of the TPC.

Figures 5.13a & 5.13c show that although the other parameters do not have high information gain relative to the rise time parameter they are still used in tree construction. The S2 width vs drift time RQ (Figure 5.10a) for instance is the feature with the highest weight i.e. it is used to make splits in the most trees. This is expected as the S2 width and the drift time are correlated for real scatters but uniform random for accidentals as shown in Figure 5.11.

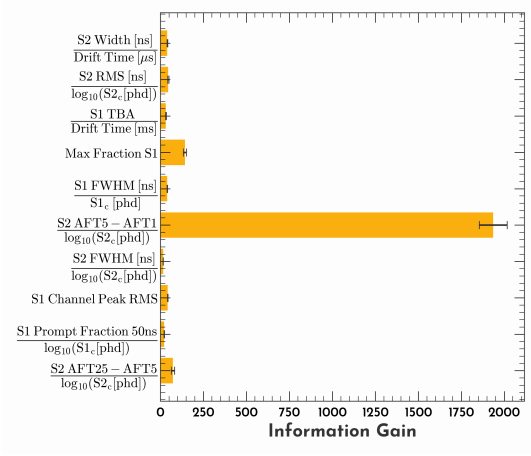
Features with the lowest weight were also temporarily removed to assess whether this removal had an impact on the performance. Removing any of the parameters listed in Table 5.1 led to a worse performance as measured by the area under the curve (AUC) for the WIMP ROI ROC curves shown in Figure 5.15 and discussed in the next section.

5.3.2 Evaluation and Classification

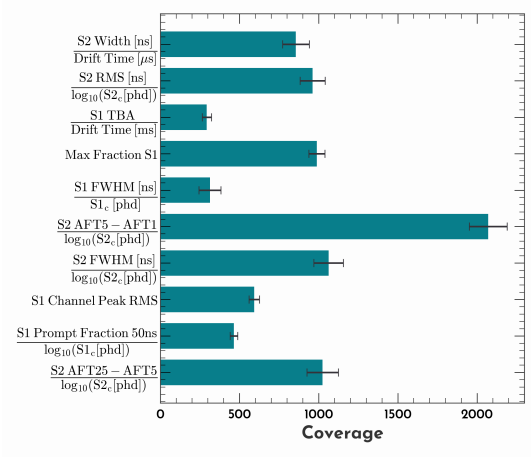
With the full labelled data now scored, an appropriate acceptance threshold must be selected to classify events. This is selected by creating an ROC which shows



(a)



(b)



(c)

Figure 5.13: Feature importance for the input parameters used by the BDT analysis measured using three metrics.(a) shows the *weight* of each feature which measures the number of times a feature is used to split across all trees. (b) shows the *information gain* which measures the loss function reduction achieved by a given feature. And (c) shows the coverage which measures the number of events that are split using a given variables across all trees.

the accidental leakage and signal acceptance for range of threshold values. In this study, rather than using a single threshold value t the threshold is set as a function $t(S1c, \log_{10}\text{S2c})$, resulting in adjustable signal efficiency and accidental leakage maps. The threshold map $t(S1c, \log_{10}\text{S2c})$ is set by calculating ROC curves for events in each $(S1c, \log_{10}\text{S2c})$ bin in the binning map shown in Figure 5.14, where the same binning as the weighting procedure was chosen. Figure 5.15 shows three ROC curves corresponding to the three bins highlighted in cyan in Figure 5.14 that cover most of the $30\,\text{GeV}/\text{c}^2$ WIMP ROI. The signal efficiency and accidental leakage achieved for events in the selected bin using the SR1 analysis is also plotted for comparison. The ROC curves show that the BDT outperforms the SR1 analysis since there exists a threshold for which the BDT can achieve a lower (higher) accidental leakage (signal efficiency) for the signal efficiency (accidental leakage) achieved by the SR1 analysis.

Additionally, the BDT based classifications allows analysers to easily lower the accidental leakage as necessary (at the cost of signal efficiency) as it involves simply tuning on threshold $t(S1c, \log_{10}\text{S2c})$. Similar adjustments to the SR1 analysis is a much longer process that requires tuning the 10 individual cuts. If upon further investigation of the accidentals a new RQ or feature is engineered that can improve the rejection, it can be added as an input parameter to the BDT rather than adding to a long list of analysis cuts.

The ROC curves in Figure 5.15a also indicate that accidental rejection is hardest for near threshold events with $3 < \text{S1c[phd]} < 10$ and $2.8 < \log_{10}(\text{S2c[phd]}) < 3.8$ and the performance starts to improve with S1 size. As the performance improves the area under the ROC curves increases as in Figure 5.15b and 5.15c. It is possible that the relatively low number of NR events (964) compared with the ER (4291) is a contributing factor to sub-optimal performance seen in the NR band near threshold. This should improve in the next run (SR2) as more AmLi calibration data will be taken allowing one to generate additional Chopstitch NR events.

The results presented in this study use thresholds that were set to achieve an approximately four fold reduction in the accidentals leakage compared to the SR1

analysis. This demonstrates the ability to reduce the accidentals rate to an acceptable level for the longer exposures planned for future science runs.

5.3.3 Binomial Proportion Estimation

Using the BDT classifications and the true class labels for data, the accidental leakage and signal efficiency for each bin in $S1c$ and $\log_{10}$ S2c is estimated within a 68% confidence interval. The number of signal (accidental) events that are accepted k_s (k_a) out of a total N_s (N_a) signal (accidental) events is binomially distributed with the success probability given by the signal efficiency ε_s (accidental leakage ε_b). The success probability ε is simply given by k/n, however the range $[\varepsilon_{low}, \varepsilon_{high}]$ within which the true value lies with a 68% probability is estimated using the *Clopper Pearson Exact Method* [111]. This method uses the cumulative distribution function (CDF) of the binomial distribution $I_{1-\varepsilon}(n-k, k+1)$ where $I_q(a,b)$ is the regularised incomplete beta distribution, to calculate the confidence interval. For an observed number of accepted events k out of total events n, the upper limit on the

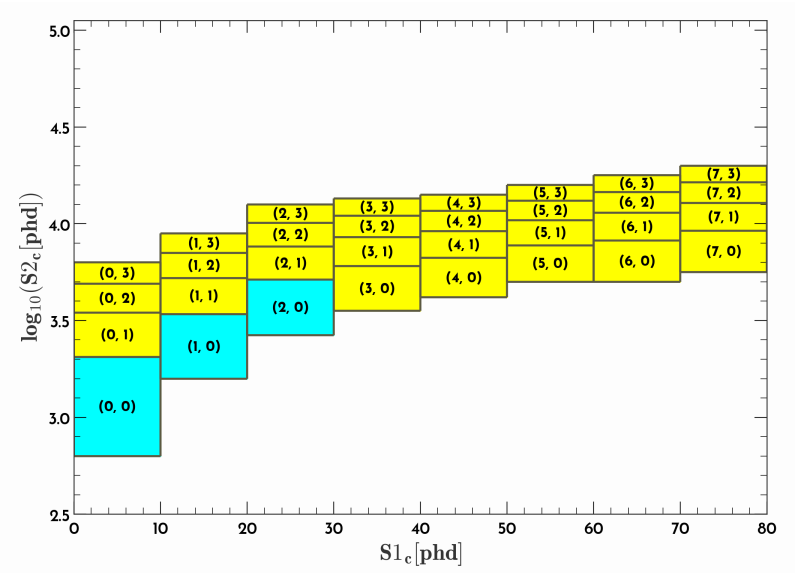
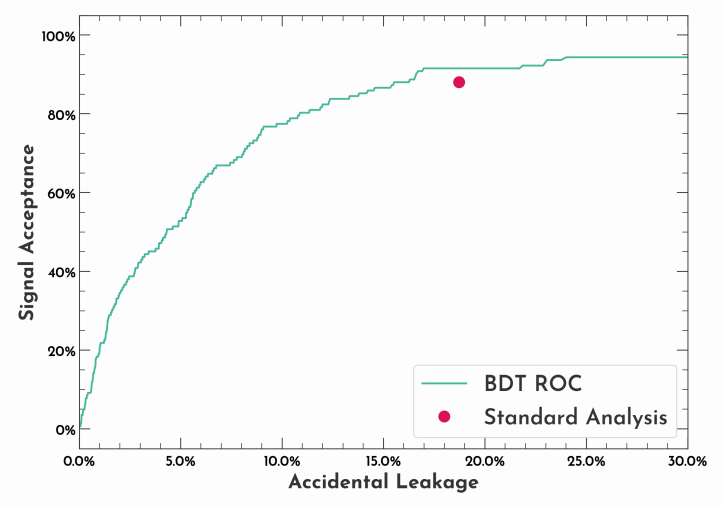
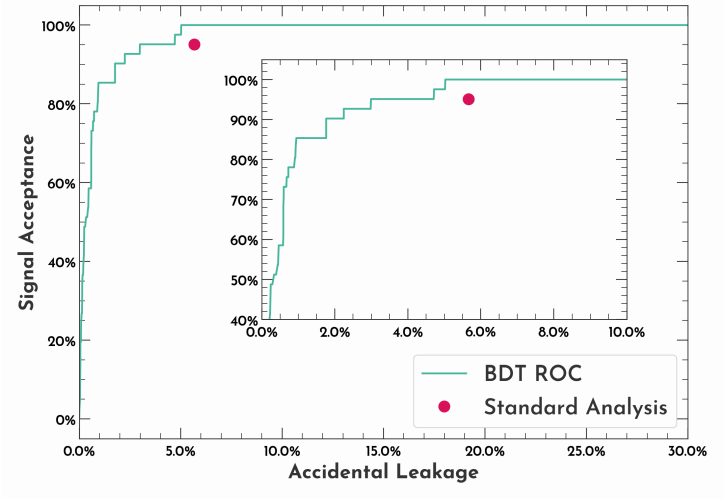


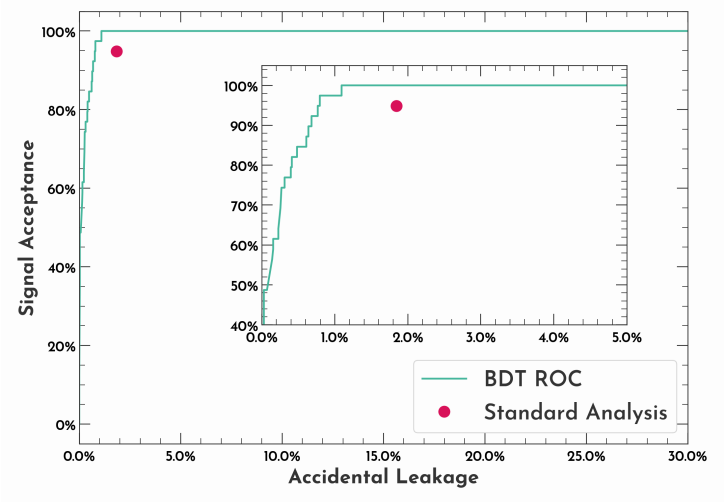
Figure 5.14: A map showing the binning used to evaluate the ROC curves and set the BDT acceptance thresholds. The ROC curves for the three bins highlighted in cyan are shown in Figure 5.15.



(a) S1c[phd] < 10 and $2.8 < \log_{10}(\text{S2c[phd]}) < 3.8$.



(b) $10 <$ S1c[phd] < 20 and $3.2 < \log_{10}(\text{S2c[phd]}) < 3.85$.



(c) $20 <$ S1c[phd] < 30 and $3.4 < \log_{10}(\text{S2c[phd]}) < 4.1$.

Figure 5.15: ROC curves from the three bins highlighted in cyan in Figure 5.14. The orange marker shows the signal efficiency and accidental leakage achieved by the SR1 analysis. The inset figures in (b) and (c) show the ROC curves zoomed in near the SR1 analysis value. (a) shows the (0,0) bin; (b) shows the (1,0) bin ; and (c) shows the (2,0) bin.

acceptance probability ε_{high} satisfies:

$$I_{1-\varepsilon_{high}}(n-k,k+1) = \alpha/2 = 0.16; \;\; \alpha = 1-0.68 \tag{5.4}$$

And similarly the lower limit ε_{low} satisfies:

$$1 - I_{1-\varepsilon_{low}}(n-k,k+1) = 0.16 \tag{5.5}$$

Since a CDF $F_K(k;\varepsilon)$ of a binomially distributed variable K gives the probability $P(K \leq k)$, in Eq 5.4, for the upper limit of the binomial success probability ε_{high} we are looking for a value such that the probability that at most k events are accepted out of n is 0.68. Similarly, in Eq 5.5 one minus the CDF gives the probability $P(K > k)$ so for the lower limit we looking for ε_{low} such that the probability that at least k events are selected is 0.68.

Using this method ε_s and ε_a for each bin are calculated resulting in 2D signal efficiency and accidental leakage maps $\varepsilon_s(S1c, \log_{10}(S2c))$ and $\varepsilon_a(S1c, \log_{10}(S2c))$, respectively. The upper and lower limits are also calculated and later used to estimate the uncertainty in the number of accidental events remaining after all cuts are applied.

5.4 Accidental Leakage and Signal Efficiency

Taking the initial distribution of Chopstitch accidental events within the ROI normalised to the number of UDT events in the same region (as shown in Figure 5.16) and then applying the accidental leakage to each bin we arrive at an estimate for the distribution of accidental events with the BDT analysis. The post cut distributions of accidental events for both the SR1 analysis and the BDT analysis are shown in Figure 5.17. Figure 5.18 is a visualisation of the reduction in the expected number of accidental events and plots the expected accidental distribution after the BDT analysis as a fraction of the expected accidental distribution obtained using the SR1 analysis. An approximately four-fold reduction in the expected number of accidental events in the single scatter region is observed and as expected, no change is seen

in bins outside the BDT selection region as events in this region are classified using the SR1 cuts in both approaches. The reduction in the accidental rate of course comes with a reduction in signal efficiency.

The signal efficiency maps for the SR1 and BDT based analysis are shown in Figure 5.19, where the BDT efficiency divided by the SR1 efficiency shown in Figure 5.20. This shows a $\sim 60\%$ reduction in the signal efficiency compared to the SR1 analysis for events near the detector threshold and much a lower reduction in the signal efficiency for events in bins corresponding to higher energy NR interactions.

Using these accidental distributions and signal efficiency maps the number of accidentals expected in the entire ROI and within the NR band for a 60 day and 1000 day run was calculated. These results are shown in Table 5.3 along with values obtained by using the SR1 analysis.

The number of accidental events expected in a 60 day run is reduced to 0.54 ± 0.3 in the full ROI and 0.09 ± 0.05 inside the NR band. Compared to 1.2 ± 0.3 and 0.36 ± 0.1 remaining using the SR1 analysis in the full ROI and NR Band, respectively. This represents a four-fold reduction in the expected number of accidental events in the WIMP-ROI with a $\sim 20\%$ average reduction in signal efficiency in the NR band. Moreover, this brings the expected number of accidental events for a 1000 day run much closer to the requirement of $\mathcal{O}(1)$ event for the experiment to reach its projected sensitivity.

The improvement is higher in the single scatter region for ERs and NRs compared to the full ROI since events outside the BDT selection region are classified using the SR1 cuts in the BDT analysis and therefore no change to the number of accidental events is expected. Although there is a reduction in the signal efficiency to reach the four fold improvement, the advantage of using the BDT analysis lies in its ability to both leverage the power of a multivariate classification as well enabling analysers to easily tune the acceptance threshold $t(\log_{10}\text{S2c-S1c})$ to set the acceptable accidental leakage.

5.4.1 Sources of Uncertainty

There are five sources of uncertainty in the estimated number of accidentals. These are: the binomial uncertainty on the accidental leakage, the poisson uncertainty on the number of UDT events, systematic uncertainty in the congruence between the Chopstitch accidental events and the UDT accidental events, systematic uncertainty between the UDT accidental events and the physical drift time accidental events, and systematic uncertainties from the lack of accurate S1 and S2 corrections in accidental events. These individual uncertainties are estimated and added in quadrature to give the uncertainties in Table 5.3.

Analysis	Number of Events		Average NR Efficiency
	60 Days	1000 Days	
SR1 (Full ROI)	1.2 $\pm$0.3	19.9 $\pm$5	92.3 $\pm$ 0.2%
SR1 (Inside NR Band)	0.36 $\pm$0.1	6.00$\pm$1.5	
BDT (Full ROI)	0.54 $\pm$0.3	8.93 $\pm$5	73 $\pm$ 0.2%
BDT (Inside NR Band)	0.09 $\pm$0.05	1.45 $\pm$0.8	

Table 5.3: Expected number of accidentals in the SR1 WIMP-search data for 60 and 1000 live days using the BDT analysis and the SR1 analysis. The expected NR efficiency for both analyses is also shown.

5.5 BDT Analysis on the SR1 Data

5.5.1 BDT Analysis on the WIMP Search Data

Using the trained BDT models and the classification procedure outlined, the WIMP-search data was analysed. Figure 5.21 shows events remaining in the fiducial volume after all livetime and physics cuts are applied, as well as the BDT preselection (see Sec 5.2.4) cuts which are common to both the BDT and SR1 analysis. These are: the stinger cut, the above anode cut and the XY reconstruction cut. A total of 429 events remained and events within the BDT region were assigned the average score of the 5 BDTs. These events were accepted or rejected using the thresholds $t(S1c, \log_{10}\text{S2c})$, and events outside this region were selected using the SR1 cuts. The positions of these events in the TPC are shown in Figure 5.22.

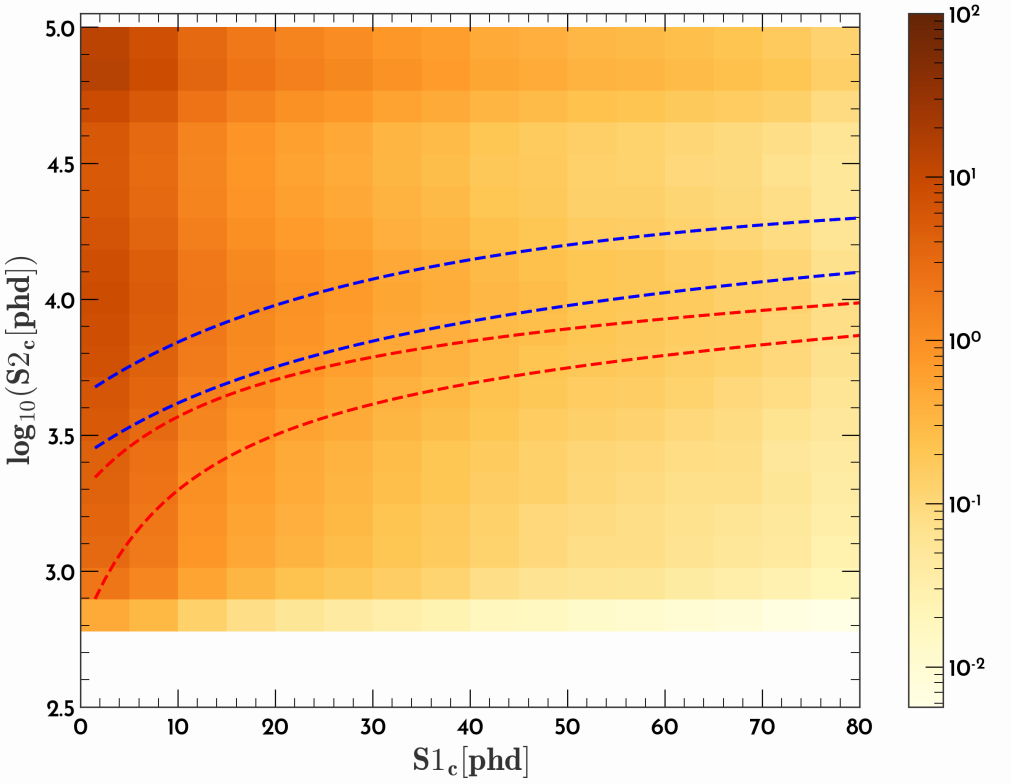
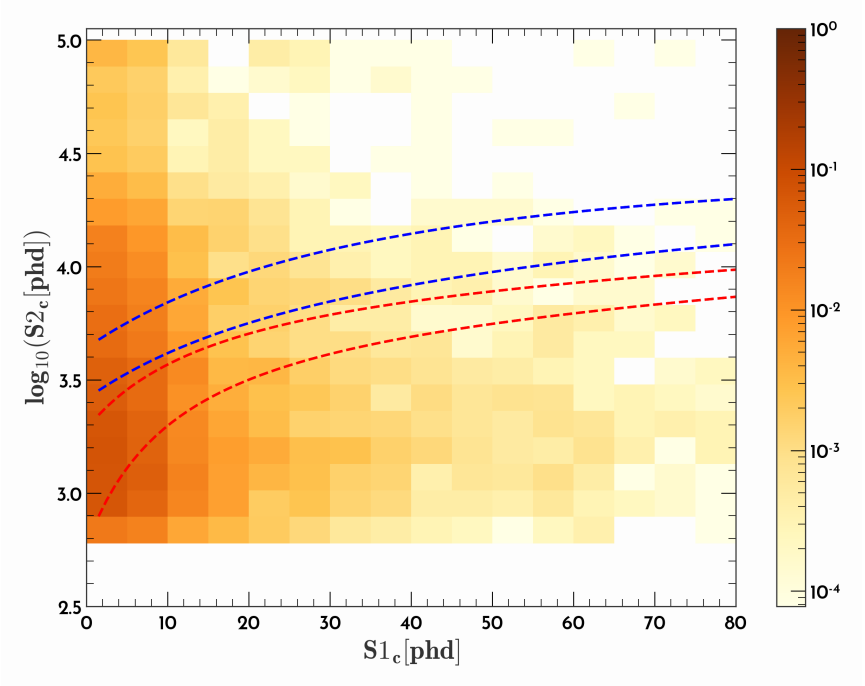


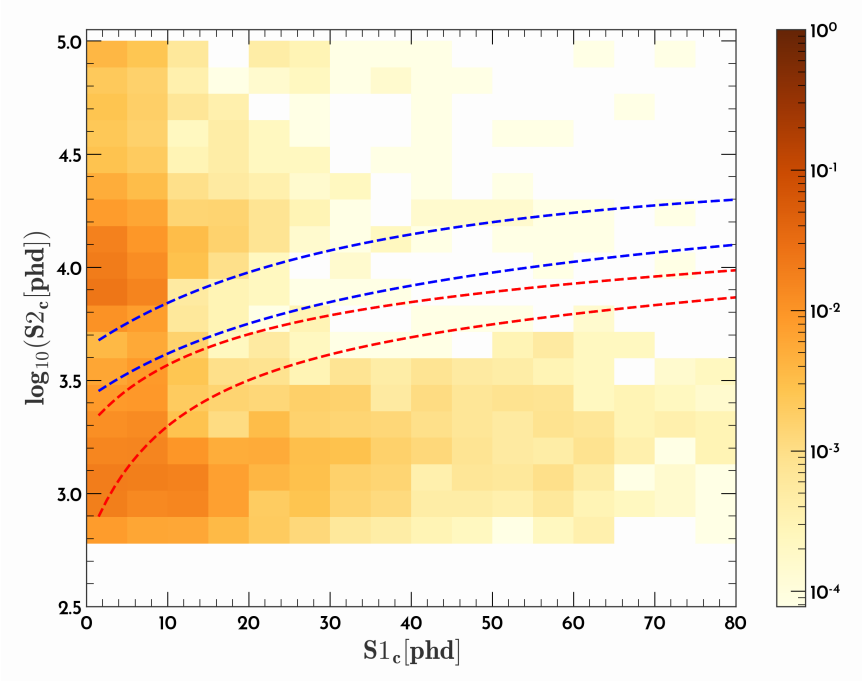
Figure 5.16: The initial accidentals distribution in the ROI with no cuts applied. This is given by the distribtution of Chopstitch accidentals (Figure 4.5) in the ROI normalised to the number of UDT events in the same region.

The BDT classification of the WIMP-search data is shown in Figure 5.23a. All 429 events are plotted with the accepted events shown in green and the rejected events shown in orange. A total of 311 events are accepted by the BDT compared with 335 events surviving all SR1 cuts. Figure 5.24 shows the how the above 429 events are classified with the SR1 analysis. The events in green in Figure 5.24 are the same events as the SR1 result shown in Figure 4.10 of the previous chapter.

Events that are classified differently by the BDT analysis and the SR1 analysis are shown in Figure 5.25. The BDT rejects 42 events that are accepted by the SR1 analysis (shown in yellow) and accepts 18 events that are rejected by the SR1 analysis (shown in purple), leading to 24 fewer events remaining in the ROI. The BDT score for the 18 events that are accepted by the the BDT but rejected by the SR1 analysis were in all cases close to the acceptance thresholds t with an average score of 0.99 and are unlikely to be accidental events given the 4 times lower accidental



(a)



(b)

Figure 5.17: The estimated distribution of accidental in the ROI after the SR1 cuts are applied is shown in (a) and (b) shows the distribution after the BDT analysis is applied.

leakage at the selected thresholds. A more likely explanation for these are single scatter events rejected by the SR1 analysis and are accounted for in the signal effi-

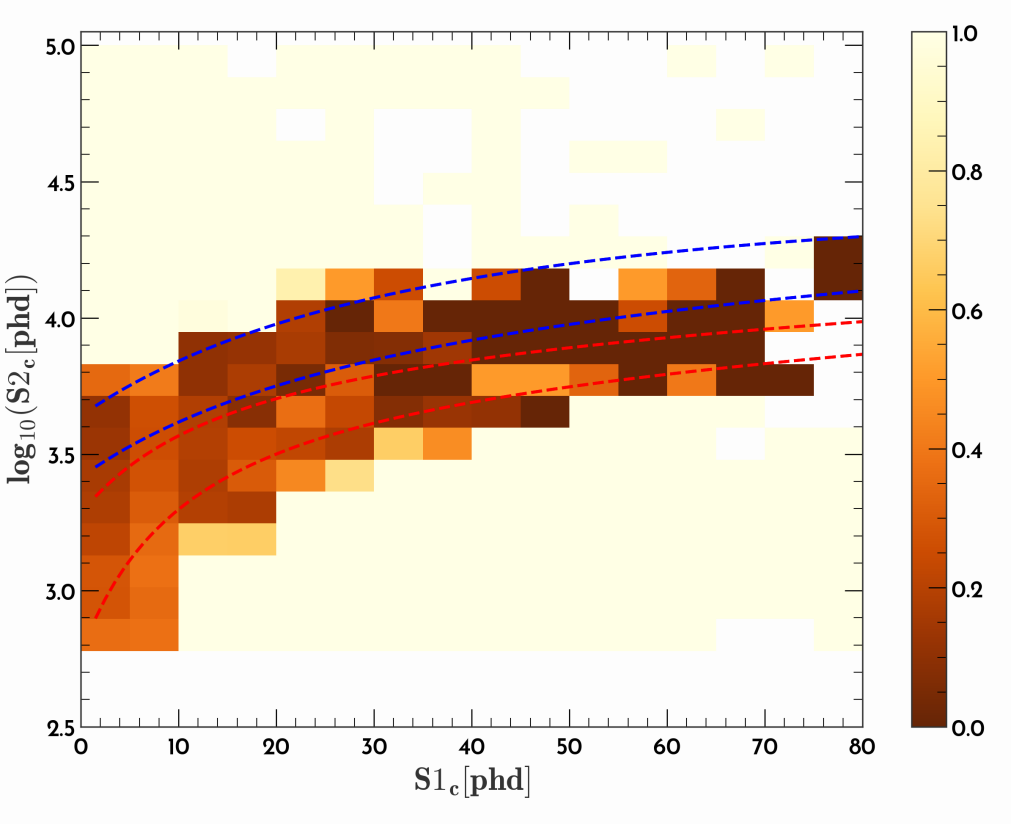
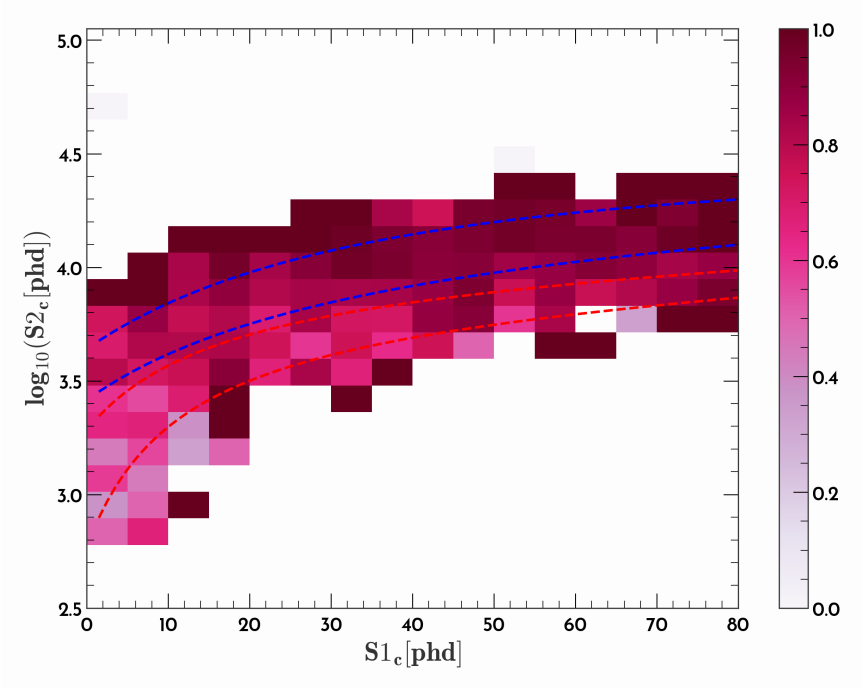


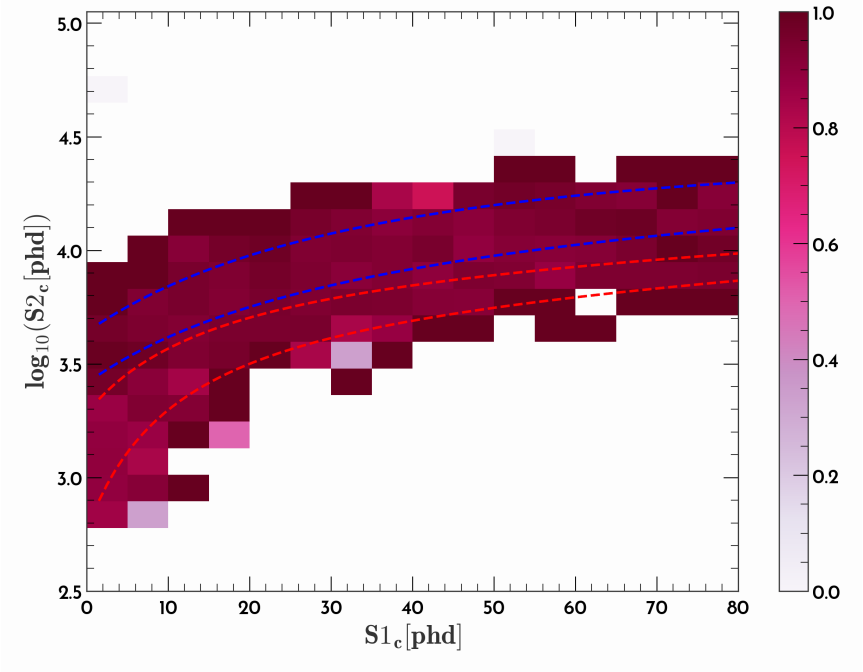
Figure 5.18: The plot shows the ratio of accidentals in each bin using the BDT analysis to the SR1 analysis. It is given by dividing the BDT accidentals distributions (Figure 5.17b) by the SR1 distribution (Figure 5.17a).

ciency. However not the same single scatter events are removed in the BDT analysis as the SR1 analysis since the cuts are not being applied in the same space. So it is more likely that the events are single scatters, however in the space in which the SR1 cuts are applied, simply adjusting the cuts to accept these events would lead to a lower purity. Though as mentioned earlier adjusting the signal acceptance is not as straightforward in the SR1 analysis as it requires loosening or tightening 10 individual cuts. Whereas with the BDT analysis since one knows the prior distribution of the BDT score for the accidental events (see Figure 5.12), assessing the type-II error (i.e. background leakage) associated with the BDT score of a given event is easily done since one is working with a 1D cut.

Of the 42 events that are rejected by the BDT but accepted by the standard analysis, all events except 1 have a BDT score above 0.5 suggesting these are likely signal events that are removed in the interest of signal purity. The one event with a BDT



(a)



(b)

Figure 5.19: The estimated signal efficiency using the BDT analysis is shown in (a) and (b) shows the signal efficiency obtained using the SR1 analysis cuts.

score below 0.5 that was accepted by the SR1 cuts was scored at 0.07. Using the prior distribution of the BDT score it was estimated that 90% of accidentals have

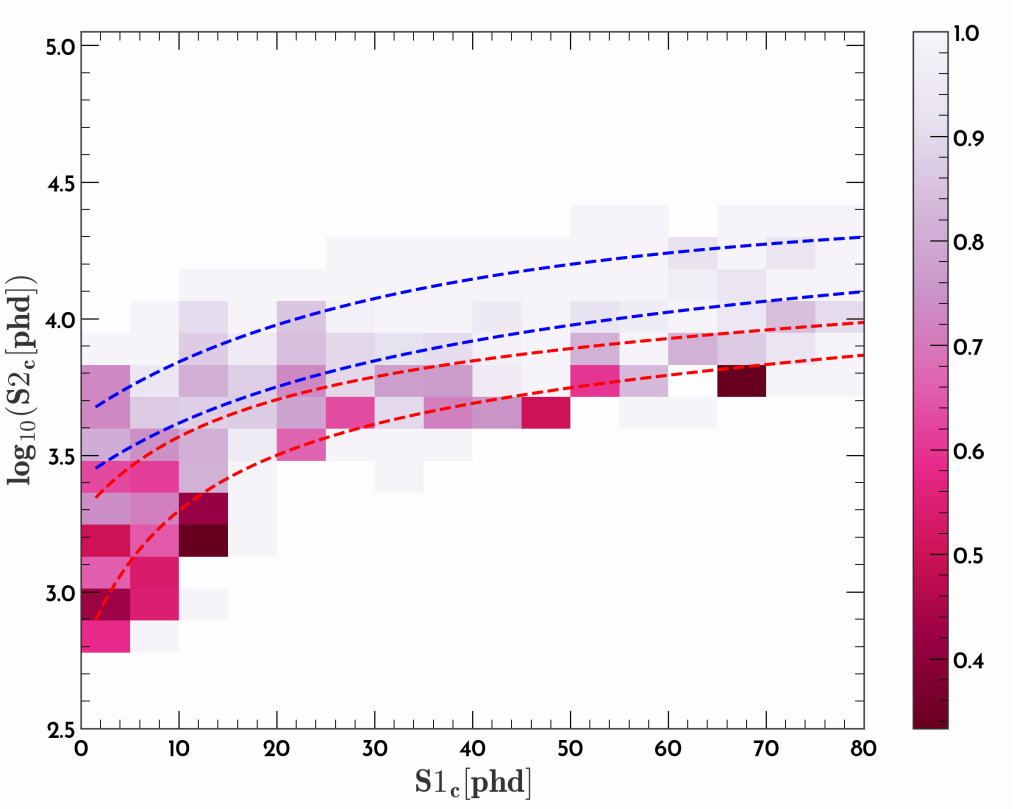


Figure 5.20: The fractional loss in signal efficiency going from the SR1 analysis to the BDT.

a BDT score that is lower than or equal to 0.07 and 0.8% of single scatter events have a score in this region. This suggests the event is more likely an accidental event that was removed by the BDT but left intact by the SR1 analysis. Removing this event is consistent with the accidental rate estimated after cuts in the SR1 analysis and consistent with the expected accidental rate with the BDT analysis. A visualisation of this event using the official LZ offline event display is shown in Figure 5.26. Accidental backgrounds appear deceptively similar to single scatters and are hence rejected systematically through statistical analyses like the one presented in this study.

To recap, the SR1 analysis expected 1.2 ± 0.3 accidental events in the ROI and the BDT analysis expects 0.54 ± 0.3. Only a single event that was rejected by the BDT but accepted by the SR1 analysis has a BDT score that is more likely to be an accidental event than a single scatter. The number of events rejected by the BDT is

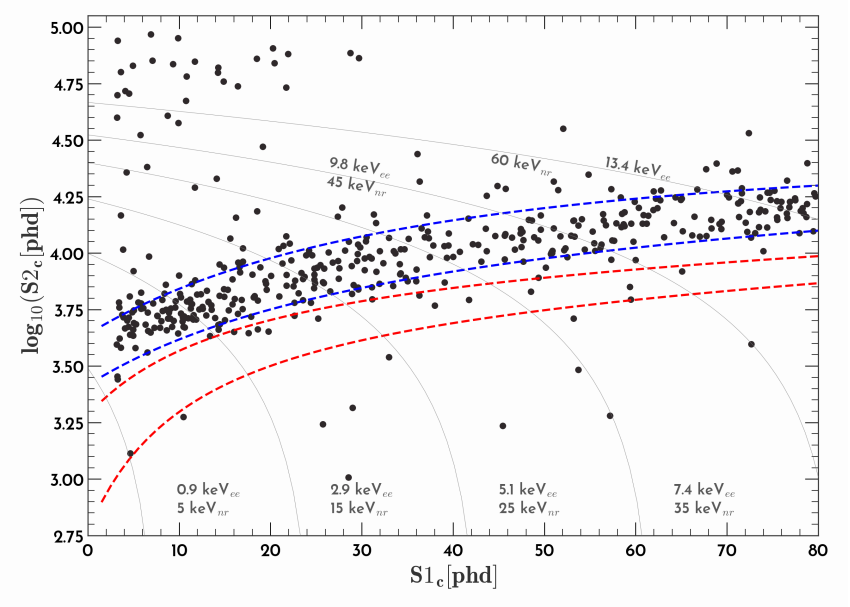


Figure 5.21: WIMP-search data in the FV after cuts common to the SR1 and BDT analysis are applied in $\log_{10}$ S2c-S1c space. The thin grey lines indicate contours of constant energy. The red dashed lines indicate the 10% and 90% quantiles of the NR region. The blue dashed lines indicate the 10% and 90% quantiles of the ER region.

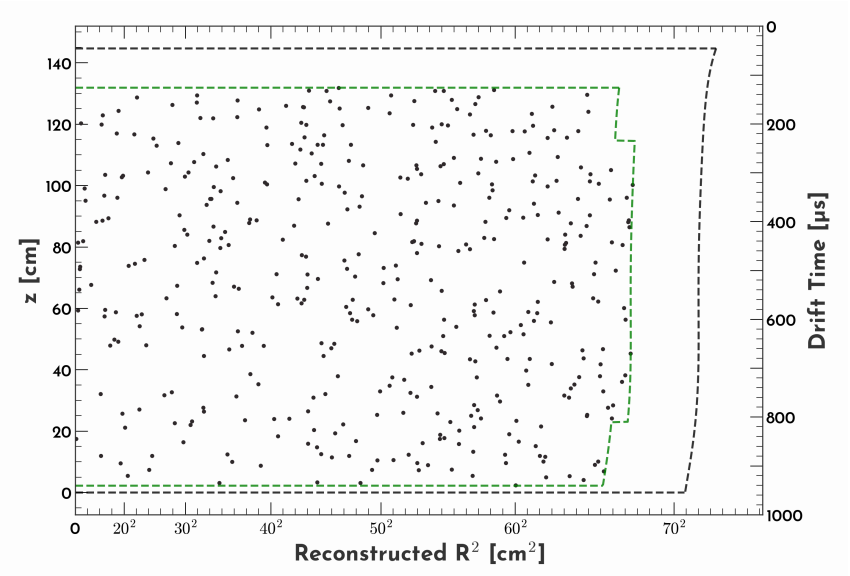
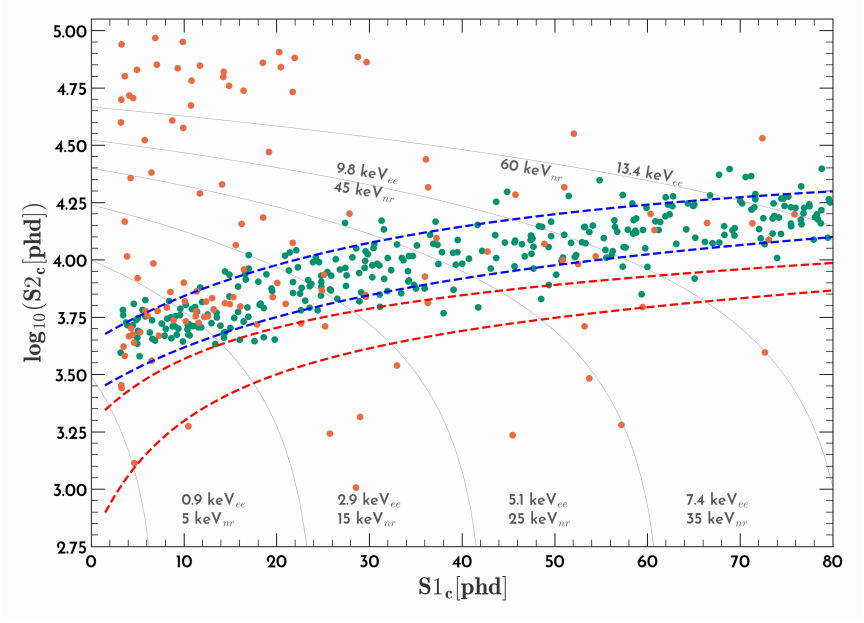
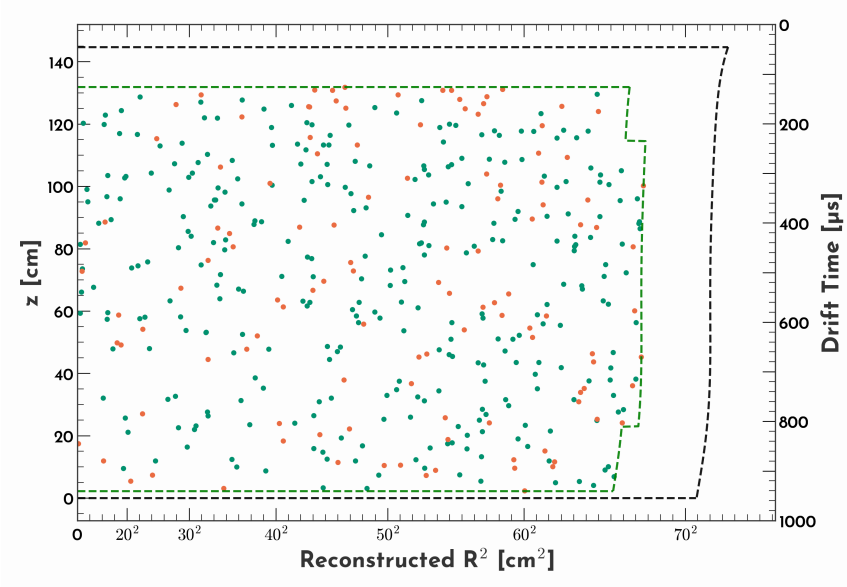


Figure 5.22: WIMP-search data in the FV after cuts common to the SR1 and BDT analysis are applied in reconstructed r^2 and z. The dashed black line shows the extent of the active TPC and the dashed green line shows the FV.

consistent with the signal efficiency of the BDT analysis within the 95% confidence limit calculated. Using the analysis presented in this study, in a 1000 day exposure LZ can expect 1.45 ± 0.8 events in the NR band compared with 6.00 ± 1.5 events



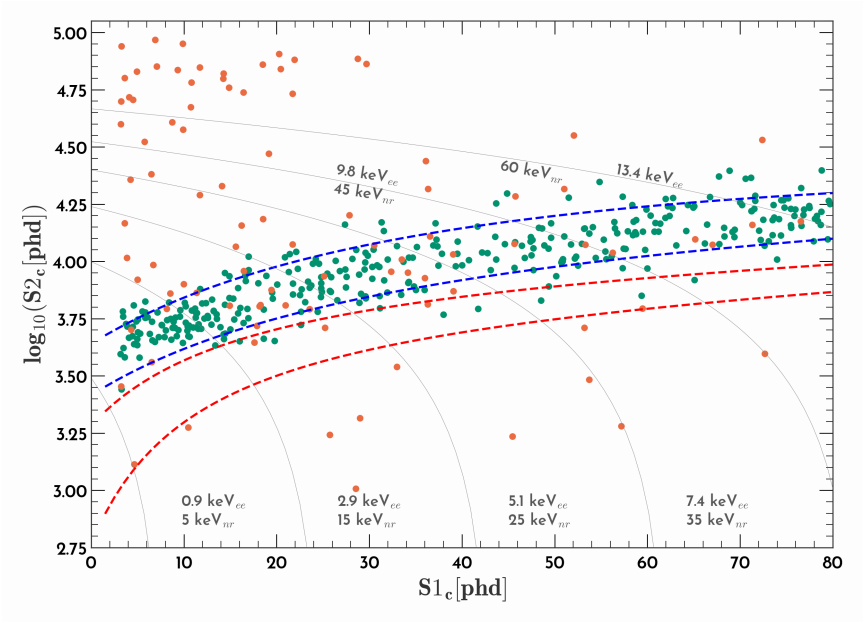
(a)



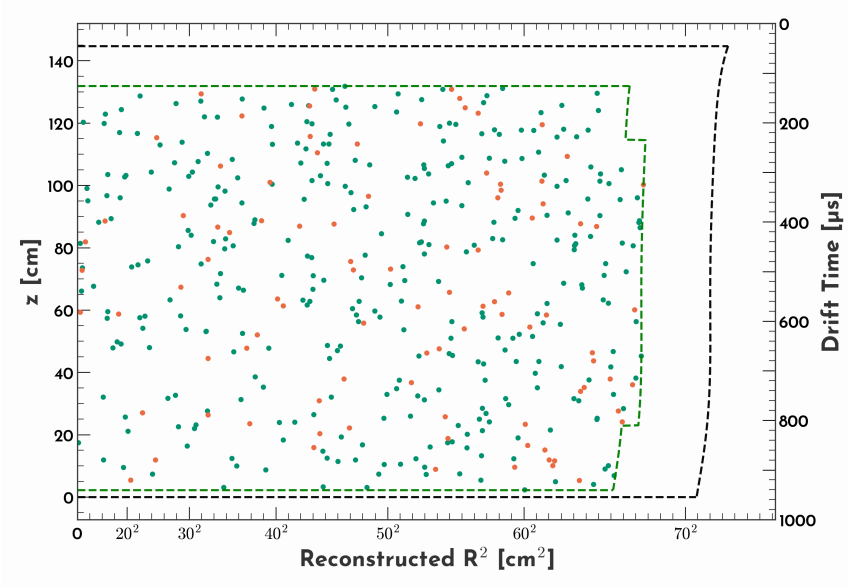
(b)

Figure 5.23: Results of BDT analysis on the WIMP-search data in the FV after cuts common to the SR1 and BDT analysis are applied in $\log_{10}$ S2c-S1c space in (a) and in in reconstructed r^2 and z in (b). The green events are accepted by the BDT analysis and the orange events are rejected.

using the SR1 analysis cuts thus meeting the required accidental rate to reach the projected sensitivity. This is achieved with a relatively small loss in signal efficiency that can be compensated for by a moderate increase in run time.

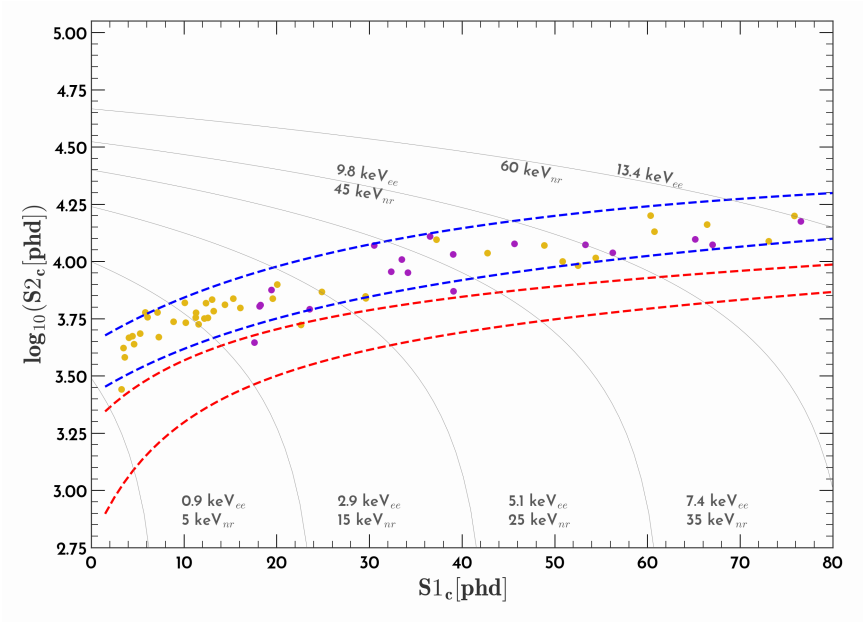


(a)

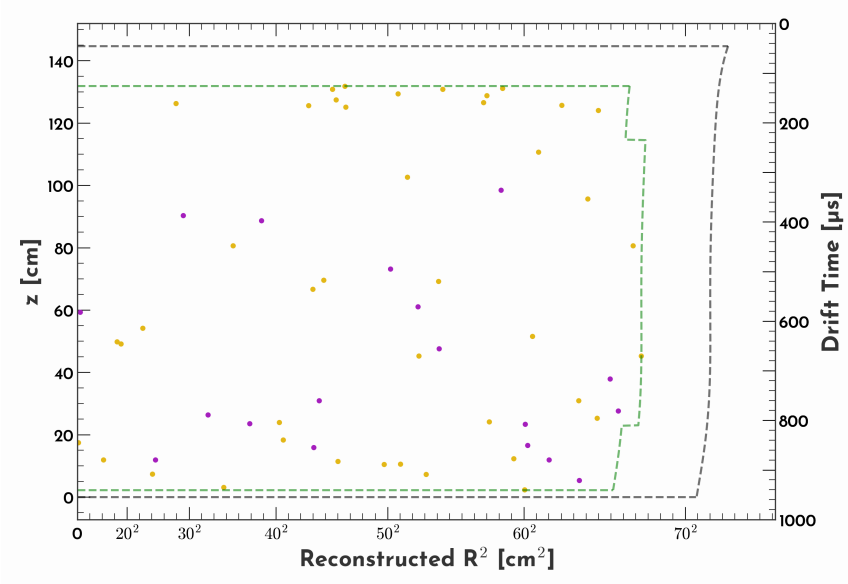


(b)

Figure 5.24: Result of the SR1 cut on the WIMP-search data in the FV after cuts common to the SR1 and BDT analysis are applied in $\log_{10}$ S2c-S1c space in (a) and in in reconstructed r^2 and z in (b). The green events are accepted by the SR1 analysis and the orange events are rejected.

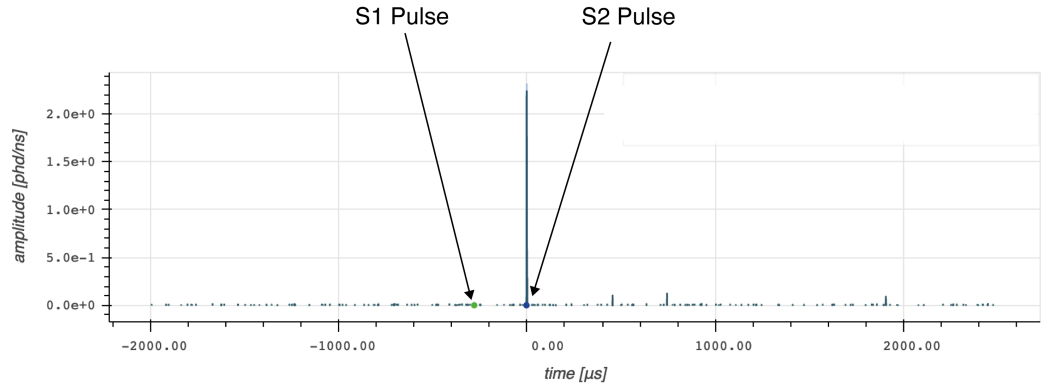


(a)

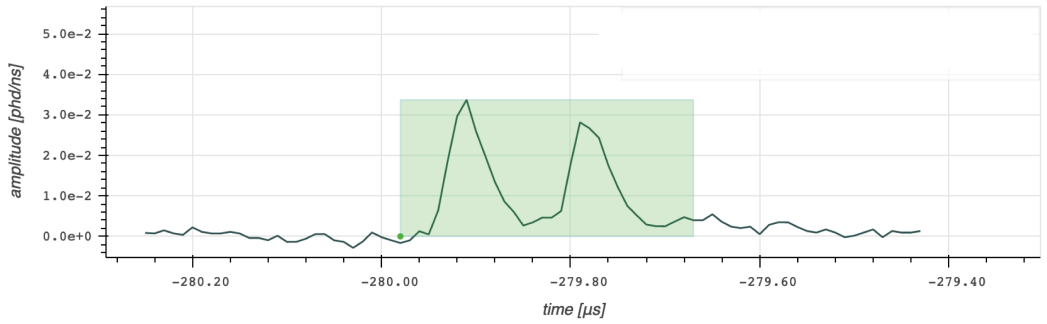


(b)

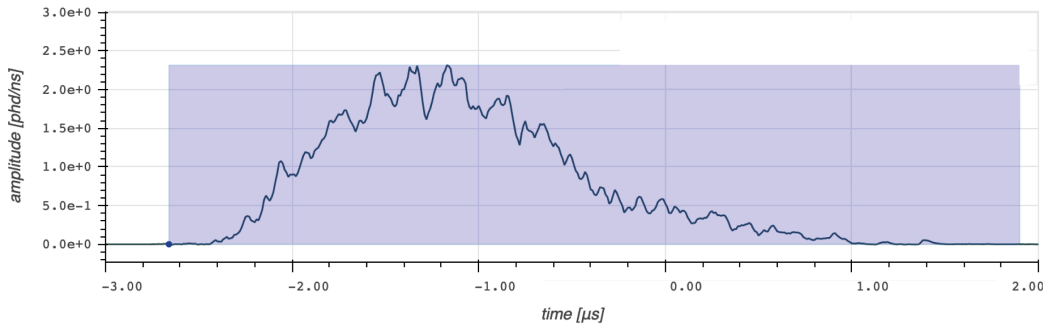
Figure 5.25: WIMP-search events that are classified differently by the two analyses shown in in $\log_{10}$ S2c-S1c space in (a) and in in reconstructed r^2 and z in (b). Events accepted by the BDT but rejected by the SR1 analysis are shown with purple markers and events that are rejected by the BDT but accepted by the SR1 analysis are shown with yellow.



(a)



(b)



(c)

Figure 5.26: A visualisation using the LZ event display of the event from the SR1 WIMP-search data identified by the BDT as an accidental (a) shows the full 4 µs event window with the position of the S1 and S2 pulse highlighted; (b) shows the S1 pulse with the pulse boundaries defined by the green box (c) shows the S2 pulse with the pulse boundaries defined by the purple box.

Chapter 6

Conclusion

The work presented in this thesis outlines some of my contributions to the LUX-ZEPLIN experiment, a dual-phase liquid xenon time projection chamber (TPC) located at the Sanford Underground Research Facility in South Dakota. In addition to the analyses presented, I also had the opportunity to contribute to the construction of the detector and am pleased to report that it has successfully completed a WIMP search, setting a world-leading limit on spin-independent WIMP-nucleon cross section. This is achieved by LZ through R&D efforts to improve TPC technology leading to better control and mitigation of backgrounds, more efficient detection of light and charge, and the successful operation of a detector with nearly 20 times the target volume of its predecessor.

Results from the first science run with a 60 day exposure using a fiducial mass of 5.5 tonnes have been presented in Chapter 4. A PLR analysis shows the data to be consistent with a background-only hypothesis. The highest sensitivity to spin-independent WIMP-nucleon scattering for masses greater than $9\,\mathrm{GeV/c^2}$ is reported with the most stringent limit set for spin-independent scattering at $36\,\mathrm{GeV/c^2}$, rejecting cross sections above $\sigma_{\mathrm{SI}} = 9.2 \times 10^{-48}\,\mathrm{cm^2}$ at the 90% confidence level.

However, the rate of accidental coincidence background events seen in SR1 (after analysis cuts are applied) would lead to an expected 6.00 ± 1.5 events in the NR band in a 1000 day exposure. This is well beyond the ~ 1 event requirement outlined in the Technical Design Report [10] for LZ to reach its projected sensitivity. In the SR1 analysis, accidentals were rejected using a set of sequential cuts which

were developed by understanding the sources of isolated S1 and S2 events and identifying RQs that can be used for rejection. In chapter 5, a BDT based analysis is presented that builds on the understanding of the underlying physics and uses machine learning to optimise the discrimination between signal and accidental events. This technique gives a four fold improvement in background rejection at the cost of a slight loss in signal efficiency. Using the procedure outlined LZ can expect 1.45 $\pm$ 0.8 events in a 1000 day exposure, thereby reducing the accidentals rate to an acceptable level.

The merits of this analysis, though, go beyond simply improving the rejection of accidentals. The use of a variable classification threshold as a function $t(S1c, \log_{10}\text{S2c})$ allows analysers to easily adjust signal efficiency and accidental leakage maps. This is possible because the BDT approach only requires the tuning of a cut on a single parameter, the BDT score. As opposed to tuning several cut thresholds simultaneously to achieve the desired background rejection and signal efficiency. Furthermore, as new RQs are identified they can easily be combined with the existing input parameters whilst still only tuning a single classification threshold.

Although the BDT analysis presented does improve the sequential cuts approach in SR1 and reduces the accidental rate to an acceptable level, it should be viewed as a blueprint containing the necessary steps and considerations needed for applying this machine learning technique to reject accidentals.

For the next science run, the main constraint in the development of this analysis, the limited amount of calibration data available for training, will no longer be present. And with more training data available one can expect to see an improvement in the performance of the BDT. Finer binning for the classification threshold $t(S1c, \log_{10}\text{S2c})$ could also be used, giving analysers the ability to tune the signal efficiency and accidental leakage maps with greater precision whilst maintaining or reducing the statistical uncertainty in each bin. With more data, bias mitigation can also be improved by using finer binning when assigning sample weights (for a flat distribution in $\log_{10}$S2c-S1c) to training events. Lastly, as our understanding

of isolated S1 and S2 pulses improves, enhancements to accidental rejection could also be found by developing new RQs and feeding these into the BDT.

With or without these improvements, in the future, the BDT model should be maintained by an expert and provided as an off-the-shelf classifier for analysers who can then, based on the requirements of their analyses, select their own thresholds ($t(S1c, \log_{10} \mathrm{S2c})$) to reject events, produce signal efficiency and accidental leakage maps, and conduct statistical inference.

Despite LZ seeing a higher accidental background rate than what was projected in the Technical Design Report [10], it has been demonstrated that these events can be reduced to acceptable levels using the machine learning analysis presented, and possibly reduced further with other innovative techniques and algorithms. In the coming years, LZ will continue to publish compelling, world leading results that will make significant contributions to the study of dark matter.

Bibliography

[1] K. G. Begeman, A. H. Broeils, and R. H. Sanders, "Extended rotation curves of spiral galaxies: dark haloes and modified dynamics," *Monthly Notices of the Royal Astronomical Society*, vol. 249, pp. 523–537, 04 1991.

[2] E. Hubble, "Photograph of galaxy cluster 0024+1654 as a gravitational lens from the hubble space telescope." `https://esahubble.org/images/opo9610a/`.

[3] N. Hubblesite, "A composite image of galaxy cluster 1e 0657-56 also known as the "Bullet Cluster"." `https://hubblesite.org/contents/news-releases/2006/news-2006-39.html`.

[4] Planck Collaboration, "Planck 2018 results - iv. diffuse component separation," *A&A*, vol. 641, p. A4, 2020.

[5] Planck Collaboration, "Planck 2018 results - vi. cosmological parameters," *A&A*, vol. 641, p. A6, 2020.

[6] ATLAS Collaboration, "Search for associated production of a z boson with an invisibly decaying higgs boson or dark matter candidates at s=13 tev with the atlas detector," *Physics Letters B*, vol. 829, p. 137066, 2022.

[7] M. Schumann, "Direct detection of wimp dark matter: concepts and status," *Journal of Physics G: Nuclear and Particle Physics*, vol. 46, p. 103003, aug 2019.

[8] S. Alsum, *Effective Field Theory search results from the LUX Run 4 data set, and construction of the LZ system test platforms*. PhD thesis, Wisconsin U., Madison, 2020.

[9] K. M. Stifter, *The LZ dark matter experiment : from detector development to early data*. PhD thesis, Stanford U., 2021.

[10] B. J. Mount *et al.*, "LUX-ZEPLIN (LZ) Technical Design Report," 2017.

[11] D. Akerib *et al.*, "The LUX-ZEPLIN (LZ) radioactivity and cleanliness control programs," *The European Physical Journal C*, vol. 80, p. 1044, Nov 2020.

[12] J. Aalbers *et al.*, "First dark matter search results from the lux-zeplin (lz) experiment," *Phys. Rev. Lett.*, vol. 131, p. 041002, Jul 2023.

[13] Y. Meng *et al.*, "Dark Matter Search Results from the PandaX-4T Commissioning Run," *Phys. Rev. Lett.*, vol. 127, p. 261802, Dec 2021.

[14] E. Aprile *et al.*, "Dark matter search results from a one ton-year exposure of XENON1T," *Phys. Rev. Lett.*, vol. 121, no. 11, p. 111302, 2018.

[15] D. Akerib *et al.*, "Results from a search for dark matter in the complete LUX exposure," *Phys. Rev. Lett.*, vol. 118, no. 2, p. 021303, 2017.

[16] P.-A. Amaudruz *et al.*, "First results from the DEAP-3600 dark matter search with argon at SNOLAB," *Phys. Rev. Lett.*, vol. 121, no. 7, p. 071801, 2018.

[17] V. C. Rubin and J. Ford, W. Kent, "Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions," , vol. 159, p. 379, Feb. 1970.

[18] R. Massey, T. Kitching, and J. Richard, "The dark matter of gravitational lensing," *Reports on Progress in Physics*, vol. 73, p. 086901, jul 2010.

[19] F. Zwicky, "Die Rotverschiebung von extragalaktischen Nebeln," *Helv. Phys. Acta*, vol. 6, pp. 110–127, 1933. [Gen. Rel. Grav.41,207(2009)].

[20] D. Fabricant, M. Lecar, and P. Gorenstein, "X-ray measurements of the mass of M 87.," , vol. 241, pp. 552–560, Oct. 1980.

[21] K. Garrett and G. Duda, "Dark Matter: A Primer," *Adv. Astron.*, vol. 2011, p. 968283, 2011.

[22] M. Markevitch, "Chandra observation of the most interesting cluster in the universe," *ESA Spec. Publ.*, vol. 604, p. 723, 2006.

[23] C. L. Bennett, D. Larson, J. L. Weiland, N. Jarosik, G. Hinshaw, N. Odegard, K. M. Smith, R. S. Hill, B. Gold, M. Halpern, E. Komatsu, M. R. Nolta, L. Page, D. N. Spergel, E. Wollack, J. Dunkley, A. Kogut, M. Limon, S. S. Meyer, G. S. Tucker, and E. L. Wright, "Nine-year wilkinson microwave anisotropy probe (wmap) observations: Final maps and results," *The Astrophysical Journal Supplement Series*, vol. 208, p. 20, sep 2013.

[24] W. J. Percival, S. Cole, D. J. Eisenstein, R. C. Nichol, J. A. Peacock, A. C. Pope, and A. S. Szalay, "Measuring the Baryon Acoustic Oscillation scale using the Sloan Digital Sky Survey and 2dF Galaxy Redshift Survey," *Monthly Notices of the Royal Astronomical Society*, vol. 381, pp. 1053–1066, 10 2007.

[25] S. D. McDermott, H.-B. Yu, and K. M. Zurek, "Turning off the lights: How dark is dark matter?," *Physical Review D*, vol. 83, mar 2011.

[26] B. Audren, J. Lesgourgues, G. Mangano, P. D. Serpico, and T. Tram, "Strongest model-independent bound on the lifetime of dark matter," *Journal of Cosmology and Astroparticle Physics*, vol. 2014, pp. 028–028, dec 2014.

[27] J. F. Navarro, C. S. Frenk, and S. D. M. White, ""the structure of cold dark matter halos"," , vol. 462, p. 563, May 1996.

[28] J. R. Primack, "The nature of dark matter," 2001. available at arXiv:astro-ph/0112255.

[29] P. W. J.M Overduin, *Dark Sky, Dark Matter*. London: CRC Press, 2002.

[30] D. Cerdeno, "Dark matter 101: from production to detection," 2016. Lecture series, Higgs Centre School of Theoretical Physics, available at https://api.semanticscholar.org/CorpusID:214637651.

[31] G. Jungman, M. Kamionkowski, and K. Griest, "Supersymmetric dark matter," *Physics Reports*, vol. 267, no. 5, pp. 195–373, 1996.

[32] J. F. Navarro, C. S. Frenk, and S. D. M. White, "A universal density profile from hierarchical clustering," *The Astrophysical Journal*, vol. 490, p. 493, dec 1997.

[33] Y. Shi, Z.-Y. Zhang, J. Wang, J. Chen, Q. Gu, X. Yu, and S. Li, "A cuspy dark matter halo," *The Astrophysical Journal*, vol. 909, p. 20, mar 2021.

[34] D. Merritt, A. W. Graham, B. Moore, J. Diemand, and B. Terzić, "Empirical Models for Dark Matter Halos. I. Nonparametric Construction of Density Profiles and Comparison with Parametric Models," *The Astronomical Journal*, vol. 132, p. 2685, nov 2006.

[35] W. J. G. de Blok, "The core-cusp problem," *Advances in Astronomy*, vol. 2010, 2010. Article ID 789293.

[36] M. Taoso, G. Bertone, and A. Masiero, "Dark matter candidates: a ten-point test," *Journal of Cosmology and Astroparticle Physics*, vol. 2008, p. 022, mar 2008.

[37] S. Dodelson and L. M. Widrow, "Sterile neutrinos as dark matter," *Phys. Rev. Lett.*, vol. 72, pp. 17–20, Jan 1994.

[38] S. Weinberg, "A new light boson?," *Phys. Rev. Lett.*, vol. 40, pp. 223–226, Jan 1978.

[39] F. Wilczek, "Problem of strong p and t invariance in the presence of instantons," *Phys. Rev. Lett.*, vol. 40, pp. 279–282, Jan 1978.

[40] R. D. Peccei and H. R. Quinn, "CP conservation in the presence of pseudoparticles," *Phys. Rev. Lett.*, vol. 38, pp. 1440–1443, Jun 1977.

[41] R. L. Workman and Others, "Review of Particle Physics, Sec 90: Axions and Other Similar Particles," *PTEP*, vol. 2022, p. 083C01, 2022.

[42] R. D. Peccei, *The Strong CP Problem and Axions*, pp. 3–17. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008.

[43] S. J. Asztalos, L. J. Rosenberg, K. van Bibber, P. Sikivie, and K. Zioutas, "Searches for astrophysical and cosmological axions," *Annual Review of Nuclear and Particle Science*, vol. 56, no. 1, pp. 293–326, 2006.

[44] J. L. Feng, "Dark matter candidates from particle physics and methods of detection," *Annual Review of Astronomy and Astrophysics*, vol. 48, no. 1, pp. 495–545, 2010.

[45] J. M. Gaskins, "A review of indirect searches for particle dark matter," *Contemporary Physics*, vol. 57, no. 4, pp. 496–525, 2016.

[46] ADMX Collaboration, "Search for invisible axion dark matter in the $3.3 - -4.2$ μeV mass range," *Phys. Rev. Lett.*, vol. 127, p. 261803, Dec 2021.

[47] A. Derevianko, V. A. Dzuba, V. V. Flambaum, and M. Pospelov, "Axio-electric effect," *Phys. Rev. D*, vol. 82, p. 065006, Sep 2010.

[48] LZ Collaboration, "Projected sensitivities of the LUX-ZEPLIN experiment to new physics via low-energy electron recoils," *Phys. Rev. D*, vol. 104, p. 092009, Nov 2021.

[49] XENON Collaboration, "Search for new physics in electronic recoil data from xenonnt," *Phys. Rev. Lett.*, vol. 129, p. 161805, Oct 2022.

[50] PandaX-II Collaboration vol. 38, p. 011301, Jan 2021.

[51] J. Billard, M. Boulay, S. Cebrián, L. Covi, G. Fiorillo, A. Green, J. Kopp, B. Majorovits, K. Palladino, F. Petricca, L. R. (chair), and M. Schumann,

"Direct detection of dark matter—appec committee report*," *Reports on Progress in Physics*, vol. 85, p. 056201, apr 2022.

[52] J. D. Lewin and P. F. Smith, "Review of mathematics, numerical factors, and corrections for dark matter experiments based on elastic nuclear recoil," *Astroparticle Physics*, vol. 6, pp. 87–112, Dec. 1996.

[53] S. Shaw, *Dark Matter Searches with the LUX and LZ Experiments*. PhD thesis, U. Coll. London, 2016.

[54] J. ENGEL, S. PITTEL, and P. VOGEL, "Nuclear physics of dark matter detection," *International Journal of Modern Physics E*, vol. 01, no. 01, pp. 1–37, 1992.

[55] V. Chepel and H. Araújo, "Liquid noble gas detectors for low energy particle physics," *Journal of Instrumentation*, vol. 8, p. R04001, apr 2013.

[56] R. Platzman, "Total ionization in gases by high-energy particles: An appraisal of our understanding," *The International Journal of Applied Radiation and Isotopes*, vol. 10, no. 2, pp. 116–127, 1961.

[57] A. Mangiarotti, M. Lopes, M. Benabderrahmane, V. Chepel, A. Lindote, J. Pinto da Cunha, and P. Sona, "A survey of energy loss calculations for heavy ions between 1 and 100kev," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 580, no. 1, pp. 114–117, 2007. Proceedings of the 10th International Symposium on Radiation Physics.

[58] P. Sorensen and C. E. Dahl, "Nuclear recoil energy scale in liquid xenon with application to the direct detection of dark matter," *Phys. Rev. D*, vol. 83, p. 063501, Mar 2011.

[59] J. Mock, N. Barry, K. Kazkaz, D. Stolp, M. Szydagis, M. Tripathi, S. Uvarov, M. Woods, and N. Walsh, "Modeling pulse characteristics in xenon with nest," *Journal of Instrumentation*, vol. 9, p. T04002–T04002, Apr 2014.

[60] D. Akerib *et al.*, "Identification of radiopure titanium for the lz dark matter experiment and future rare event searches," *Astroparticle Physics*, vol. 96, pp. 1–10, 2017.

[61] B. L. Paredes, H. Araújo, F. Froborg, N. Marangou, I. Olcina, T. Sumner, R. Taylor, A. Tomás, and A. Vacheret, "Response of photomultiplier tubes to xenon scintillation light," *Astroparticle Physics*, vol. 102, pp. 56–66, nov 2018.

[62] Hamamatsu, "Photomultiplier Tubes and Assemblies." `https://www.hamamatsu.com/eu/en/product/optical-sensors/pmt/pmt_tube-alone/head-on-type/R11410-20.html`.

[63] D. Akerib *et al.*, "Projected WIMP sensitivity of the LUX-ZEPLIN dark matter experiment," *Phys. Rev. D*, vol. 101, p. 052002, Mar 2020.

[64] D. Akerib *et al.*, "Tritium calibration of the lux dark matter experiment," *Phys. Rev. D*, vol. 93, p. 072009, Apr 2016.

[65] D. Akerib *et al.*, "Chromatographic separation of radioactive noble gases from xenon," *Astroparticle Physics*, vol. 97, pp. 80–87, 2018.

[66] J. Aalbers *et al.*, "Background Determination for the LUX-ZEPLIN (LZ) Dark Matter Experiment," *Phys. Rev. D*, vol. 108, p. 012010, Jul 2023.

[67] M.-M. Bé *et al.*, *Table of Radionuclides*, vol. 7 of *Monographie BIPM-5*. Pavillon de Breteuil, F-92310 Sèvres, France: Bureau International des Poids et Mesures, 2013.

[68] M.-M. Bé, V. Chisté, C. Dulieu, M. Kellett, X. Mougeot, A. Arinc, V. Chechev, N. Kuzmenko, T. Kibédi, A. Luca, and A. Nichols, *Table of Radionuclides*, vol. 8 of *Monographie BIPM-5*. Pavillon de Breteuil, F-92310 Sèvres, France: Bureau International des Poids et Mesures, 2016.

[69] J. Aalbers *et al.*, "Cosmogenic production of Ar-37 in the context of the LUX-ZEPLIN experiment," *Phys. Rev. D*, vol. 105, no. 8, p. 082004, 2022.

[70] D. Akerib *et al.*, "Radiogenic and muon-induced backgrounds in the lux dark matter detector," *Astroparticle Physics*, vol. 62, pp. 33–46, 2015.

[71] C. Zhang, D.-M. Mei, V. Kudryavtsev, and S. Fiorucci, "Cosmogenic activation of materials used in rare event search experiments," *Astroparticle Physics*, vol. 84, pp. 62–69, 2016.

[72] J. Back and Y. Ramachers, "Activia: Calculation of isotope production cross-sections and yields," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 586, no. 2, pp. 286–294, 2008.

[73] D. Akimov *et al.*, "Observation of coherent elastic neutrino-nucleus scattering," *Science*, vol. 357, pp. 1123–1126, sep 2017.

[74] N. Ackerman *et al.*, "Observation of two-neutrino double-beta decay in ^{136}Xe with the exo-200 detector," *Phys. Rev. Lett.*, vol. 107, p. 212501, Nov 2011.

[75] A. Gando *et al.*, "Measurement of the double-β decay half-life of ^{136}xe with the kamland-zen experiment," *Phys. Rev. C*, vol. 85, p. 045504, Apr 2012.

[76] M. Arthurs, *Radon Reduction and the First Science Results of the LZ Experiment*. PhD thesis, Michigan U., 2022.

[77] C. Faham, V. Gehman, A. Currie, A. Dobi, P. Sorensen, and R. Gaitskell, "Measurements of wavelength-dependent double photoelectron emission from single photons in vuv-sensitive photomultiplier tubes," *Journal of Instrumentation*, vol. 10, p. P09010, sep 2015.

[78] B. López Paredes, H. Araújo, F. Froborg, N. Marangou, I. Olcina, T. Sumner, R. Taylor, A. Tomás, and A. Vacheret, "Response of photomultiplier tubes to xenon scintillation light," *Astroparticle Physics*, vol. 102, pp. 56–66, 2018.

[79] V. N. Solovov *et al.*, "Position reconstruction in a dual phase xenon scintillation detector," *IEEE Trans. on Nucl. Sci.*, vol. 59, no. 6, pp. 3286–3293, 2012.

[80] D. Baxter *et al.*, "Recommended conventions for reporting results from direct dark matter searches," *Eur. Phys. J. C*, vol. 81, no. 10, pp. 1–19, 2021.

[81] M. Agostini *et al.*, "Simultaneous precision spectroscopy of pp, ^{7}Be, and pep solar neutrinos with Borexino Phase-II," *Phys. Rev. D*, vol. 100, p. 082004, Oct 2019.

[82] M. Agostini *et al.*, "Simultaneous precision spectroscopy of pp, ^{7}Be, and pep solar neutrinos with Borexino Phase-II," *Phys. Rev. D*, vol. 100, p. 082004, Oct 2019.

[83] N. Vinyoles *et al.*, "A new generation of standard solar models," *Astrophys J.*, vol. 835, no. 2, p. 202, 2017.

[84] B. Aharmim *et al.*, "Combined analysis of all three phases of solar neutrino data from the Sudbury Neutrino Observatory," *Phys. Rev. C*, vol. 88, p. 025501, Aug 2013.

[85] M. Berglund and M. E. Wieser, "Isotopic compositions of the elements 2009 (iupac technical report)," *Pure Appl. Chem.*, vol. 83, no. 2, pp. 397–410, 2011.

[86] E. Aprile *et al.*, "Observation of two-neutrino double electron capture in 124xe with xenon1t," *Nature*, vol. 568, no. 7753, pp. 532–535, 2019.

[87] J. B. Albert *et al.*, "Improved measurement of the $2\nu\beta\beta$ half-life of ^{136}Xe with the EXO-200 detector," *Phys. Rev. C*, vol. 89, p. 015502, Jan 2014.

[88] D. Akerib *et al.*, "Ultralow energy calibration of lux detector using ^{127}Xe electron capture," *Phys. Rev. D*, vol. 96, p. 112011, Dec 2017.

[89] T. Fruth, *PMT studies and loop antenna development for the LZ dark matter experiment*. PhD thesis, Oxford U., 2019.

[90] G. R. Araujo, T. Pollmann, and A. Ulrich, "Photoluminescence response of acrylic (PMMA) and polytetrafluoroethylene (PTFE) to ultraviolet light," *The European Physical Journal C*, vol. 79, aug 2019.

[91] R. Linehan, *High Voltage Electrode Development and the LZ Experiment's First WIMP Search*. PhD thesis, Stanford U., 2022.

[92] N. Abgrall *et al.*, "Muon flux measurements at the davis campus of the sanford underground research facility with the majorana demonstrator veto system," *Astroparticle Physics*, vol. 93, pp. 70–75, 2017.

[93] J. Lewin and P. Smith, "Review of mathematics, numerical factors, and corrections for dark matter experiments based on elastic nuclear recoil," *Astropart. Phys.*, vol. 6, no. 1, pp. 87–112, 1996.

[94] M. C. Smith *et al.*, "The RAVE survey: constraining the local galactic escape speed," *Mon. Not. R. Astron. Soc.*, vol. 379, no. 2, pp. 755–772, 2007.

[95] C. McCabe, "The earth's velocity for direct detection experiments," *J. Cosmol. Astropart. Phys.*, vol. 2014, no. 02, p. 027, 2014.

[96] R. Schönrich, J. Binney, and W. Dehnen, "Local kinematics and the local standard of rest," *Mon. Not. R. Astron. Soc.*, vol. 403, no. 4, pp. 1829–1833, 2010.

[97] J. Bland-Hawthorn and O. Gerhard, "The galaxy in context: Structural, kinematic, and integrated properties," *Annu. Rev. Astron. Astrophys.*, vol. 54, no. 1, pp. 529–596, 2016.

[98] Abuter, R. *et al.*, "Improved gravity astrometric accuracy from modeling optical aberrations," *Astron. Astrophys.*, vol. 647, p. A59, 2021.

[99] D. Akerib *et al.*, "Simulations of events for the LUX-ZEPLIN (LZ) dark matter experiment," *Astropart. Phys.*, vol. 125, p. 102480, 2021.

[100] J. Allison *et al.*, "Recent developments in Geant4," *Nucl. Instrum. Methods Phys. Res., Sect. A*, vol. 835, pp. 186–225, 2016.

[101] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Power-Constrained Limits," 5 2011. available at: arXiv:1105.3166v1.

[102] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," *CoRR*, vol. abs/1603.02754, 2016.

[103] DØ Collaboration, "Evidence for production of single top quarks and first direct measurement of $|V_{tb}|$," *Phys. Rev. Lett.*, vol. 98, p. 181802, May 2007.

[104] ATLAS Collaboration, "Measurement of higgs boson decay into b-quarks in associated production with a top-quark pair in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector," *Journal of High Energy Physics*, vol. 2022, jun 2022.

[105] LUX Collaboration, "Improving sensitivity to low-mass dark matter in lux using a novel electrode background mitigation technique," *Phys. Rev. D*, vol. 104, p. 012011, Jul 2021.

[106] PandaX-II Collaboration, "Study of background from accidental coincidence signalsin the pandax-ii experiment*," *Chinese Physics C*, vol. 46, p. 103001, oct 2022.

[107] Y. Freund and R. E. Schapire, "A desicion-theoretic generalization of on-line learning and an application to boosting," in *Computational Learning Theory* (P. Vitányi, ed.), (Berlin, Heidelberg), pp. 23–37, Springer Berlin Heidelberg, 1995.

[108] J. Friedman, T. Hastie, and R. Tibshirani, "Additive logistic regression: a statistical view of boosting (With discussion and a rejoinder by the authors)," *The Annals of Statistics*, vol. 28, no. 2, pp. 337 – 407, 2000.

[109] T. Chen and C. Guestrin., "XGBoost Package." `https://github.com/dmlc/xgboost`, 2022.

[110] T. Chen and C. Guestrin., "XGBoost Documentation." `https://xgboost.readthedocs.io/`, 2022.

[111] C. J. Clopper and E. S. Pearson, "The use of confidence or fiducial limits illustrated in the case of the binomial," *Biometrika*, vol. 26, no. 4, pp. 404–413, 1934.