Application of Machine Learning Techniques to Direct Detection Dark Matter Experiments

Omar Jahangir

Thesis submitted for the degree of

Doctor of Philosophy

of

University College London.

Department of Physics and Astronomy
University College London

March, 2022

Abstract

Determining the nature of Dark Matter has been one of the biggest mysteries over the past few decades. Cosmological models predict a universe comprising of 26% Dark Matter, with Weakly Interacting Massive Particles (WIMPs) being one of the leading candidates to explain its nature. The LUX-ZEPLIN (LZ) experiment aims to explore the nature of Dark Matter. Using a dual-phase liquid xenon time projection chamber placed 4850 feet underground at the Sanford Underground Research Facility (SURF) in SD USA, LZ hopes to reach groundbreaking sensitivities of 1.6×10^{-48} cm^2 for a 40 GeV/c^2 WIMP mass.

To prepare for the LZ data taking at the end of 2021, novel techniques used in Machine Learning (ML) are used to develop and improve on existing data analysis methods currently employed. ML, which is a sub-field of Artificial Intelligence (AI), has seen some of the biggest growth over the past decade. The first part of this thesis will concentrate on using standard ML techniques to improve on bismuth-polonium (BiPo) tagging, which is vital to be able to constrain the backgrounds generated by radon. Using a Random Forest classifier, simulated BiPo events are trained upon, with the aim to find missed BiPo events not within the classical regions of interest. This allows for further constraints on the total radon contribution.

The second part of this thesis will look at using Deep Learning - a subset of ML, to explore position reconstruction techniques particularly important for events near the walls of noble liquid TPCs, with the aim to increase the usable fiducial volume. The impact of implementing such methods is illustrated using the LZ experiment.

Impact Statement

Research into confirming the existence of Dark Matter is a hot topic within the High Energy Physics community, with the most sensitive Direct Detection Dark Matter experiment about to start taking data within a few months. The LUX-ZEPLIN (LZ) experiment will be the largest dark matter detector which operates a xenon target, on earth. It will look for Weakly Interacting Massive Particle (WIMP) nucleon interactions and will be the most sensitive experiment yet. Many challenges were overcome in making the LZ experiment come to fruition, with engineering challenges arising from using over 10 tonnes of liquid xenon, and placing it over a mile underground in the Homestake gold mine at the Sanford Underground Research Facility, South Dakota US.

The use of Artificial Intelligence - another hot topic within the technology industry, will be applied within this thesis. By using Artificial Neural Networks, built to mimic the neurons inside the human brain, we were able to explore position reconstruction techniques tested on simulated datasets of the LZ detector. This improvement may allow experiments such as LZ to have an even greater sensitivity to WIMPs than predicted. Furthermore, other aspects of AI were utilised to show the power that it can have in the analysis of data from physics experiments; namely the use of Random Forests classifiers to classify Radon events.

Beyond academia, the use of AI will have a major impact on society, with possible applications in healthcare, energy and transport sectors. While using Neural Networks for position reconstruction are already being used to control autonomous cars.

Acknowledgements

بسم اللّه الرحمان الرحيم

الحمد للّهِ الذي خلق السماوات بقدرة، وجعل الأرض بلطفة، و صوّر الانسان بنطفة، و كلّ شيء هالك الا وجهه . هو الملك الجبّار، العزيز الغفّار، الذي خلق كل شيء بمقدار.

وصلاة وسلام على سيّد الأنبياء والمرسلين، وعلى آله واصحابه اجمعين ، وعلى من تبعه بإحسان إلى يوم الدين، آمين .

أشهد ان لا اله الا اللّه، و أشهد انّ محمدا عبده و رسوله .

اما بعد،

All praise is due to Allah, The Creator, The Sustainer, The Originator of all that is, and all that will be. And peace and blessings be upon all His Messengers, and those who follow them in goodness - Ameen.

There are numerous people who have helped in allowing me to complete this PhD. Firstly, I'd like to thank my mother, father, sister and wife for their infinite love in encouraging and supporting me to pursue this PhD. Without their duas, it definitely would not have been possible. I am also eternally grateful to all my teachers throughout the years, who have kept me in their prayers, supported me, and helped me become the person I am today.

I'd like to extend my deepest gratitude and appreciation to my supervisor Professor Chamkaur Ghag for all the support, supervision, feedback and guidance he has given throughout the last four years. Our discussions on all matters - whether scientific, philosophical or theological, will forever be cherished. This gratitude is also extended to the entire UCLDM team to whom I was fortunate enough to call my friends and colleagues. The support specifically given by T.Fruth and J.Dobson truly was invaluable, and I cannot thank them enough.

Being a member of the LZ collaboration was an honour, and I am genuinely grateful to all the hardwork and effort given by the whole collaboration in making the LZ project a success. I am also grateful to everyone at the CDT in Data Intensive Science for giving me the opportunity to combine my love of Physics and AI into one project.

I pray that the creator of the Heavens and the Earth - and all that is within it, allow this PhD to be a means of benefit to the Ummah, and to allow it to be a source of goodness for me and my family in this dunya and the next. Ameen.

ربنا تقبّل منا، إنك أنت السميع العليم. و تب علينا، انّك أنت التواب الرحيم. وصلى
الله على خاتم الأنبياء والمرسلين ، وعلى آله واصحابه اجمعين. آمين يا رب العالمين.

عمر بن محمد جهانجير
١٧ صفر ١٤٤٣

Declaration

I, Omar Jahangir, 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 thesis. The following section will highlight the chapters of this thesis, specifying my contributions to the broader work that makes up the entirety of this work.

In Chapter 1, the theoretical topic of dark matter within a cosmological paradigm is introduced. Published results from many leading experiments will be referenced, with some theoretical motivations being discussed.

In Chapter 2, the principles of direct detection dark matter experiments is introduced, particularly dual-phase liquid xenon detectors. The LUX-ZEPLIN experiment will be primarily referenced, as it is the experiment upon which the work in this thesis is based. A part of this PhD was devoted to understanding the LZ experiment, and developing and constructing parts of the detector which is now underground. This included carrying out the cabling work necessary to connect all the PMTs in unison.

In Chapter 3, the topic of Machine Learning is introduced, with the fundamentals of the learning process being discussed. Neural networks will be introduced, with a focus on Convolutional Neural Networks (CNNs) and how they work is explained.

In Chapter 4, the Bismuth-Polonium (BiPo) tagging method currently implemented within the LZ experiment is reviewed, including the theoretical reasoning behind the cuts. This work was built upon the work already done by the LZ collaboration. However, this classical approach was applied from scratch from me using Python, with new time differences between the S1 and S2 signals in a BiPo event being applied. Furthermore, the innovative application of Machine Learning techniques to the problematic radon backgrounds mentioned in this chapter was my work. The data used within this chapter was produced from the Mock Data Challenge 3 and was available to all collaborators within the LZ experiment.

In Chapter 5, the current position reconstruction algorithm implemented by the LZ collaboration is introduced, with an explanation of the underlying theory. My contribution was the development of the 1D and 2D Convolutional Neural Networks (CNNs) using the Keras framework. The positioning of the PMTs for the 2D-ML input were developed by me and another LZ collaborator. Calculations on the effect of having a larger fiducial volume on the sensitivity was carried out by using code developed by the LZ collaboration, while the backgrounds of a larger fiducial volume were carried out with the aid of LZ collaborators. The improvement in sensitivity resulting from the larger fiducial volume achievable with this technique was estimated.

In Chapter 6, the application of machine learning algorithms specific to time-series datasets is applied onto the raw data produced by the PMT signals. The labels for the dataset within this chapter were taken from a hand scanning campaign in which over 15 LZ collaborators took part, of which I was one. The application of the ML methods was carried out by me.

Omar Jahangir

September, 2021

Contents

List of Figures

List of Tables

Chapter 1

Dark Matter

Surely, in the creation of the Heavens and the Earth, and in the alternation of night and day, are signs for the people of wisdom.

Quran [3:190]

Ever since the dawn of creation, humans have looked up and marvelled at the infinite vastness that lay before them. With the advancements of science and technology, this endeavour has led mankind to look further and gain a deeper understanding of the universe, and more importantly, their place within it.

Through these observations, an unknown substance seems to present itself to observers. This dark matter, as astronomers have coined, has eluded identification by researchers, despite observations of its gravitational interference with ordinary matter indicating to its presence [1]. The best cosmological models of the universe suggest a universe comprised of 26% of dark matter. In this chapter, the evidence for dark matter, proposed candidates and search methods are reviewed.

1.1 Evidence of Dark Matter

In this section, discussions for the evidence of dark matter will be carried out, starting with observations carried out in the 1930s, and ending with results from measurements of the Cosmic Microwave Background (CMB) radiation.

1.1.1 Galaxy Clusters

During the 1930s, Fritz Zwicky was measuring galaxy clusters, when he noticed that the mass that can be inferred by the stellar material in the Coma Galaxy cluster was different to what can be inferred by looking at velocity dispersions of the galaxy (see [2] for an example of the velocity dispersions of the 21 cm line observations of the Coma cluster).

By counting the number of galaxies, and estimating the average mass of each galaxy, Zwicky was able to calculate the potential energy. By using the virial theorem to relate the total kinetic energy of a galaxy cluster to the gravitational potential energy within the cluster, and applying it to velocity dispersions measured by the galaxy, Zwicky found a mass-to-light ratio of over 400 solar masses [3].

The virial theorem relates the average total energy of discrete particles within a stable system to the potential energy of the system, and can be written as

$$\langle T \rangle = -\frac{1}{2} \sum_{k=1}^{N} \langle F_k \cdot r_k \rangle \tag{1.1}$$

where $\langle T \rangle$ relates the total time-averaged kinetic energy of the system, F_k the force on a galaxy k, and r_k representing the position of the particle within the system.

The large mass-to-light ratio seen by Zwicky would indicate the existence of a type of matter which would interact gravitationally, hence affecting the gravitational potential energy of the galaxy, whilst not releasing any light, hence not contributing to the luminosity of the galaxy [4] [5]. Further observations would go on to confirm what Zwicky found, with observations of galaxy rotation curves being one of the first to confirm the existence of a non-luminous type of matter.

1.1.2 Galaxy Rotations

Vera Rubin was one of the first pioneers to present evidence for dark matter in the 1970s [6]. By studying the rotation curve measurements of galaxy clusters, Rubin's results showed an increase in stellar velocity up to an initial radius from the galaxy core, with an almost constant velocity value thereafter the further one got from the galaxy core.

However, according to Newtonian classical dynamics, the velocity of objects as a function of radial distance r from the centre of a galaxy should be

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

with the mass

$$M(r) = 4\pi \int \rho(r) r^2 dr \tag{1.3}$$

Since most of the matter is expected to be contained within the galaxy core, outside the core, the dark matter density should fall as $\rho(r) \propto \frac{1}{\sqrt{r}}$ when approaching large radii, as seen in equation 1.2.

In Figure 1.1, which shows the rotation curve of dwarf spiral galaxy NGC 6503, at large distances from the galactic core, the velocity remains constant (dash-dotted line). Hundreds of rotation curves have been measured [8] [9], and though there are some anomalous results, the general pattern is that rotation velocities become flat outside galactic cores, whereas from Equation 1.2 simple Newtonian dynamics without dark matter implies that they should fall off as $1/\sqrt{r}$.

This would suggest the existence of a halo of matter with a density profile derived as follows [10]:

$$\frac{GMm}{r^2} = \frac{mv_c^2}{r} \tag{1.4}$$

where M is the mass of the galaxy, and m the mass of the object in orbit, r the

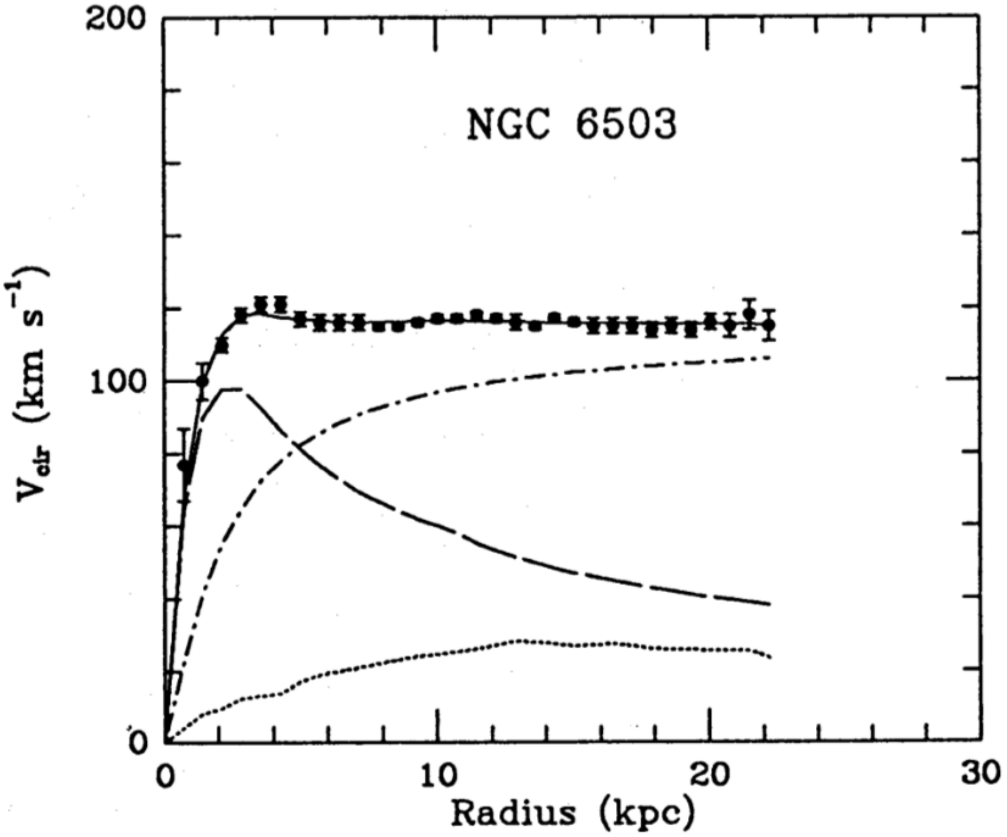


Figure 1.1: Obervational data of Galaxy curves. Dashed line indicates contribution from optical disk, dotted line from gas, and dash-dotted line from the dark matter halo. Image taken from [7].

distance between them, and v_c the constant velocity observed at large distances. Since a constant rotational velocity at large distances v_c is observed (as seen in figure 1.1), rearranging the above equation, the function of the mass distribution as a function of radius can be written as:

$$M(r) = \frac{v_c^2 r}{G} \tag{1.5}$$

Taking a radial shell of thickness dr at a radius r from the centre of a galaxy core, the differential mass enclosed within this shell is

$$\begin{gathered} dM(r) = 4\pi r^2 \rho(r) dr \\ \frac{dM(r)}{dr} = 4\pi r^2 \rho(r) \end{gathered} \tag{1.6}$$

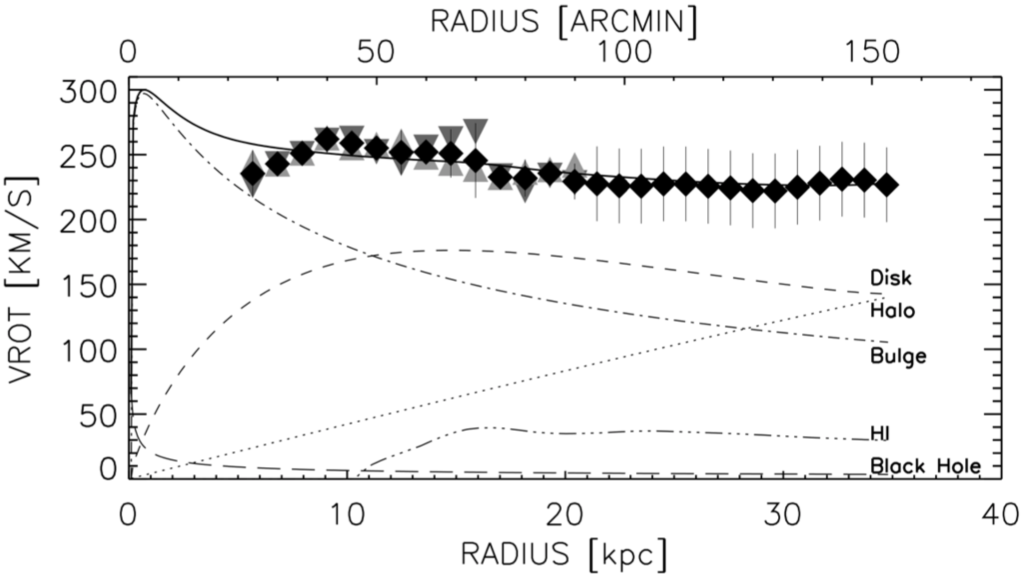


Figure 1.2: Plot showing the galaxy rotation curve for the M31 Andromeda galaxy. The solid line shows the best fit to the data (square points). Dashed lines show the contribution from different models, as labelled on the plot. Plot taken from [11]. More galaxy rotation curves can be found in [9].

where $\rho(r)$ represents the density of the galaxy. Substituting equation 1.5 for M(r) into above equation, and differentiating will give the following

$$\begin{aligned} 4\pi r^2 \rho(r) &= \frac{v_c^2}{G} \\ \rho(r) &= \frac{v_c^2}{4\pi r^2 G} \end{aligned} \tag{1.7}$$

Hence, this gives a density profile of the following form

$$\rho(r) \propto \frac{1}{r^2} \tag{1.8}$$

This sort of behaviour is present in other galaxies, with Andromeda (figure 1.2) and the Milky Way also showing a constant velocity profile at further distances from their respective galaxy centres [11][12]. All this evidence suggests a presence of a dark matter halo, which would account for a large proportion of the mass within a galaxy.

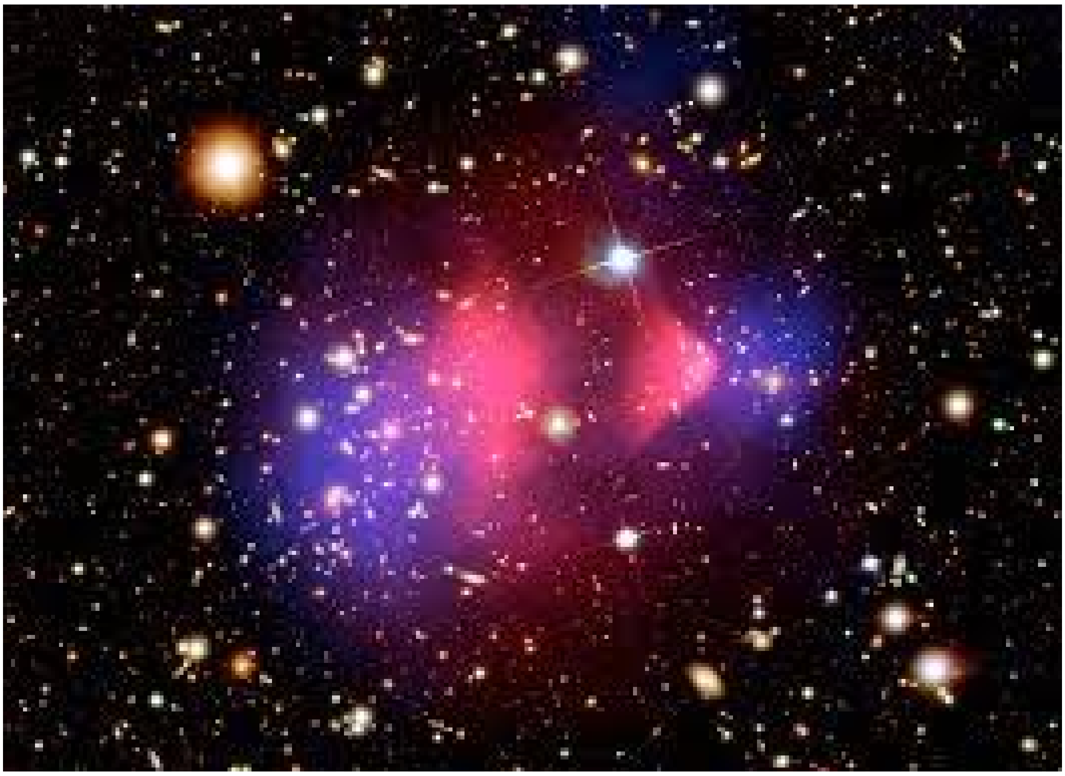


Figure 1.3: Image of Bullet Cluster, with blue/purple region showing areas of greatest gravitational potential mapped using gravitational lensing, and pink indicating interstellar gas regions. The image indicates the nature of dark matter being non-interacting [13]. Image taken from [14].

1.1.3 Galaxy collisions

Another famous evidence for dark matter is the collision of two galaxies, as seen in figure 1.3, and is an image of the Bullet Cluster galaxy [14]. Galaxies are comprised of both luminous matter ($\sim 1\% - 2\%$ [15]) and plasmas ($\sim 5\% - 15\%$ [16]). During a collision, the galaxies act as collisionless particles, while the plasma experiences a ram-pressure [14], which slows down the movement of the plasma. Therefore the galaxies within the collision decouple from the plasma.

Mapping the location of the plasma by measuring the emitted X-rays, a comparison between the location of luminous matter and the plasma can be seen. The plasma in each galaxy, interacting with each other, will remain closer to the point of interaction.

In the absence of non-luminous dark matter, the gravitational potential of the col-

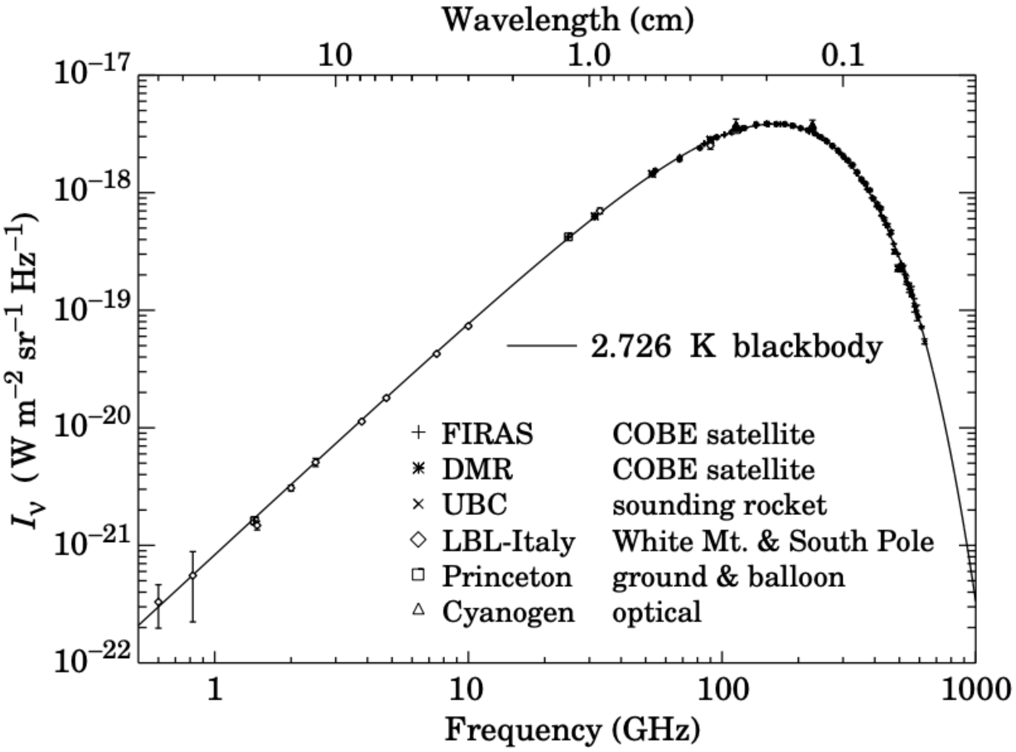


Figure 1.4: Plot showing the thermal spectrum of the CMB. Points indicate various inputs of data, with solid line representing 2.73 K blackbody. Image taken from [17].

liding galaxies should trace the dominant visible component caused by the X-ray emitting plasmas. However, in the presence of collisionless dark matter dominating the mass component of the galaxies, the gravitational potential will follow the location of the dark matter during the collision. This can be seen in figure 1.3, where the location of the potential (as measured by weak gravitational lensing), is located near the brightest point of the cluster (seen in purple). This indicates that the majority of the galaxy consists of dark matter, thus interacting gravitationally but without interacting with other particles in the collision [13].

1.1.4 Cosmological evidence

The Cosmic Microwave Background (CMB) is primordial radiation from the early universe. According to the Hot Big Bang Model, the early universe is full of particles constantly scattering and producing radiation, with the last scattering taking place when the universe was only 300,000 years old, after which the photons have

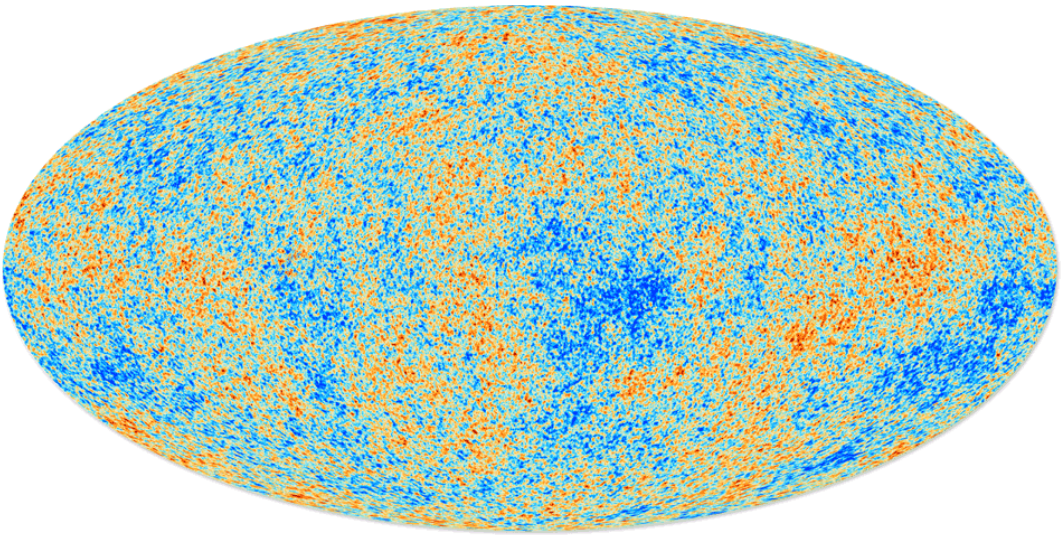


Figure 1.5: A map of the CMB radiation. Image shows the temperature isotropies, with the temperature difference between blue and red regions only a few thousandths of a Kelvin. These correspond to matter density fluctuations within the distribution. Image taken from [18].

been travelling through the universe freely. This remnant radiation is still present today but has been redshifted due to the expansion of the universe. Photons belonging to the CMB can be detected with a spectral distribution following that of the black-body function with T= 2.725K [19], the spectrum of which can be seen in figure 1.4.

On large scales, the CMB is isotropic and homogenous. But subtracting the Milky Way galaxy's microwave foreground and dipole due to the galaxy motion relative to the rest frame of the CMB radiation, the CMB is found to be uniform to 1 part in 10^5. Fluctuations at the level of 1 part in 10^5 as a function of position on the sky are a rich source of data for astrophysicists, who have used the so-called cosmic background radiation data to estimate the visible and dark matter contribution to the overall matter content of the Universe. Figure 1.5 shows these small temperature fluctuations (called anisotropies) as measured by the Planck collaboration [18]. These anisotropies can be used to determine cosmological parameters.

By looking at figure 1.5, a mapping can be found at different angular scales. The angular scales represent the different physical sizes of the anisotropies and can be parameterised as the multipole moments of spherical harmonics. Spherical har-

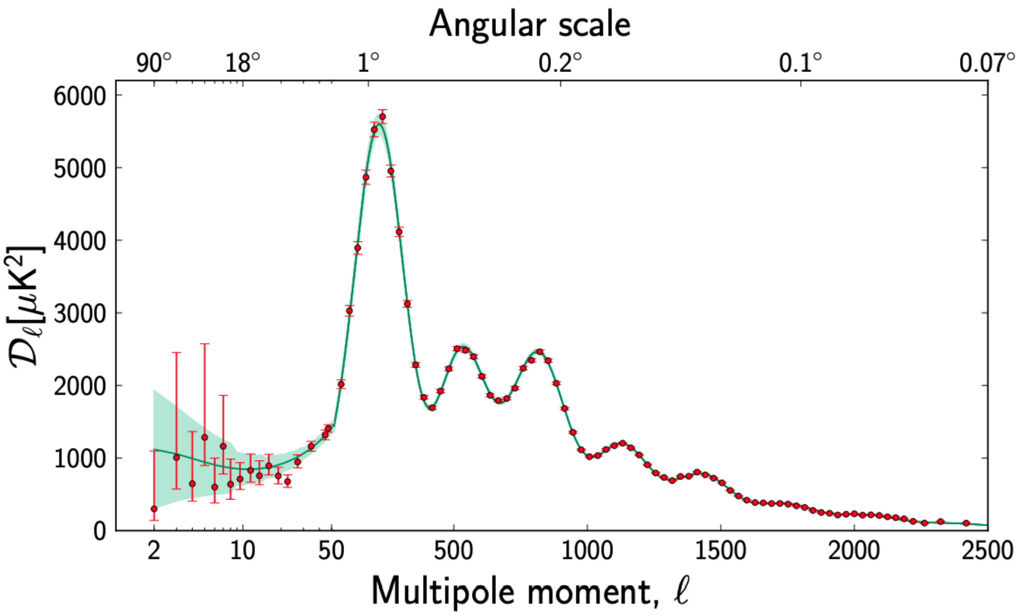


Figure 1.6: Image showing the angular power spectrum of the Cosmic Microwave Background (CMB) temperature fluctuations, as measured by the Planck collaboration. Green line shows the acoustic peaks of the ΛCDM model. Image taken from [20].

monic wavelengths of $\lambda = \pi deg/l$ can be used to characterise these CMB fluctuations, with small values in the multipole moment l corresponding to large distances in the sky, whilst large values of l corresponding to smaller distances. These harmonics can then be used to determine the wavelengths of the modes of the CMB on a sphere. By looking at the mapping at different scales, a power spectrum can be derived from the CMB background, thus showing the temperature fluctuations as a function of angular size.

The density fluctuations within baryonic matter are imprinted onto the CMB moments before recombination on the surface of last scatter. These arise from the effects of gravity and radiation pressure.

Figure 1.6 shows the power spectrum of the universe, as measured by the Planck collaboration [18]. The location of the first peak can be used to find the total energy density, with a multipole moment $l \sim 220$ indicating a flat universe due to the sound horizon at last scattering [21]. If light were to travel in a straight line (as would be expected in a flat universe), then a Doppler peak would be found at 1^o [22], which can be seen in figure 1.6. Peaks 2 and 3 of figure 1.6 can be used to give information regarding the baryon density and dark matter density in the universe [21].

By scanning the sky between the microwave to sub-millimetre spectrum, the Planck collaboration were able to use the power spectrum from the CMB radiation to calculate the baryon density $\Omega_{\mathrm{b}}\mathrm{h}^2 = 0.0224 \pm 0.0001$ and the dark matter density, $\Omega_{\mathrm{c}}\mathrm{h}^2 = 0.120 \pm 0.001$ [18], where h is a normalising constant given in equation 1.9, and H_o the current Hubble constant [18].

$$h = \frac{H_o}{100km\, s^{-1}Mpc^{-1}} \tag{1.9}$$

The Lambda Cold Dark Matter Model of Cosmology (ΛCDM) model is considered the most accurate model of the universe currently. The model can be used to determine the existence of non-baryonic and non-relativistic (cold) particles that constitute the existence of dark matter in the universe.

By interacting predominantly through gravitational couplings, and not electromagnetically, non-baryonic dark matter played an important role in the early formation of the universe. Cold dark matter would clump due to small density perturbations growing due to Jeans instabilities [23], whilst baryonic matter resisted clustering for longer because of its coupling to radiation. This explains large scale structures [24] existing today within the CMB. However, if the dark matter in the early universe was relativistic, it would fail to form these gravitational wells, and the observed large scale structure would be suppressed. This class of dark matter is referred to as cold dark matter (CDM).

This damping would lead to no large scale structure being observed in the universe [25]. Since large-scale structure formation is seen within the CMB, dark matter particles must have been non-relativistic in the early universe [26].

By looking at the information within the CMB and the power spectrum derived from it, a universe with $\sim 84\%$ of its total mass being dark matter is found.

1.1.5 Summary

Looking at all the observational and cosmological pieces of evidence indicates a universe consisting of dark energy, dark matter and baryonic matter. Whilst the Standard Model of Particle Physics can be used with a high degree of accuracy to model the baryonic matter found within the universe, it does not offer any suggestions as to the structure of dark matter. This leaves more than $\sim 85\%$ of the mass of the universe unaccounted for.

1.2 Dark Matter Candidates

There are three main categories of solution to the dark matter problem. These arise through the introduction of new particles not previously described by the Standard Model of Particle Physics; astrophysical objects such as primordial black holes; and proposing new theories of cosmology through alternative gravity. This section will look at the first category and will detail proposed candidates such as Axions, Neutrinos and Weakly Interacting Massive Particles (WIMPs).

1.2.1 Axions

Axions are pseudo-Nambu-Goldstone bosons [27], and were first proposed by Peccei and Quinn (PQ) as a way to reconcile the strong Charge-Parity (CP) problem [28]. Nambu-Goldstone bosons are massless bosons which arise when the global PQ symmetry is spontaneously broken at high mass scales. When adding in QCD effects, the axion attains a small mass component, thus becoming a pseudo-Nambu-Goldstone boson [29][27]. Within a mass range of 10^{-6} to 10^{-2} eV, axions produced in the early universe could be a solution for the abundance of dark matter.

The Axion Dark Matter Experiment (ADMX) collaboration uses a microwave cavity experiment to look for the currently undetected axion by probing the axion-photon coupling with a strong magnetic field [30]. Since Axions are expected to convert to monochromatic photons, these can be detected using an antenna

[31].

1.2.2 MACHOs and Primordial Black Holes

Another source that could impact the size of the undetected matter in the universe is from Massive Astrophysical Compact Halo Objects (MACHOs). These are large objects within galaxies that emit little radiation, thus being difficult to detect directly. They could range from faint stars, star remnants and substellar objects; with searches via microlensing indicating that they do contribute to the unseen mass of galactic halos. However, these alone cannot be used to explain the large mass difference of $\sim 25\%$ of the galactic halos [32] [33] [34].

Primordial Black holes (PBH) could also contribute to dark matter, with the black holes being produced before Big Bang Nucleosynthesis. However, since these have masses below the sensitivity of microlensing surveys [35], looking for them has been challenging, with predictions showing that primordial black holes only contribute a small fraction of the mass of a galaxy [36].

In 2016, the Laser Interferometer Gravitational-Wave Observatory (LIGO) was able to observe the gravitational waves produced by the collision of two black holes, with the masses of both black holes greater than 30 solar masses [37]. This observation by LIGO could indicate a new species of black holes which were formed in the early universe. These new species of black holes could contribute to the matter associated with dark matter if the mass range is between $20\mathrm{M_s} \leq \mathrm{M} \leq 100\mathrm{M_s}$ ($\mathrm{M_s}$ being one solar mass) [38] [39].

1.2.3 Sterile Neutrinos

The standard model neutrino was an early candidate for dark matter, as it is stable, weakly-interacting and long-lived. However, by simulating relativistic neutrinos in the early universe, large scale structure formation is not seen, hence indicating that a new form of a particle is needed [40].

This could call for a new type of neutrino called a sterile neutrino which would be

Majorana in nature with the mass arising from a seesaw mechanism [41]. These sterile neutrinos could be postulated to only interact with standard model particles via a small mixing angle, and could be a simpler explanation to what constitutes dark matter [42].

1.2.4 WIMP Dark Matter

Another candidate for a particle that could constitute dark matter is Weakly Interacting Massive Particles (WIMPs). These are theorised particles beyond the standard model which are stable, long-lived, massive and non-relativistic. These particles would also interact weakly with the current standard model particles, and gravitationally, as well as possibly via another force carrier.

WIMPs are a strong candidate for constituting dark matter, as due to early universe physics, WIMPs initially modelled to be in chemical equilibrium result in the correct abundance of Ω_{CDM}, the cold dark matter density in the early universe.

In the early universe, the particles contained within this universe all exist within a thermal plasma [43]. Since dark matter is assumed to be in thermal equilibrium with baryonic matter in the early universe due to these high temperatures, after expansion and cooling, these reaction rates would fall below the threshold required to remain in thermal equilibrium. What is left is a universe with the relic abundance of dark matter equal to what is observed today. Assuming an annihilation cross-section on the order of the weak scale, a dark matter density similar to what is observed cosmologically would be observed [44].

Extensions could be made to the family of particles in the standard model using supersymmetry (SUSY), which introduces a symmetric "super-partner" to the particles within the standard model. These super-partners have the same internal quantum number as their standard model counterparts, with their spin differing by one-half. Hence every fermion has an associated supersymmetric bosonic partner, and for every standard boson, there is an associated supersymmetric fermionic partner [45]. The lightest neutralino provides a good WIMP candidate for particle dark matter, having been created with other SUSY particles in the early universe [46].

The past decade of research into WIMP searches within the GeV mass range has excluded a large amount of parameter space. The results from the LHC indicate a lack of observations of non-standard model particles thus further constraining existing models [47] [48] [49], with the parameter space required for WIMPs to give the correct relic abundance decreasing with these new limits [50]. This has meant future searches for dark matter are now focusing towards constraining WIMP masses within the sub-GeV range. These will be discussed in the following section.

1.3 WIMP Detection

There are three methods of detecting WIMP dark matter particles. The first of which is indirect detection of dark matter. These mainly consist of astrophysical searches in places of high dark matter density such as the core of galaxies or galaxy clusters [22]. These searches look for the annihilation of dark matter particles to produce detectable quantities such as gamma-rays, electron-positron pairs, and neutrinos.

WIMP searches can also be conducted in production-decay experiments at detectors, without relying on the hypothesised local halo abundance. This avenue of dark matter detection relies on detecting any missing transverse energy, as any dark matter particles produced in a collision is expected to leave the detector without being detected. An example of this are the events produced at the LHC detectors, where any dark matter produced would leave the detector due to not interacting electromagnetically [51].

The final option, which is the detection of WIMPs from the local halo undergoing collisions with a dedicated target in a direct search experiment apparatus, is the subject of the remainder of this thesis. In particular, discussions for the LZ experiment will be discussed.

1.3.1 Direct Detection Principles

The main principle behind direct detection is that dark matter particles can interact with atomic nuclei by scattering elastically. Because WIMPs are massive particles, their scattering process can be modelled as a non-relativistic two-body scattering problem [52]. Direct detection relies on the dark matter present in the Milky Way, with the distribution of dark matter in the Earth's neighbourhood affecting the rate of interaction. The density and velocity of WIMPs and relative motion of the Earth with respect to the dark matter halo is needed to estimate the expected scattering rate, with the halo assumed to be a standard isothermal and Gaussian, and with a Maxwellian velocity distribution assumed [53].

WIMP density and velocity

Understanding the local WIMP dark matter density is necessary for direct detection dark matter experiments. ρ_o, which is the local WIMP dark matter density, gives the average mass of WIMPs over a few hundred parsecs [54]. The WIMP density is directly proportional to the expected differential rate, with uncertainties in the density affecting the constraints on scattering cross-sections of WIMP dark matter.

The local WIMP density $\rho_{\mathrm{o}} = 0.3$ GeVcm^{-3} is assumed when comparing different direct detection results, with this estimate being accurate to $\sim O(2)$ [55]. However, recent Gaia estimations give a value of $\rho_{\mathrm{o}} \sim 0.47$ GeV cm^{-3} [56].

The WIMP velocity distribution is calculated using the local circular velocity of the sun around the galaxy core $\upsilon_{\mathrm{c}} = 220$ kms^{-1}, and the local escape velocity $\upsilon_{\mathrm{esc}} = 544$ kms^{-1} and the mean velocity of the earth $\upsilon_{\mathrm{earth}} = 245$ kms^{-1} relative to the dark matter halo.

Using these values, a Standard Model Halo (SHM) can be assumed, which is modelled to be an isothermal sphere with an isotropic Maxwellian velocity distribution [57], and can be written as:

$$f(\upsilon) = \frac{1}{(2\pi\sigma^2)^{1/2}} exp(-\frac{|\upsilon|^2}{2\sigma^2}) \tag{1.10}$$

where σ is the speed dispersion, and is related to the local circular speed as $\sigma = \rho_c\sqrt{3/2}$ [58]. Although including baryon physics within the velocity distribution would alter the shape of the velocity distribution away from being Maxwellian, [59] showed that these would not alter the results when only a Maxwellian velocity distribution is used.

WIMP-Nucleon scattering

Although WIMPs are theorised to only interact weakly with standard model particles, a large flux of WIMPs in the Milky Way halo allows for a small percentage of interactions to occur elastically with atomic nuclei. The cross-section σ is needed when calculating the interaction between a WIMP and a nucleus,

$$\sigma_o = A^2 \left(\frac{\mu_N}{\mu_n}\right)^2 \sigma_n \tag{1.11}$$

with σ_o being the spin-independent zero-momentum cross-section which describes the coherent interaction between a WIMP and the whole nucleus, A being the atomic number of the target nucleus, μ_N and μ_n being the WIMP-nuclear and WIMP-nucleon reduced-mass respectively, and σ_n the scalar WIMP-nucleon cross-section. This indicates that the scattering rate is larger for larger atomic nuclei, due to the A^2 dependence.

The differential recoil rate can then be given as:

$$\frac{dR}{dE_R} = F^2(q)\frac{\rho_o\sigma_o}{2m_W\mu_N^2}\int_{v_{min}}\frac{f(\upsilon)}{\upsilon}d\upsilon \tag{1.12}$$

where ρ_o is the local WIMP density, with other values described above. The $f(\upsilon)/\upsilon$ term takes into account the velocity of WIMPs within the dark matter halo to the relative motion of the earth, with the solutions to the velocity integral found in [60], and an example of which can be found in equation 1.13

$$\int_{v_{min}} \frac{f(\upsilon)}{\upsilon} = \frac{1}{\upsilon_o y} \tag{1.13}$$

with $y = \frac{\upsilon_{earth}}{\upsilon_o}$, and $\upsilon_o = \sqrt{\frac{2}{3}}\sigma_o$.

$F^2(q)$ is a form factor [61], and accounts for the suppression of event rates for heavier nuclei, which is dependent on the momentum transfer $\mathrm{q} = \sqrt{2\mathrm{m_N E_R}}$. The form factor must be included to take into account the inability to approximate the nucleus as a homogeneous sphere.

The minimum velocity υ_{min}, is defined as the minimum velocity of a WIMP which can be detected by a given detector threshold $\mathrm{E_{thr}}$ of a detector [62], and is given by

$$\upsilon_{min} = \sqrt{\frac{m_W E_{thr}}{2\mu_N^2}} \tag{1.14}$$

Since the recoil rate depends on the mass of the WIMP as well as the mass of the target nuclei, a comparison can be made to compare the rate for different target nuclei. Figure 1.7 shows the comparison for the event rate of a 100 GeV/c^2 WIMP with a cross-section $\sigma_\mathrm{o} = 10^{-45}$ cm^{-2} for different nuclei. It can be seen that xenon has a higher interaction rate compared to other materials, due to the A^2 term in equation 1.11 contributing to a larger cross-section for xenon compared to argon, thus making it a good target nucleus to detect WIMP interactions.

1.3.2 Direct Detection Experiments

The main aim of a dark matter direct detection experiment is to find the interaction signal caused by a dark matter particle within the dark matter halo of the Milky Way, as the Earth travels through it. This signal can produce three distinct signatures which are detectable: scintillation, ionisation, and heat (phonons), each of which is referred to as a detection channel. Since the first proposal of direct detection experi-

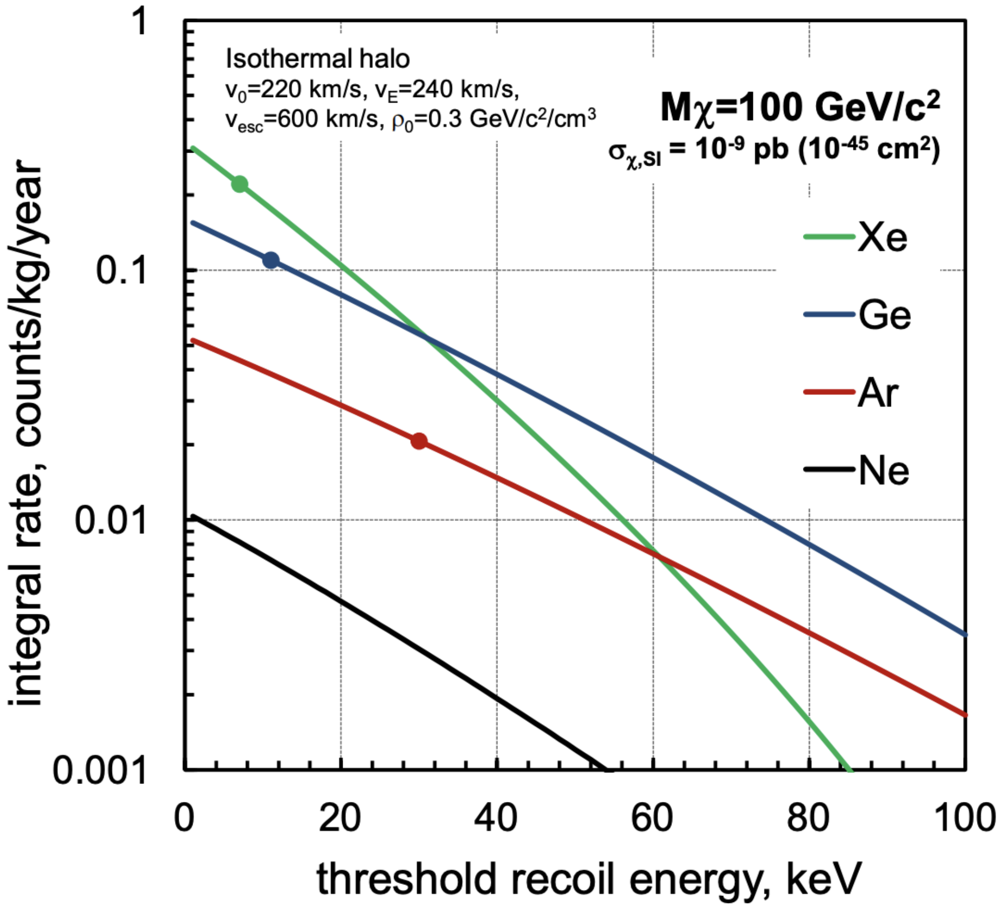


Figure 1.7: Plot showing the WIMP-nucleon scattering against recoil energy for a 100 GeV WIMP, with $\sigma = 10^{-45}$ cm^{-2} for Xe, Ge, Ar and Ne. A perfect energy resolution and an isothermal dark matter halo model is assumed. Coloured dots indicate the current thresholds for each technology. Image taken from [63].

ments in the 1980s [52], many experimental concepts have been developed to probe WIMP-nucleon interactions, and detect at least one of the three signatures.

The energies of an expected WIMP interaction are of the order $\sim$ O$(10 - 100$ KeV$)$, with the expected rate for a WIMP mass of $\sim$ 100 GeV/c^2 is $\leq$ 1 event kg^{-1} year^{-1} [64]. Thus the challenge in designing an experiment to detect such an event signal is to reduce backgrounds as much as possible. This is accomplished by placing experiments underground to reduce backgrounds originating from cosmogenic sources (discussed in chapter 4 and 5), as well as reducing any trace radioactivity from the detector components used in the detector construction.

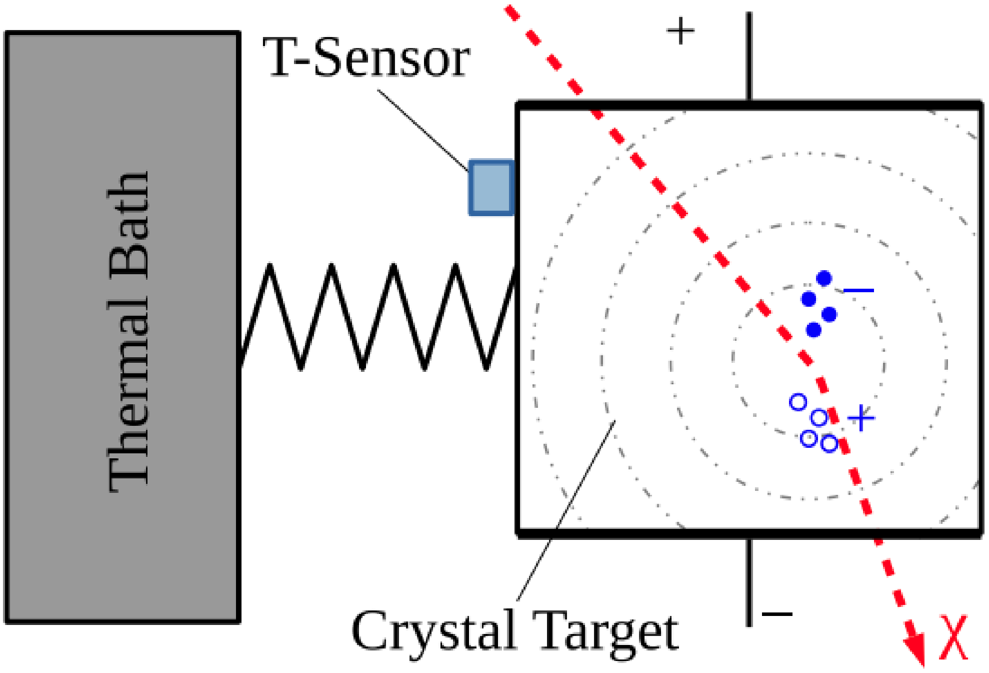


Figure 1.8: Image showing the principle of cryogenic detectors. They are cooled to sub mK-temperatures. The presence of a thermal bath allows the detector to be weakly coupled to it. Red line shows the trajectory of a particle χ interacting with the dielectric crystal. Image taken from [65].

Combining experimental methods for the different signatures produced by a dark matter interaction with a target nucleus can allow different interactions to be distinguished. Nuclear recoils (NR) can be caused by the neutrons and the daughter nuclei produced in radioactive decays of particles within the detector, whilst electron recoils (ER) can be caused by gamma radiation.

Cryogenic and Solid-State Detectors

Some categories of detectors are operated below room temperature to reduce noise or to enable use of certain targets. For example, noble liquids such as xenon must operate at $\sim$ 90 K. Other detectors, such as those sensitive to phonons from recoils may need to operate much colder, $\sim$ 4 K. Cryogenic detectors are an example of such detectors. They look for the detection of heat, and the scintillation light produced due to ionisation during a particle interaction [65]. Figure 1.8 shows an image of a cryogenic detector.

By using a different combination of signal channels (heat and ionisation), cryogenic detectors can determine between electron and nuclear recoils. This is seen in the EDELWEISS [66] and the CDMS experiments [67]. Because these experiments use crystals as their target material, scaling up to larger masses makes it difficult, thus achieving a higher exposure is challenging.

The SuperCDMS experiment [67], based in the Soudan Mine, US, uses a germanium detector to detect both ionisation and phonons. An electric field is applied to the germanium detector, thus allowing the electrons to be extracted from the surface. Sensors are placed at the faces of the germanium crystal to detect phonons. The 2018 results gave an upper limit of 1.4×10^{-44} cm^2 for a 1690 kg-day exposure for the WIMP-nucleon spin-independent cross section [68].

Noble Gas-Liquid Detectors

Detectors made from liquified noble gases for direct detection dark matter experiments are some of the most sensitive detectors to detect WIMPs within the $(10-100)$ KeV range. Two main detector mediums, liquid-xenon (LXe) and liquid-argon (LAr) have seen more prevalent uses compared to other noble gases when looking for scintillation and ionisation signals. With the ability to scale up to larger fiducial masses easier with LXe and LAr - as the fiducial mass is dependent on the amount of LXe or LAr within the detector, they are an excellent choice for experiments probing dark matter and neutrinoless double beta decay.

Experiments employing LXe in particular as the detector medium can offer excellent [x,y,z] vertex reconstruction and ER/NR discrimination by using a two-phase time-projection chamber configuration such as LUX [69], XENON [70] and the ZEPLIN (II/III) experiments [71] [72]. The scintillation light signal (S1) is detected using photomultiplier tubes (PMTs) placed at the top and bottom of the detector, with LXe scintillating at $\sim$ 178 nm, and the PMTs having a quantum efficiency of 30.6% [73]. By placing an electric drift field of known strength, the ionisation charge produced during an interaction can be transported to the gas region, where it

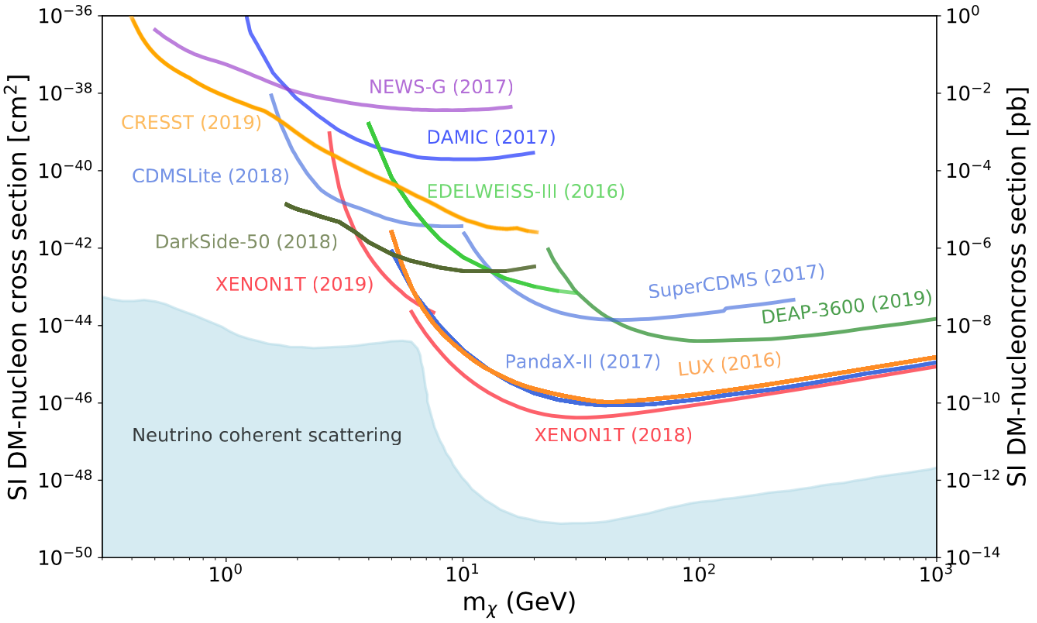


Figure 1.9: Plot shows the upper limits at 90% Confidence Level (CL) on spin-independent WIMP-nucleon cross sections for a range of WIMP masses for the leading dark matter experiments (as of end of 2020). Larger masses are constrained by noble gas TPCs (red and orange lines), with cryogenic crystals (yellow line) give the best limit for low WIMP masses. Blue shaded region shows the irreducible background from neutrinos (often denoted as the neutrino-floor) [74]. Image taken from [75]

then induces an electroluminescence light (S2) which is also detected by the PMTs. LXe also allows for fiducialisation, whereby ERs and wall backgrounds are reduced by rejecting events outside a central detection region. PMTs, the signals detected and their quantum efficiencies are discussed in further detail in chapter 2.

Equation 1.11 shows that the cross-section is dependant on the atomic number of the atom used as a medium for WIMP interactions. Since xenon has a larger atomic number compared to argon, the A^2 term allows for a better cross-section to be achieved. The current best limits for SI-WIMP-nucleon interactions are shown in figure 1.9 showing limits placed on a large range of WIMP masses, with detectors employing LXe technology (red and orange lines in figure 1.9) giving the best limits over the past decade.

1.3.3 Next Generation Experiments

Construction of the next generation (Generation-Z) of LXe direct detection dark matter experiments is complete, with results from the XENONnT [76] and LUX-ZEPLIN (LZ) [77] experiments expected in late 2021/early 2022. This shows the promise of LXe technology, with results expected to close much of the parameter space to the neutrino floor for WIMPs of mass greater than 10 GeV. The neutrino-floor is caused due to nuclear recoils from neutrino-nucleus scattering [74]. This reduces the improvement in sensitivity for WIMPs within that region. This can already be seen at low WIMP masses ($m_\chi \ \leq 10$ GeV/c^2), where the presence of the neutrino backgrounds caused by ^{8}B is limiting the sensitivity to WIMPs. The next chapter will detail the LZ experiment, and will be the main topic of discussion throughout this thesis.

Further beyond these experiments, work has already started on the next generation of experiments using LXe TPC technology, with the DARWIN collaborations [78] merging to build a 'Generation-3' (G3) experiment with over 50-tonnes of LXe being employed. With such a large active mass region, DARWIN plans to penetrate into the neutrino floor produced by atmospheric, solar and supernova neutrinos to definitively test the standard WIMP hypothesis [79]. DARWIN will also have sensitivity to well-motivated alternative (non-WIMP) models of dark matter, exotic neutrino physics, and Beyond Standard Model (BSM) searches including neutrinoless double-beta decay with world-leading sensitivity.

Chapter 2

The LUX-ZEPLIN Experiment

Verily, We have created all things with due proportion

Quran [54:49]

The LUX-ZEPLIN (LZ) experiment is a direct detection dark matter experiment which uses a two-phase liquid xenon time-projection chamber (LXe TPC). Since the origins of time-projection chambers from the 1970s, this technology has been pushing the boundary of direct detection WIMP searches [65], with the LUX experiment one of the first to report strong limits on WIMP dark matter.

The next generation LZ experiment is located at the Davis Cavern in the Sanford Underground Research Facility (SURF) in Lead, South Dakota USA. It is placed 4850 feet underground, and aims to build upon LUX to set more stringent limits on WIMP-nucleon interactions in order to detect dark matter. This section will introduce particle interactions in liquid and gaseous xenon, the theory behind signal generation, as well as looking at the LZ detector and some of its design features.

2.1 Liquid-Xenon Time Projection Chambers

LXe TPCs have been improving the sensitivity to WIMP-nucleon interactions, with many experiments employing the same technology over the years. Since their origins in the 1970s [80] [81], LXe TPCs use liquid and gaseous xenon as a scintillation medium to measure particle interactions. This has allowed strong limits to be place on spin-independent and spin-dependent WIMP-nucleon cross-sections for a range of WIMP masses. As the main factors which determine the sensitivity of TPCs is the background rate of the nearby surroundings as well as the amount of xenon within the experiment, scale-up of liquid xenon experiments is possible so long as the purity of the target and the granularity of the readout can be maintained at scale.

Xenon is a good material to be used in TPCs for low-background WIMP searches. It is sensitive to WIMP-nucleon cross-sections of the order of $(10-100$ GeV$)$ because of its relatively high atomic number $A = 131$, as well as not having any long-lived radioactive isotopes. Isotopes with an odd number of neutrons render such detectors sensitive to spin-dependent WIMP-nucleon interactions [82].

Particle interactions with LXe atoms produce prompt scintillation photons as well as ionisation electrons. Since xenon is transparent to its scintillation light, the signal created is detectable, with low energy interactions producing a measurable signal [63].

Having a large atomic number also allows xenon to have excellent self-shielding. This is where backgrounds originating from outside the detector or from surfaces within the instruments are stopped in a short distance within the liquid xenon.

In general, dual-phase LXe TPCs are constructed with a large mass of liquid xenon placed under an electric field, with a small gas region at the top of the detector. Scintillation photons are detected with photomultiplier tubes (PMTs) placed at the top and bottom of the liquid-gaseous volume. The presence of the vertical electric field helps extract the ionisation electrons to the gas region, with the strength of the electric field being several 100s V/cm. Upon reaching the liquid-gaseous surface, these ionisation electrons are accelerated within the gas region, thus producing electroluminescence light. This light is also detected by the PMTs. By combining the signals of the scintillation light with the time delay of the electroluminescence light,

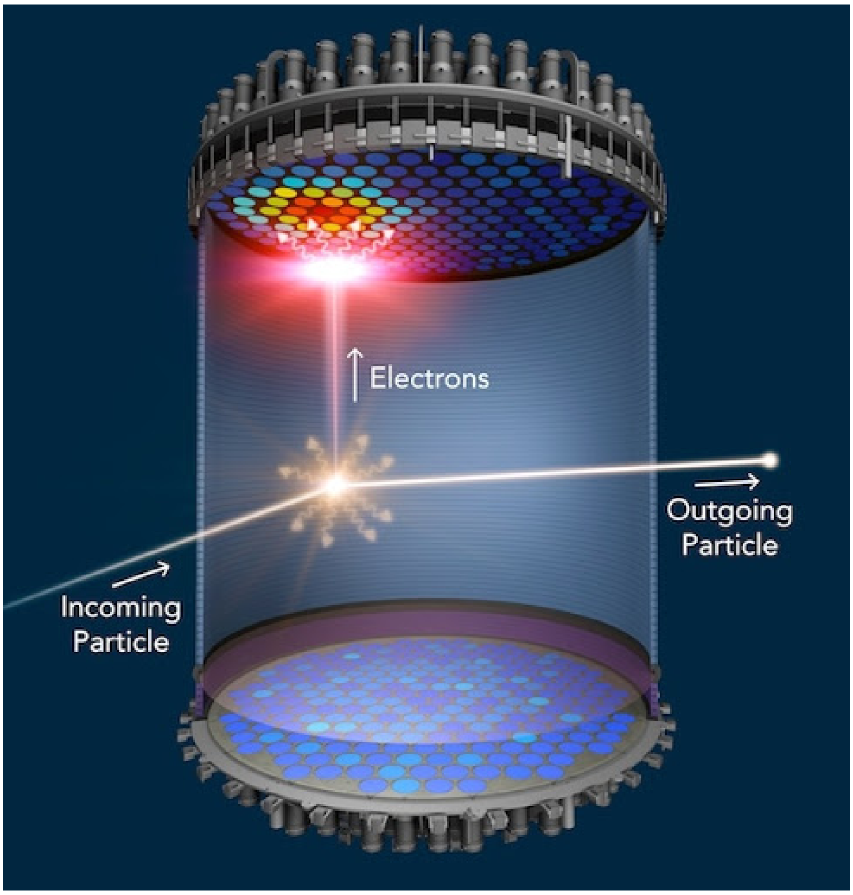


Figure 2.1: Schematic image showing a particle interaction inside the LZ TPC. The incoming particle produces the S1 signal (shown in orange) with the electrons drifting towards the gas layer via an applied Electric field to produce the S2 signal. PMTs are placed at the top and bottom of the detector. 3D position reconstruction is achieved by looking at the hit pattern of the PMTs at the top for x-y reconstruction, with the z-position determined by the time difference between the S1 and S2 signals [73].

good XYZ position reconstruction and background discrimination can be achieved. Position reconstruction and background discrimination will be discussed further in chapter 5.

These will be discussed further in the following sub-sections.

2.2 Particle interactions and Detection in LXe TPCs

Interactions within a LXe TPC produce two characteristic recoils separately, both of which are used to determine the nature of the interacting particle. Backgrounds originating from radioactive isotopes emitting beta or gamma particles undergo electron recoils predominantly resulting in electron recoils. An electron recoil (ER) is the collision of the incoming particle with one of the electrons of a xenon atom in the target. Neutral particles like WIMPs or neutrons undergo nuclear recoils (NR), where the incoming particle collides with the atomic nucleus of a xenon atom. The recoiling particle in both ER and NR interactions scatter with neighbouring electrons and nuclei of other xenon atoms, thus causing scintillation photons, ionisation electrons and heat. Current TPC designs are only capable of detecting scintillation photons and ionisation electrons; the heat produced in collisions dissipates in the target without detection.

2.2.1 Scintillation (S1) Signals

Primary scintillation (S1) signals are generated following the collision of the incident particle with an electron or nucleus in the target. Either NR or ER processes result in a xenon atom in an excited state. This excited xenon atom forms a molecular dimer with a neighbouring xenon atom, with excited states with vibrational excitation different from purely electronic excitation. This excimer then de-excites to the ground state, thus producing a vacuum ultra-violet (VUV) photon. This can be seen by the following set of equations [63] [83]:

$$P_i + Xe \rightarrow Xe^* + P_i \tag{2.1}$$

$$Xe^* + Xe \rightarrow Xe_2^{*,\nu} \tag{2.2}$$

$$Xe_2^{*,\nu} + Xe \rightarrow Xe_2^* + Xe \tag{2.3}$$

$$Xe_2^* \rightarrow Xe + Xe + \gamma \tag{2.4}$$

whereby P_i is an incoming particle producing an electronic or nuclear recoil (e.g gamma ray or WIMPs), Xe a xenon atom, with ν being used to determine purely electronic excitation ($\nu = 0$) from vibrationally excited states, which mostly de-excite non-radiatively, but can emit infrared photons.

Equation 2.1 shows the excitation of a xenon atom, equation 2.2 is the formation of a dimer with an excited vibrational degree of freedom, equation 2.3 is the decay of a vibrationally excited dimer to a xenon molecule with an electron in an excited state only, and equation 2.4 is the decay of the excited xenon dimer to xenon atoms with the release of a gamma photon.

The second way in which scintillation photons are released is by the recombination of ionisation electrons [63] [83], with the end result being the release of a gamma photon the same as the process above, and can be seen below:

$$P_i + Xe \rightarrow Xe^+ + P_i + e^- \tag{2.5}$$

$$Xe^+ + Xe + Xe \rightarrow Xe_2^+ + Xe \tag{2.6}$$

$$Xe_2^+ + e^- \rightarrow Xe^{**} + Xe \tag{2.7}$$

$$Xe^{**} + Xe \rightarrow Xe^* + Xe + heat \tag{2.8}$$

$$Xe^* + Xe \rightarrow Xe_2^{*,\nu} \tag{2.9}$$

$$Xe_2^{*,\nu} + Xe \rightarrow Xe_2^* + Xe \tag{2.10}$$

$$Xe_2^* \rightarrow Xe + Xe + \gamma \tag{2.11}$$

where equation 2.5 shows the ionization of xenon, equation 2.6 shows the production of di-atomic xenon atoms when ionised xenon interacts with a neighbouring xenon atom, equation 2.7 shows the recombination of di-atomic xenon to produce an excited atomic state (Xe^{**}), equations 2.8-2.9 show the formation of a vibrationally excited dimer. The decay of the dimer (equations 2.10 and 2.11) follows an identical de-excitation path similar to the processes seen in equation 2.3 and 2.4, ending with the release of a gamma photon.

The number of ionisation electrons undergoing this process is a function of the initial recoil energy of the interaction and the strength of the electric field applied [84]. Having a stronger field drifts free electrons away from the interaction point, therefore constraining this recombination luminescence.

The VUV scintillation light released has a wavelength of 178 nm, with a FWHM of 14 nm. Impurities such as water molecules cause absorption of the light, necessitating the use of very high purity liquid xenon. In LZ, an absorption length of 30-100 nm is expected, with the scintillation photons (also referred to as the S1-signal) being detected at PMTs located at the top and bottom of the detector [63].

Higher energy interactions

Particles with a larger amount of energy (i.e. α particles $\geq$ MeV) can sometimes lead to processes of higher order. These higher order processes can decrease the number of scintillation photons released (hence a lower S1 signal is found). The particle tracks formed due to these higher energy interactions increases the density of the excited xenon atoms present, thus the probability of two excited atoms interacting is greater. Therefore a higher ionisation rate is found within the tracks formed. This process is called bi-excitonic quenching [85], with the process being:

$$Xe^{*} + Xe^{*} \rightarrow Xe^{+} + Xe + e^{-} \tag{2.12}$$

If the single excited xenon atom has a large enough energy to ionise a neutral xenon atom from the ground state, then the following process is found, where

$$Xe^{*} + Xe \rightarrow Xe^{+} + Xe + e^{-} \tag{2.13}$$

This process is called penning ionisation.

Both these processes can lead to a suppression of the S1 signal at higher energies, and will be an important feature when using Machine Learning to identify processes which have high energy interactions.

2.2.2 Secondary (S2) Signals

The remaining electrons which escape recombination drift upwards towards the liquid surface due to an applied vertical electric field, hence moving away from the point of interaction. An electric field of several kV/cm is applied at the liquid surface to allow the drifting electrons to overcome the potential barrier at the surface level, thus allowing them to escape into the gas phase. This electric field is generated by placing grids on either side of the surface level [86].

In the gas region, the electrons are accelerated to enable them to excite the xenon gas atoms within the gas region, with the de-excitation of these gaseous xenon atoms producing the secondary S2 (electroluminescence) signal. This S2 signal is detected by both the top and bottom PMT arrays, with the number of photons within the S2 signal dependent on the strength of the applied electric field within the TPC.

The S2 signal size is of orders of magnitude larger compared to the S1 signal, with a time delay between the two signals being used to determine the z-position of the interaction. This requires detailed information regarding the strength of the electric field applied.

In LZ, a drift field of $E = 310$ V/cm is used, hence electrons travelling from the bottom of the TPC (height = 1.46m) takes $\sim 800\ \mu s$ to reach the gas region at the top of the TPC.

The ionisation electrons within the TPC also diffuse in the transverse (x-y) plane, alongside the diffusion in the longitudinal (z) direction (D_L). This transverse diffusion (D_T) however has no first-order effect on the S2 signal size, with the ratio between both diffusions $\mathrm{D_L/D_T} \sim 0.1$. This is realised by an expected diffusion of 2.2 mm in the transverse direction [73].

However, impurities within the LXe can lead to uncertainties in the S2 signal size. This is as electronegative molecules such as O_2, H_2O or N_2O present in the LXe can affect the mobility of the ionisation electrons. By capturing the drifting electrons, these molecules can then become negative ions with a lower mobility. This is dependent on the concentration of the impurities within the LXe, with the impurities present within LXe by the outgassing of detector components. This shows the importance of purification of LXe, with contaminant levels ~O(ppb) or lower required for lower energy events.

2.2.3 Light and Charge Yields

Incoming particles producing either electronic or nuclear recoils deposit energy E. This energy can be expressed by the number of excitons, n_e, and electron-ion pairs, n_i, as

$$E = LW(n_e + n_i) \tag{2.14}$$

where L is the called the Lindhard factor, and accounts for the energy lost due to heat, and W in equation 2.14 is the energy needed to liberate a single electron or electron-ion pair, and is measured [84] to be

$$W = 13.7 \pm 0.2\ eV \tag{2.15}$$

For electronic recoils, an $L = 1$ is assumed, as all the energy is assumed to go into the electronics system. For nuclear recoils, the Lindhard model states that

$$L = \frac{kg(\varepsilon)}{1 + kg(\varepsilon)} \tag{2.16}$$

and describes the fraction of recoil energy which is transferred to electrons, with $g(\varepsilon)$ describing the ratios between electronic and nuclear stopping powers. k is a proportionality constant between the electronic stopping power and velocity of a recoiling nucleus [87], with the Lindhard model describing interactions with LXe atoms below 2 keV accurately [88].

Measurable Quantities

In order to parameterise the detected signals well, measurable quantities specific to LXe TPCs must be taken into account in order to measure the detectable S1 and S2 signals produced. Hence, the number of VUV photons n_γ and the number of electrons n_e released during an interaction can be expressed as:

$$n_\gamma = N_{ex} + N_i r \tag{2.17}$$

$$n_e = N_i(1 - r) \tag{2.18}$$

where N_{ex} is the total number of excited xenon atoms, N_i the total number of ionised xenon atoms, and r being the fraction of ions recombining. Hence, the size of the S1 and S2 signals (in terms of detected photons) can be written as:

$$S_1 = n_\gamma g_1 \tag{2.19}$$

$$S_2 = n_e \varepsilon N_{ph} g_{1,gas} \tag{2.20}$$

where g_1 is the light collection efficiency, with $g_{1,gas}$ the light collection efficiency in gas. N_{ph} is the number of electroluminescence photons released per electron, and

ε being equal to the electron extraction efficiency. Hence,

$$g_2 = \varepsilon N_{ph} g_{1,gas}$$

can be defined as the electron detection efficiency. Therefore the Energy E from equation 2.14 would become

$$E = LW \left(\frac{S_1}{g_1} + \frac{S_2}{g_2} \right) \tag{2.21}$$

Using calibration sources, both g_1 and g_2 can be measured, as was done in for the LUX experiment [89]. Both electron and nuclear recoils have a difference in the light and charge yields they produce, with the ratio of exciton to electron-ion pairs ≤ 0.2 [90] for ER events, and for NRs, the ratio being ~ 1. Hence, different calibration sources are used, with γ-rays being used for ERs. An example of this difference would be that an ER of energy 6 keV would produce the same charge and light yield as a $\sim$ 60 keV recoil from an NR.

2.2.4 Discrimination

For low-background direct detection experiments, it is important to understand the backgrounds which are present within the detector. These background levels must be kept low enough to ensure a good signal in the parameter space that is being observed. LXe TPCs allow for good distinction which allow for background reduction by being able to discriminate between ER and NR events.

When plotting events in S1-S2 space, ERs and NRs form two distinct bands due to the difference in ratios between ionisation electrons and scintillation photons. This can be seen in figure 2.2, with ER events being shown in blue, and NR events in orange. ER events leaking into the NR region (as shown by the blue points going into regions of the NR band) limit the discrimination power. LZ is expected to achieve an ER/NR discrimination of 99.5% [91].

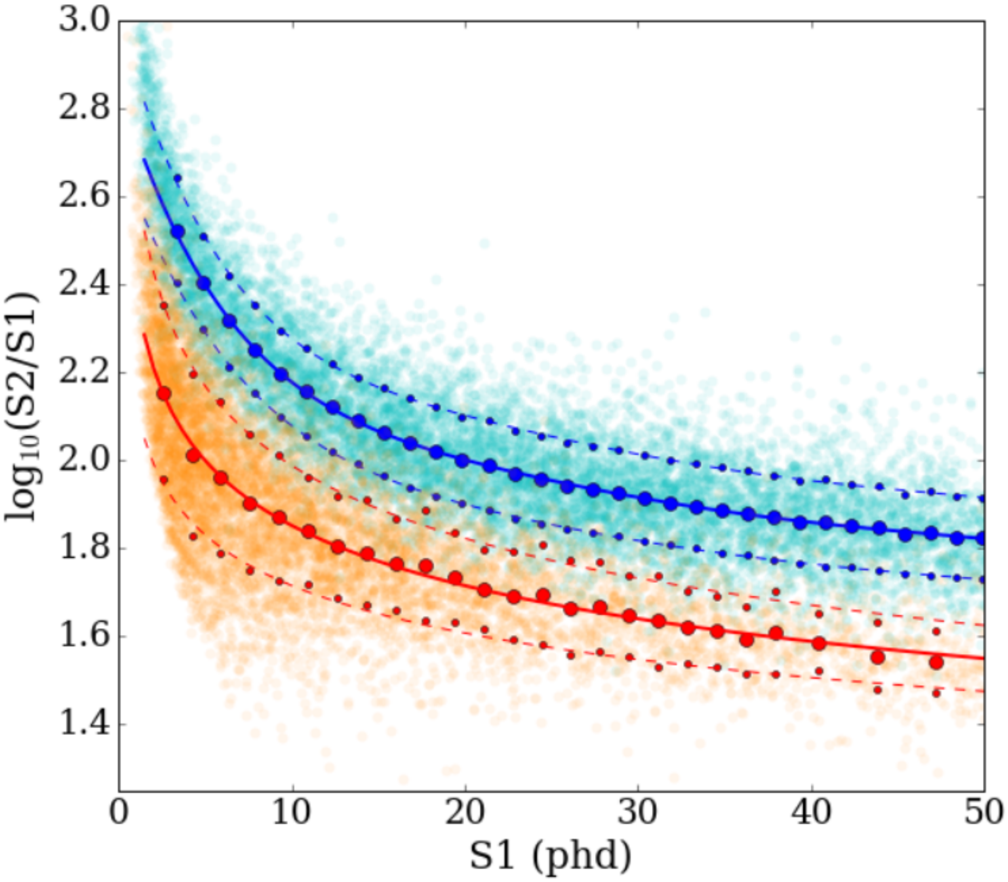


Figure 2.2: ER and NR discrimination, as taken from the LUX experiment, with the S1 signal measured in detected photons (phd). Cyan represents the ER band, whilst Orange shows the NR band. Filled circles show the Gaussian mean, whilst smaller circles show the $\pm\sigma$ to the Gaussian mean. Lines (solid and dashed) represent power law fits to the means and $\pm\sigma$. Image taken from [89].

2.3 The LUX-ZEPLIN Detector

The LZ Experiment is a worldwide collaboration of universities and academic institutions. The experiment is based at the Sanford Underground Research Facility in South Dakota, US. The detector is ~ 20 times larger than its predecessor LUX, but employs the same detector principles of a TPC for low background direct detection experiments. The active mass of liquid xenon used in LZ will be 7 tonnes, with the detector placed 4850 feet (~ 1500 m) underground within a water tank, to reduce backgrounds and maximise the sensitivity.

A LXe skin also exists which surrounds the TPC, as well as an Outer Detector

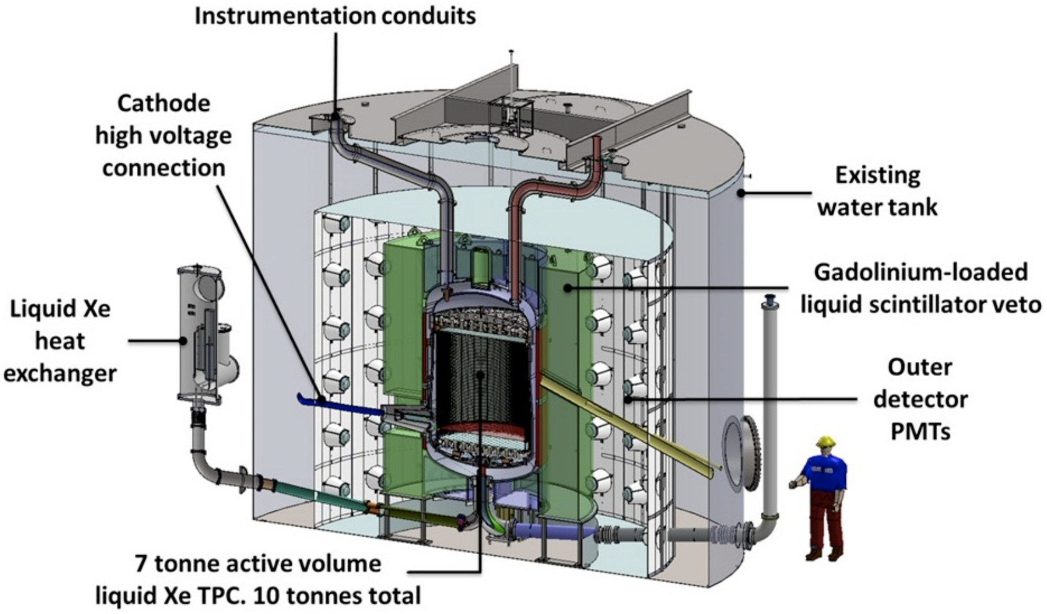

Instrumentation conduits
Cathode high voltage connection
Existing water tank
Gadolinium-loaded liquid scintillator veto
Liquid Xe heat exchanger
Outer detector PMTs
7 tonne active volume liquid Xe TPC. 10 tonnes total

Figure 2.3: Cross section of the LZ experiment, with human for scale purposes. Main detector components are labelled [73].

using gadolinium-loaded liquid scintillator (GdLS) to veto any events not within the region of interest (ROI). This section will highlight detector design of LZ, as well as the steps being taken for data taking to begin at the end of 2021.

2.3.1 The LZ TPC

The detectors used in LZ are placed within a cryostat made of double-walled vacuum-insulated titanium [92]. The Inner Cryostat Vessel (ICV) houses both the TPC and the LXe skin region, thus employing a total of 10 tonnes of liquid xenon, and is suspended inside the Outer Cryostat Vessel (OCV). The TPC is a cylinder of height and diameter equal to 1.46 m, and contains 7 tonnes of LXe, thus referred to the active region. The TPC walls are made of PTFE, with a reflectivity of $\geq 97.3\%$ when in LXe [93]. This ensures they are highly reflective thus increasing the light collection of the S1 and S2 signals as seen by the PMTs.

There are four stainless steel woven wire grids, to which high voltages are applied and placed within the TPC, thus creating three distinct field regions: the Extraction Field (electroluminescence) Region (EFR), the Drift Field Region (DFR) and the

Table 2.1: Dimentions of the LZ detector [91].

Parameter	Value
TPC height	1.46 m
TPC inner diameter	1.46 m
Active LXe mass	7000 kg
Xenon skin thickness	4.0 - 8.0 cm
Inner Cryostat diameter	1.58 - 1.66 m
Inner Cryostat height	2.59 m
Outer Cryostat diameter	1.83 m
Outer Cryostat height	3.04 m
GdLS tank outer radius	1.64 m
GdLS mass	17.3 tonnes
Water tank diameter	7.62 m
Water tank height	5.92 m
Water tank mass	228 tonnes

Table 2.2: Design voltages and field strengths for the grids used in the LZ experiment [73].

Grid/Region	Voltage/Field Strength
Cathode Grid	-50 kV
Gate Grid	-4 kV
Anode Grid	+4 kV
Electroluminescence Region	10.6 kV/cm
Drift Field Region	0.31 kV/cm

Reverse Field Region (RFR).

The DFR contains the drift field in which the active region of the LXe will be. It is between the cathode, which is located at the bottom of the detector; and the gate, which is located just below the liquid xenon surface level. This drift field region is the main region where the majority of the ionisation electrons will drift from after an interaction.

The EFR is located at the top of the detector, between the gate and anode grids, with the anode being placed within the xenon gas region, just below the top PMT arrays. The fields in the EFR are significantly higher in this region to allow for the electrons which drift upwards to increase the electron extraction efficiency, with $\sim$ 820 electroluminescence photons being released per electron in this region for the S2 signal.

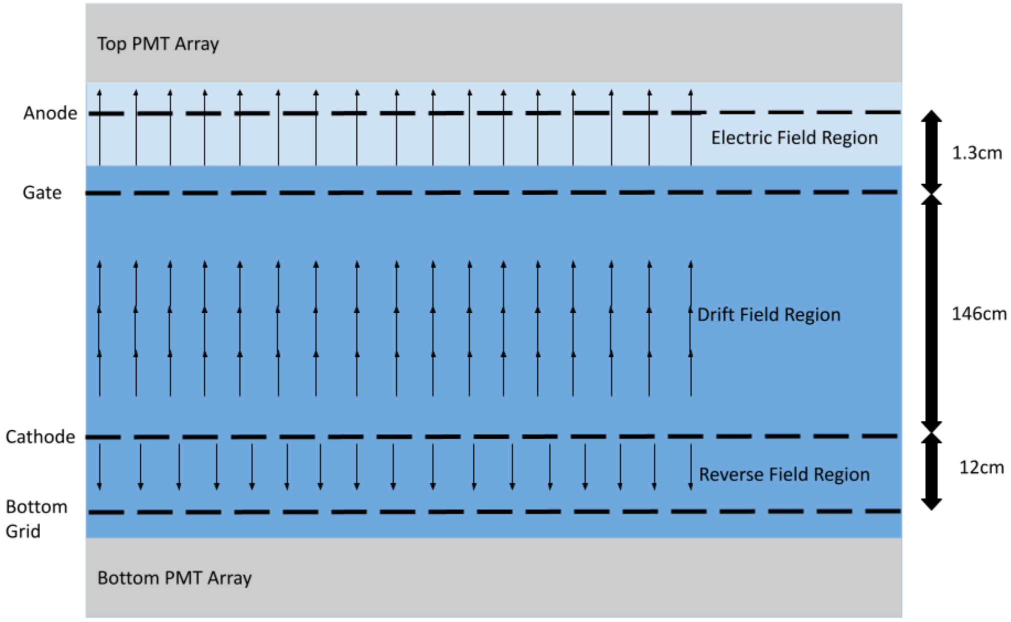


Figure 2.4: Cross section of the LZ TPC, with grids and approximate dimensions shown. Directions of the three main electric fields also shown (not to scale).

The RFR is between the cathode at the bottom of the TPC, and the bottom grid, with the field in the opposite direction to shield the bottom PMTs - which are immersed in the liquid xenon, from the strong electric field. Events which occur within this RFR region only have an S1 signal, as the ionisation electrons in this regions are not extracted. Figure 2.4 shows an image of the field directions, grids and the TPC.

2.3.2 PMTs

The LZ TPC has 494 photomultiplier tubes (PMTs) which monitor the active region of the LXe, with 253 PMTs located in the top array, and the remaining 241 located in the bottom array. The 3-inch PMTs (model no R11410-22) were developed by Hamamatsu, and were optimised for low-radioactivity searches in cold liquid xenon [73]. With an average quantum efficiency of 30.9% to VUV light, they were designed to have a low dark count and a high single photo-electron resolution.

The PMTs have 12 dynode stages, and operate at a nominal voltage of 1500 V, with a nominal gain of 5×10^6. The PMTs are installed in titanium support structures at both the top and bottom PMT arrays, with the layout of the PMT arrays optimised to achieve a high light collection efficiency.

The PMTs are laid out in a closed hexagonal pattern to maximise coverage of the TPC. Since S1 light is mostly detected by the bottom array, high detection efficiencies are needed for low-NR events, as well as giving the ability to discriminate between ER and NR events. The top array detects the S2-signal, and is also used for position reconstruction of events. To improve the position reconstruction accuracy for fiducialisation, the outermost PMTs in the top array are located above the TPC wall.

2.3.3 LZ Skin and Outer region

The skin region in the LZ experiment plays an important role in identifying any background events. The skin is a mass of LXe which sits outside the TPC, and serves as a dielectric insulation between the TPC and the ICV. 93 1-inch PMTs (model no R8520) and 20 2-inch PMTs (model no R8778) look for scintillation-only events caused by γ-rays, with a further 19 2-inch PMTs (model no R8778) observing the dome region. Since WIMPs would cause an energy deposition within the TPC by only interacting once with a xenon atom, any WIMP-like event (such as neutron scatters) within the TPC which cause other energy depositions in the outer regions must not be a WIMP, but an event caused by backgrounds.

The Outer Detector consists of 10 acrylic tanks which are filled with 17 tonnes of Gadolinium-loaded Liquid Scintillator (GdLS). It is near-hermetically sealed, and is a γ-ray and neutron veto system. It is surrounded by 120 8-inch PMTs (model n0 R5912) which are placed outside of the acrylic tanks. Linear alkylbenzene (LAB) is used as the liquid scintillator, with 0.1% of neutral Gadolinium added to improve the detection of neutrons [94].

If a WIMP-like event caused by a background such as a neutron interacts with a

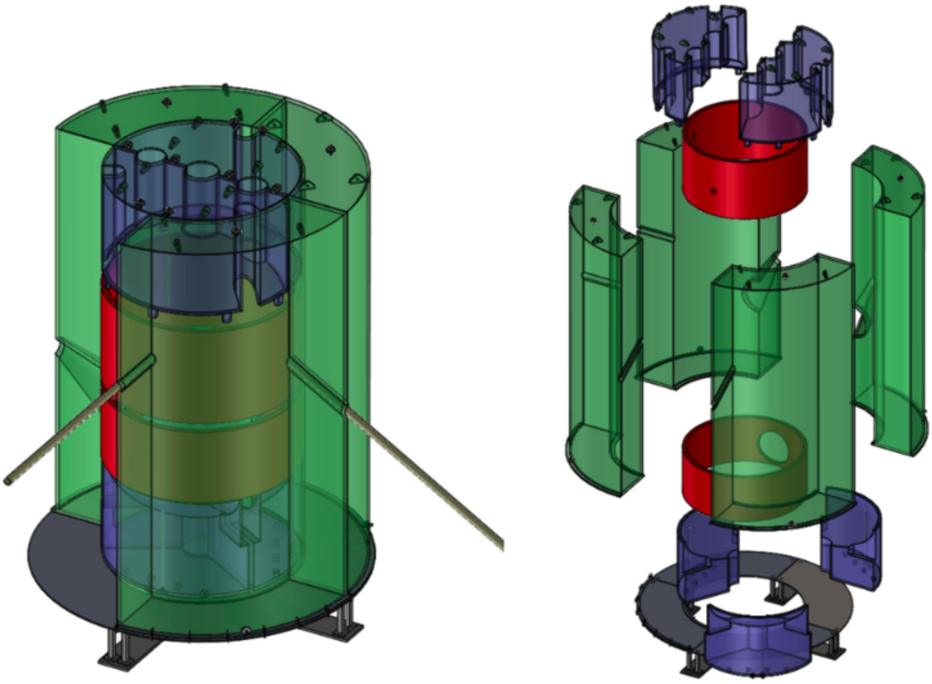


Figure 2.5: Schematic drawing of the acrylic tanks. Green tanks will cover the sides of the TPC, whilst the blue will cover the top of the TPC, and below the OD. Red shows the displacer cylinders [73].

xenon atom in the TPC, it is also likely to interact within the outer detector by capturing on the gadolinium. This would create a burst of several γ-rays of energy ~ 8 MeV, which then create a large signal within the liquid scintillator. This system works well in being a veto system to WIMP-like events, as the Outer Detector response time is within the time window of an event in the TPC [73], thus allowing an accurate correlation between NR events in the TPC.

2.3.4 LZ Xenon System

The xenon used within LZ must be kept at 175 K, and is maintained by using a thermosyphon system. These are a set of closed-loops heat pipes which use nitrogen as the process fluid. All the LXe which is held mainly within the TPC is drained via weir pipes and passes through a getter. It is then condensed back and re-circulated back into the detector (see figure 2.6). This system not only cools, but removes any impurities within the LXe which may have entered via outgassing of the materials

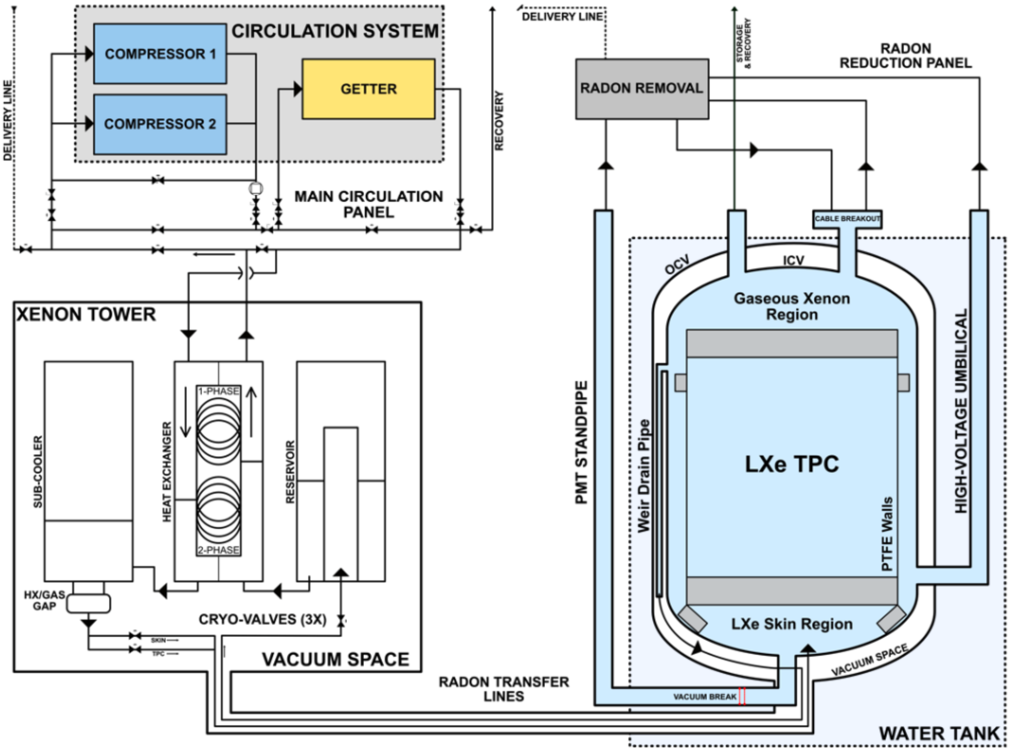


Figure 2.6: Overview of the circulation system used in LZ. The gas region in the TPC begins just after the wier used for draining the TPC. LXe travels from TPC to the xenon tower, which uses a heat-exchanger to heat the LXe. Gaseous xenon is pumped through a heated Zirconium getter, which then condenses back to the detector. The radon removal system removes radon by treating the Xe gas. Most of the xenon in LZ is held in the TPC [73].

used within the detector such as the PTFE panelling on the walls of the TPC. The total time taken to purify and cool all the LXe (10 tonnes) is approx 2.4 days, with a simplified image of the xenon system employed by LZ shows in figure 2.6.

2.3.5 Data Acquisition (DAQ) System

The signal generated from events which are captured from PMTs are amplified by analogue front-end electronics, with two channels per PMT used for different amplifications. A high-gain channel is used for low energy events, and amplifies the signal 40 times, whereas a low-gain channel is used for high energy events, and amplifies by a factor of 4.

Each signal is digitised at 100 MHz, and utilises a 14-bit resolution with a dynamic range of 2 V. Because of the longevity of LZ in being able to run for long exposure times, Pulse Only Digitisation (POD) is employed. This is where only the pulses and the minimum amount of baseline are digitised [95]. More details about the data which will be acquired throughout the LZ experiment will be presented in the next chapter.

2.4 Calibrations

An extensive calibration effort is employed to accurately measure the response of LZ. Both external calibrations (i.e. where calibration sources are placed outside the TPC) and internal calibrations (where sources are placed within the TPC) are employed, with the use of internal calibrations providing a way to calibrate the detector whilst overcoming the self-shielding of xenon. An overview of all the calibrations (both internal and external) is shown in table 2.3.

2.4.1 Internal Calibrations

For internal calibrations, specific isotopes of the relevant calibration material is injected into the LXe circulation system. This allows for uniform mixing of the isotope with the LXe in the TPC. These will be carried out for short-lived isotopes such as ^{83m}Kr, and hence be stored in solid form to allow for easier handling.

For longer lived isotopes such as ^{3}H, tritiated methane (CH_3T) is stored as pressurised gas, and is injected with the LXe into the TPC. This pioneering method was first used in the LUX experiment to calibrate the ER band to low energy thresholds. The purification system employed by LUX (and thus LZ) was able to remove any tritiated methane molecules from the TPC thus removing any backgrounds which may have been associated with it [96].

2.4.2 External Calibrations

Between the Inner Cryostat Vessel (ICV) and the Outer Cryostat Vessel (OCV), 3 vertical stainless steel tubes exist to allow for calibration sources to be deployed,

Table 2.3: The calibration sources used in the LZ experiment. The type of interaction, energy deposition, half-life, purpose and deployment is shown. Internal sources are sources which are gaseous, and are injected within the Xe circulation. Calibration sources (CSD) deploys sources between the ICV and OCV, and is mainly used for gamma and neutron sources. External sources are those which are located outside the OCV [91].

Isotope/Particle	Energy [keV]	half-life	Purpose	Deployment
^{3}H - β	18.6	12.5 y	ER band	Internal
^{14}C - β	156	5730 y	ER band	Internal
^{83m}Kr - γ	9.4, 32.1	1.83 h	TPC (x,y,z)	Internal
^{131m}Xe - γ	164	11.8 d	TPC (x,y,z), Xe skin	Internal
^{220}Rn - (α,β,γ)	various	10.6 h	Xe skin	Internal
^{22}Na - γ	511, 1275	2.61 y	TPC and OD	CSD
^{54}Mn - γ	835	312 d	ER response	CSD
^{57}Co - γ	122	0.74 y	Xe skin	CSD
^{60}Co - γ	1173, 1333	5.27 y	ER response	CSD
^{133}Ba - γ	356	10.5 y	ER response	CSD
^{228}Th - γ	2615	1.91 y	ER response	CSD
^{124}AmLi - (α,n)	1500	432 y	NR band	CSD
^{124}AmBe - (α,n)	11000	432 y	NR band	CSD
^{252}Cf - n	Watt spectrum	1.65 y	NR efficiency	CSD
^{88}YBe - (γ,n)	152	106 d	NR response	External
^{124}SbBe - (γ,n)	22.5	60.2 d	NR response	External
^{205}BiBe - (γ,n)	88.5	15.3 d	NR response	External
^{206}BiBe - (γ,n)	47	6.24 d	NR response	External
DD - n	272 - 400	-	NR light/charge yield	External
DD - n	2450	-	NR light/charge yield	External

thus allowing sources specific to neutron and gamma calibration to be introduced. These sources are used to calibrate the NR band, and characterise the response of the system.

For external calibrations, a photo-neutron source was used to calibrate the low energy NRs which would be expected in LZ, with the source located above the outer cryostat [73]. For NR calibrations, a mono-energetic Deuterium-Deuterium (D-D) source is used, with the generator setup being placed outside the water tank. The neutrons travel through the conduits in the water tank and outer detector to reach the LXe in the TPC. This was also used in LUX for NR calibrations, and allowed for in-situ calibrations at low energies [97].

2.5 Backgrounds in LZ

The ability to understand and characterize the background activity present at LZ is vital to achieve a world beating WIMP-nucleon sensitivity. This is as not being able to fully understand the background components which the detector will see can deter the ability to attribute statistical significance on any potential excess of a detected signal. This section will highlight the main sources of backgrounds that LZ faces.

2.5.1 Cosmogenic Backgrounds

These are backgrounds such as the muon flux, which at the surface level was measured to be $(1.149 \pm 0.017) \times 10^{-2}$ s^{-1}cm^{-2}sr^{-1} [98]. This is mitigated by placing the LZ detector at 4850 feet underground (4300 water equivalent), and surrounding it by the water tank reduces the muon flux by a factor of 3×10^6 relative to the surface. By using the Outer Detector, the remaining muons are vetoed due to them placing energy deposits in the water tank.

The muon flux can also generate neutrons by interacting with the cavern walls. These neutrons are problematic, as they would create more backgrounds in the detector. However, neutrons are expected to scatter multiple times within the TPC [99], as well as create an energy deposit in the outer detectors. Hence these can be vetoed thus reducing their contribution by a factor of 6 [100].

2.5.2 Surface Contaminants

Surface contaminants present a large amount of backgrounds to LZ. These can be introduced via dust accumulation on the inner side of the PTFE during the assembly process.

To try and mitigate this, LZ has set a target for plate-out of the radon daughters to be 0.5 mBq/m^2 for the walls around the TPC, with a maximum of 10 mBq/m^2 for all other areas. The total amount of dust accumulation is set to 500 ng/cm^2 to minimise the risk associated with this. More details about radon plate-out will be discussed in Chapter 4 of this thesis.

2.5.3 Trace radioactive contaminants

Due to nuclear fission weapons tests being carried out by countries throughout the 20th century, the prevalance of radioactive isotopes in the air has increased [101] [102]. ^{238}U, ^{235}U and ^{232}Th and their progeny are the most prevalent within LZ, with events within their daughter nuclei being of particular importance, since they can lead to ER and NR within WIMP energy regions of interest. Hence, a thorough screening of all detector components used in LZ has had to be undertaken [73].

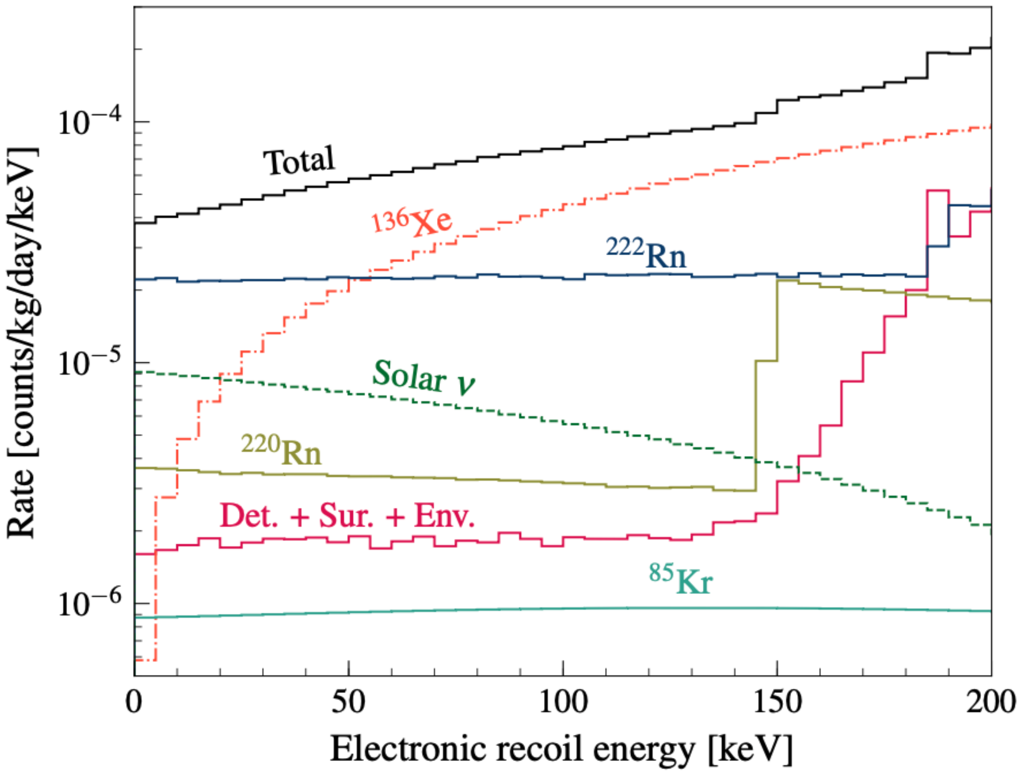
Natural xenon also contains a small amount of ^{85}Kr and ^{39}Ar. These mix uniformly throughout the LXe, and lead to ER events mainly due to β-decay. To remove these contaminants, LZ purifies the LXe using chromatography [73], with the reduction of Kr/Xe being 0.075 ppt g/g (parts per trillion grams per gram of Xe), and Ar/Xe less than 0.45 ppb g/g (parts per billion grams per gram of Xe).

Another source of backgrounds occur from cosmogenic activation of some of the isotopes of Xe, mainly ^{129m}Xe, ^{131m}Xe, ^{133}Xe, each with a half-life of 8.9 days, 11.9 days and 5.3 days respectively. However, due to their short lifetimes, their contribution to the total background rate at LZ can be minimised by implementing a cooling period of $\sim$ 8 months of the LXe underground before taking any data. This would reduce the contribution caused by these isotopes to negligible levels.

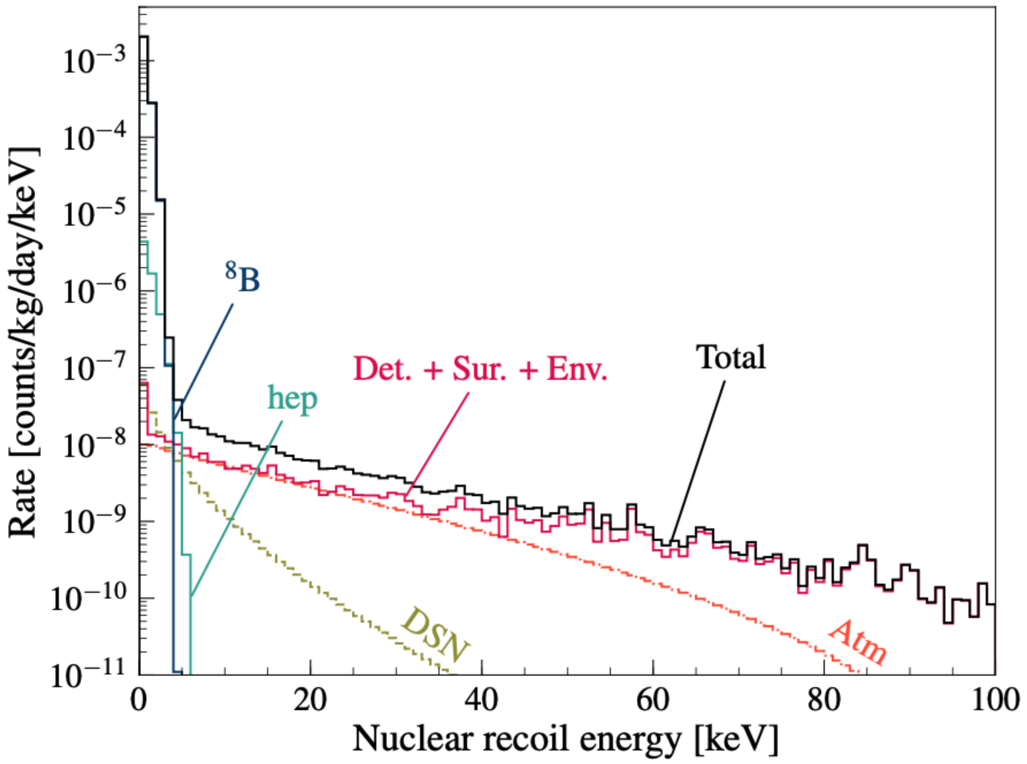
2.5.4 Irreducible backgrounds

LZ is sufficiently sensitive to measure neutrinos from many sources. Backgrounds caused by neutrinos are also expected to contribute to the total backgrounds seen by LZ, with the PP-solar neutrino flux expected to induce ER events. Additionally, atmospheric neutrinos can also induce NR events, with ^{8}B solar neutrinos causing low energy NR events [91].

The total contribution of all the predicted LZ backgrounds for a 1000 day run is



(a)



(b)

Figure 2.7: Projected backgrounds for single-scatter events only for a 1000-day run of LZ. These do not include vetos by detectors. Pink line in both plots is the summed total background due to Detector (Det), Surroundings (Sur) and Environment (Env). Brown line in 2.7b indicates backgrounds due to Diffuse Supernova Neutrinos (DSN). Orange line in 2.7b is backgrounds due to Atmospheric Neutrinos (Atm). Image taken from [91].

shown in figure 2.7 [91]

Chapter 3

Machine Learning and LZ

Read! Your Lord is the Most Gracious. Who taught by the pen. He taught man what he did not know.

Quran [96:3-5]

Algorithms that can 'learn' from a set of experiences has been the holy grail for computer scientists and engineers since the first concept of a computer was developed. Algorithms that can perform such tasks are called Machine Learning (ML) algorithms [103]. This ability to learn from experiences (data) can allow machine learning algorithms to solve problems that are too complex for classical algorithms - which rely on statically-defined steps, to compute. In this chapter, a survey of the previous attempts to use machine learning within the context of WIMP dark matter searches will be presented. An overview of machine learning concepts and ideas will also be highlighted, with an emphasis on deep learning in particular and its use within High Energy Physics. Since machine learning algorithms rely heavily on large datasets, the LZ data set will be discussed, and the data acquisition procedure of LZ.

3.1 Overview of Machine Learning

Machine learning is a field within computer science which describes the ability of an algorithm to learn features from a dataset [104]. The aim of the algorithm is to improve its ability to learn patterns within the data automatically through experience, and without the need of intervention from a user [105]. It is an active field of research within Computer Science, with research within the field increasing over the last two decades [106].

Currently, machine learning is applied to solve four broad categories of problems: clustering, classification, regression, and feature extraction and reduction [103]. In classification problems, the goal of a ML algorithm is to learn the underlying pattern which maps a set of features to a discrete label, with the aim to apply the learned pattern on new input data. An example of such a method is the Random Forest Classifier [107]. Regression problems are similar to classification tasks, with the labels instead being continuous variables. Linear Regression [108] is an example of such a regression task.

Clustering problems involves learning features within the dataset such that similar data points can be grouped together. The data provided to the ML algorithm is not labelled, hence the ML algorithm must find distinguishing features to separate the data points into clusters. An example of such a method is the use of neural networks [109] being applied to unlabelled handwritten digits.

Finally, in feature extraction and reduction problems, the goal of the ML algorithm is to reduce the number of dimensions of a dataset such that the data can be represented by its most important features [110]. This can either be done by removing a number of features that are correlated with each other or introducing a new feature which is combination of different features. An example of such a method is Principle Component Analysis (PCA) [111], which works by projecting the input data onto a lower dimensional space such that it maximises the variance within the features [112]. Although PCA can be used by itself, it is also used as a data preprocessing step to reduce the dimension of a dimensionally high dataset to allow other ML methods to be used.

The use of ML methods within different disciplines has increased over the past decade, with applications in biology [113], law [114], security [115] and healthcare [116] providing new insights into each respective field. The next section will highlight some of the applications of ML within Physics.

3.2 Machine Learning within Physics

The use of machine learning methods has been prevalent within physics experiments over the past two decades [117], and has helped experiments such as the LHC in the discovery of the Higgs boson [118]. Within Astrophysics, machine learning has been used extensively [119] since the 1980s, with techniques such as Principal Component Analysis (PCA) [111] used in the classification of spiral galaxies [120].

For cosmological simulations, machine learning has been used to help understand dark matter haloes [121] by teaching machine learning algorithms how to populate galaxies into dark matter haloes using data from the MUFASA simulations [122].

Neural networks have recently been used on data from the XENON1T experiment [123] to detect WIMP interactions [124]. There, a specific type of neural networks called convolutional neural networks - described in more detail in the next section, were applied to simulated events consisting of electron recoils (ER) as background, and 500 GeV/c^2 WIMP masses as signal, with an accuracy greater than 85% being achieved.

Within the LZ collaboration, machine learning has been used to classify the wide range of pulses recorded by the PMTs on simulated data [125]. To achieve this, a combination of Random Forests [126] and neural networks were implemented, and compared with classical methods. The results showed that the machine learning approach was able to achieve an accuracy greater than 99%, and was higher compared to the classical methods used.

3.3 Fundamentals

Machine learning algorithms are used for a wide variety of tasks, with some of the most common being classification and regression. In classification tasks, an algorithm is asked which category k a set of inputs x_i belongs to. In regression, the algorithm is asked to predict a numerical value given a set of inputs x_i.

However, the learning process by which an algorithm learns is dependent on the dataset it is given. The two main categories are then supervised learning; whereby the target values y and input features x_i are given to the learning algorithm, and unsupervised learning; where only the input features x_i is given.

This section will deal with the mathematical formulation for the learning process of machine learning algorithms, with a specific emphasis on supervised learning.

3.3.1 The Learning Process

The data that is used as input for machine learning algorithms can be represented in terms of a vector space $X = \mathbb{R}^n$. In this vector space, a *feature* is defined to be each dimension, with features being observable quantities. Hence, a dataset can be defined as a set of feature vectors x_i which are sampled from an (often unknown) underlying probability distribution P(x). Therefore, a machine learning algorithm can be defined as containing a model f and loss function L

$$f(x, \omega) \rightarrow Y \tag{3.1}$$

$$L(f, x) \rightarrow \mathbb{R} \tag{3.2}$$

where f is a function which maps from a feature vector x to an outcome Y given a vector of parameters ω; and the loss function L measures the performance of the model by taking into account the feature vectors and the outcome. An optimisation method is needed to tune the parameters of the model by taking steps within the learning process to minimise the loss.

Hence, 'learning' of a machine learning algorithm can be defined more generally as *"if its performance at tasks T, as measured by [some performance measure] P, improves with experience E"* [103].

3.3.2 Training an Algorithm

This learning process is often called 'training' and is similar to classical optimisation problems, but with an added aim of the model to find the useful properties that generalise over the training data in order to be applied to new datasets. This is different to classical optimisation problems, where the aim is to find the parameters that give the optimal loss [127].

Evaluating a machine learning model on an unseen test dataset can thus give a measure at how good the model is at generalising the patterns within the data. This test dataset should be chosen such that it is representative of the whole distribution of the dataset, and is not involved in the training process.

Mathematically, the training of machine learning algorithms takes the form of a gradient-based optimisation, whereby

$$\nabla_{\omega} L = 0 \tag{3.3}$$

Hence, steps are taken to descend the gradient of L with respect to ω to minimise L.

Although this can be done over the entire dataset, it is often impractical due to computational limitations caused by large datasets or high dimensionality. Hence, smaller batches of the training dataset are often taken, with the gradient of these smaller batches taken instead. This stochastic gradient descent (SGD) [103] method thus updates the parameters of the model as

$$\omega \rightarrow \omega - \eta \nabla_{\omega} L \tag{3.4}$$

where η is defined as the learning rate. This is a non-learned parameter that controls the step-size change of each parameter. By making changes in the parameters ω iteratively, the aim of an optimiser is to converge to a global minimum.

There are many more optimisation algorithms which work better compared to SGD, with the Adam optimiser [128] often being used widely due to it incorporating a 'momentum-like' parameter during the update process. A review of different optimisers being found in [103].

3.3.3 Generalisation and Model Capacity

For every machine learning algorithm, the space of functions available to each model to describe the observed (training) data is called the hypothesis space for the model. An example of this can be the hypothesis space fro linear regression, whereby the hypothesis space can only be for functions of the form

$$y = \sum_{i=0}^{N} \omega_i x^i \tag{3.5}$$

Hence, by having a larger hypothesis space available to a model, the capacity of the model to generalise the data is increased, thus increasing its descriptive power [103]. However, restrictions must be made to the size of the hypothesis space available to the model. A very large capacity can lead to over-fitting, whereas having a very small capacity leads to under-fitting. These errors in generalisation can be seen in figure 3.1, where various fits are done to a training dataset that is quadratic in nature.

An under-fit model will not be able to generalise the data well, hence an estimate will not fit the data. Hence, if the hypothesis space available to the model is too small, then it would not be able to capture the generalisations of the model [103]. This can be seen by fitting a linear function to the data. In an over-fit model, the

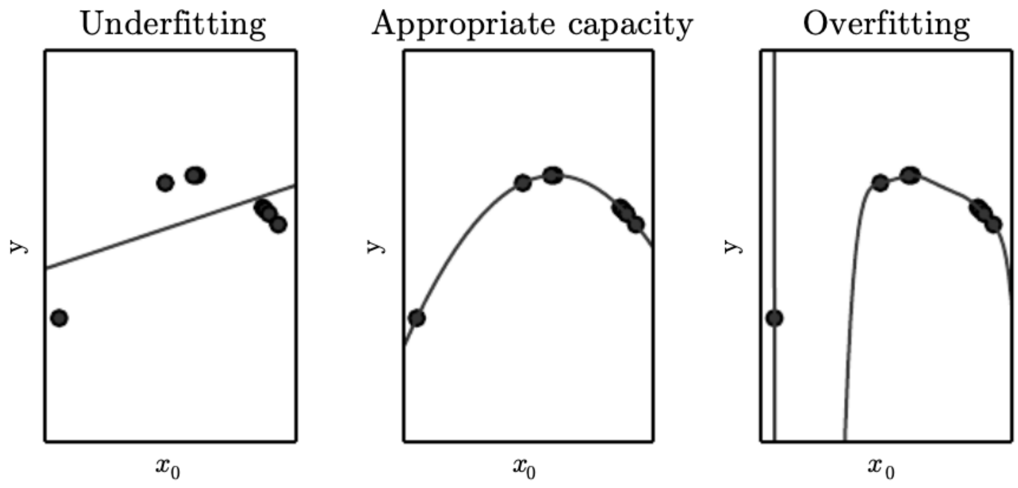


Figure 3.1: An example of Underfitting, Appropriate Capacity, and Overfitting. x was randomly sampled, with $y = x^2$ calculated. Underfitting shows a linear fit to the data, with the fit function unable to capture the curvature of the function. Appropriate Capacity shows a quadratic function fit to the data; indicating a good level of fit hence generalising well to unseen data. The overfitting plot shows a polynomial of degree 9 being used to fit the data, with the fit going through all the points exactly. More information regarding the use of Moore-Penrose pseudoinverse to solve for underdetermined equations found in [103].

model will give results that pass through all the points exactly, hence not being able to generalise the data. This can be seen by fitting a 9-th degree polynomial to the data.

To overcome this issue of model capacity, regularisation can be done to the model parameters ω, and is beyond the scope of this thesis. Further information found in [103].

3.3.4 The Random Forest

Instead of training only one machine learning algorithm on a training dataset, it can be useful to train multiple models and combine their outputs e.g by a weighted sum. An example of this is the Random Forest classifier, which is an ensemble of Decision Trees trained together [107].

Although first developed in the early 2000s [107], random forests are some of the most successful methods when used with large volumes of information. They are simple to use, with only a small number of parameters to tune for optimal results [107].

Random Forest classifiers work by finding the most important features in a random subset of features. By training multiple decision trees on these subsets of data, the random forest classifier ensembles multiple decision trees. Therefore, when unseen (test) data is analysed using the trained random forest classifier, the decision tree ensemble all make a prediction. The random forest algorithm then picks the classification with the most votes.

3.4 Deep Learning

Deep Learning is a subset of Machine Learning and is based on Artificial Neural Networks (ANNs) in a way to mimic the neurons in the brain [129] [103]. They are especially adept at working with datasets with a large number of features, with the name referring to the number of layers of artificial neurons within a model. With the advancements in hardware, namely GPUs, as well as having larger and larger datasets available, deep learning has seen a large increase in development over the past decade [106]. This section will highlight the main features of deep learning and artificial neural networks. Convolutional neural networks, a type of neural network used in computer vision, will also be discussed.

3.4.1 Artificial Neural Networks

Artificial Neural Networks mimic the neurons found in brain cells. A single neuron is called a node, which receives an input signal consisting of a feature vector x_i, a collection of weights in a weight vector ω_i, and bias b; and releases an output signal y, which is calculated by applying a non-linear activation function f on the sum of the inputs z. This can be expressed mathematically as

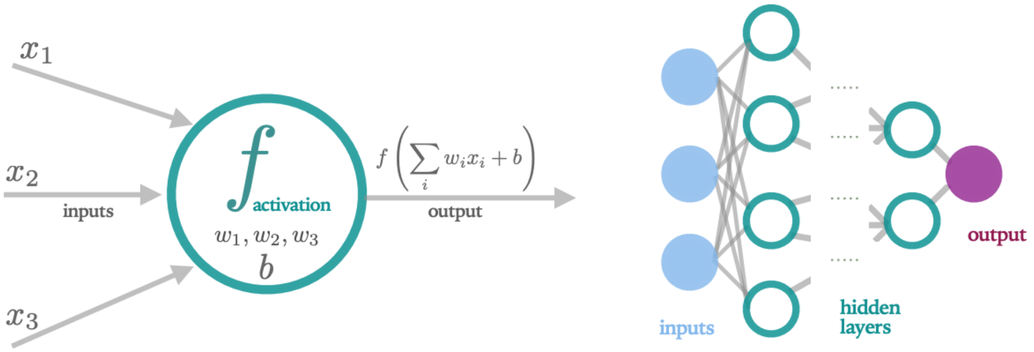


Figure 3.2: Left: Image showing the inputs to a single neuron, which takes inputs from the previous layer. These values are then combined using equation 3.6 to produce an output. This output then feeds into the next layer. **Right:** Image showing a complete neural network. Each neuron within the network is connected to every other neuron in the previous layer, and every neuron in the next layer. The process described on the *Left* takes place within every hidden neuron (shown in green) within the network. Blue-filled circles indicate input data, with purple-filled circles indicating the output of the network. Image taken from [130].

$$y = f(z) \tag{3.6}$$

$$z = \sum_{i=1}^{N} \omega_i x_i + b \tag{3.7}$$

where N is the total number of neurons within the layer. An image of a single artificial neuron can be seen in figure 3.2. The most common activation functions generally used in neural networks are the tanh, sigmoid and rectified linear unit (ReLU) functions shown below respectively.

$$f(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}} \tag{3.8}$$

$$f(z) = \frac{1}{1 + e^{-z}} \tag{3.9}$$

$$f(z) = \begin{cases} 0 & z \leq 0 \\ z & z > 0 \end{cases} \tag{3.10}$$

The need for non-linear functions used as activation functions arises due to the fact that by using a linear function, the neural network would just output a linear combination of the inputs to the neural network. Hence even by adding many layers of neurons, a linear activation function would only output a linear combination of the input features.

3.4.2 Assembling Neural Networks

Neural Networks come in very different architectures, each designed to solve a specific set of problems distinct to the dataset, and each utilising different types of neurons. In this thesis, feedforward neural networks will be discussed. These are networks where the layers of neurons are assembled linearly, with each layer feeding into the next layer until the output layer. The fully-connectedness of a layer is when each neuron in layer l is connected to every other neuron in the adjacent layers $l+1$ and $l-1$.

The number of hidden layers - layers between the input and output layers, can vary; with an increase in hidden layers leading to higher model complexity [131]. Hence, by having more hidden layers, the network can be modelled as a combination of equation 3.6, and can be expressed [132] to give the value of any neuron as

$$z_j^{n+1} = \sum_k \omega_{jk}^{n+1} y_k^n + b_l^{n+1} \tag{3.11}$$

$$y_j^{n+1} = f(z_j^{n+1}) \tag{3.12}$$

where y_k^n is the value of neuron k in hidden layer n - and is similar to equation 3.6, and the weight vector ω_{jk}^{n+1} represents how neuron k in layer n affects neuron j in layer $n+1$. Thus the output of the network is obtained by going through equations 3.11 and 3.12 layer by layer, starting at the input layer $n=0$. Hence by increasing the number of hidden layers, deep learning can allow for more complex datasets to be generalised.

3.4.3 Training Neural Networks

Training neural networks is similar to training any other machine learning algorithm whereby the set of parameters ω is optimised in order to minimise the loss function L. Hence, by using some form of gradient descent i.e. SGD, a single iteration of training a neural network involves a forward pass, a backward pass, and updating the weights.

In a forward pass, the input of feature vectors is run through the network, with the application of the activation functions, weights and biases applied and stored. This will continue until the final output y is outputted. The backward pass is similar to the forward pass, but with the output layer vectors going back to the input. At each neuron, the gradient $\partial L/\partial \omega_{ij}$ is calculated. Finally, the weights are updated following gradient descent.

3.4.4 Network Depth

Although increasing the number of hidden layers in a neural network increases the complexity, and thus allowing neural networks to solve more complex problems; an increase in depth produces issues such as the vanishing gradient problem [133].

In the training process, the gradients calculated at each neuron can have gradients within $[0,1]$. These would be multiplied by the number of layers in the network, thus causing the gradients to become very small. This would result in very slow training, with each weight update getting smaller. However, this generally occurs when using the sigmoid activation function. If gradients become too large, then the opposite issue can occur, whereby the gradients increase exponentially (thus called the exploding gradient problem), when larger than 1.

3.4.5 Convolutional Neural Networks

Convolutional neural networks (CNNs) are designed to specifically work with image data, and are used in computer vision tasks where object detection is necessary

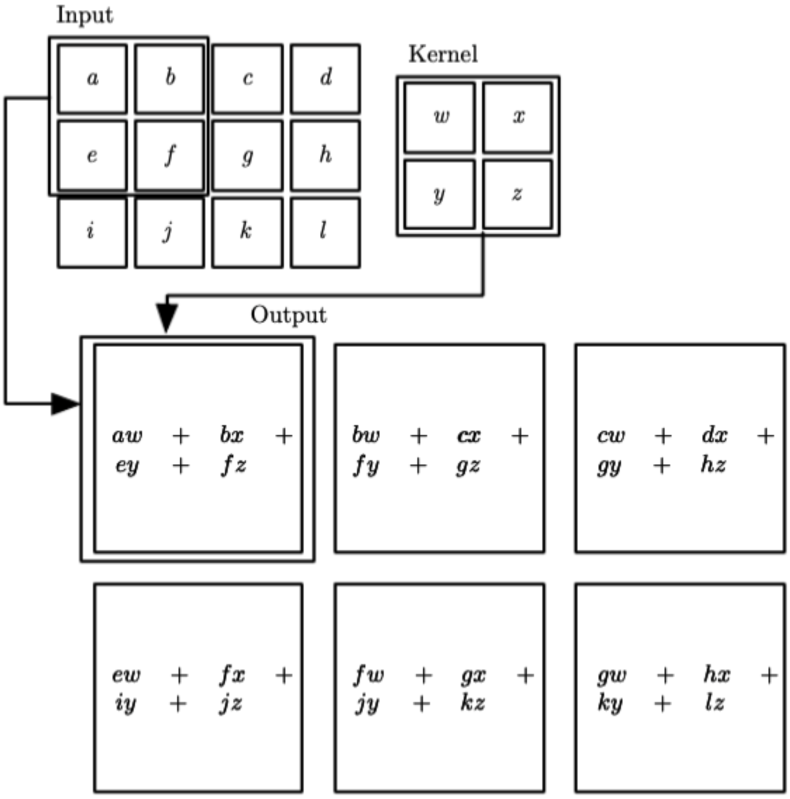


Figure 3.3: Image shows an example of 2D convolution when applied on an input data of size 4*x*3. The kernel is applied to each section where the size of the kernel can be overlaid. Different types of methods are used for when the kernel reaches the edge of the 2D input, with the image showing a "valid" convolution - where the output is only for positions where the entire kernel lies within the data. Kernel values can be taken prior, or learnt during training. Image taken from [103].

[134]. This section will outline how convolutional neural networks work.

Images can be represented in data by pixel sizes for height and width, with the depth usually referring to the red-blue-green (RGB) channels in an image. If a fully connected multi-layer perceptron (MLP) [135] was used on image data of size $(100 \times 100 \times 3)$, there would be 30000 inputs for each neuron in the next fully connected layer. This would lead to a large number of parameters to be calculated, even for smaller networks. Having a large number of parameters would then lead to a network that is difficult to train.

CNNs introduce two new types of neural network layers; convolutional and pooling layers [136]. A convolutional layer applies a convolutional operation of the

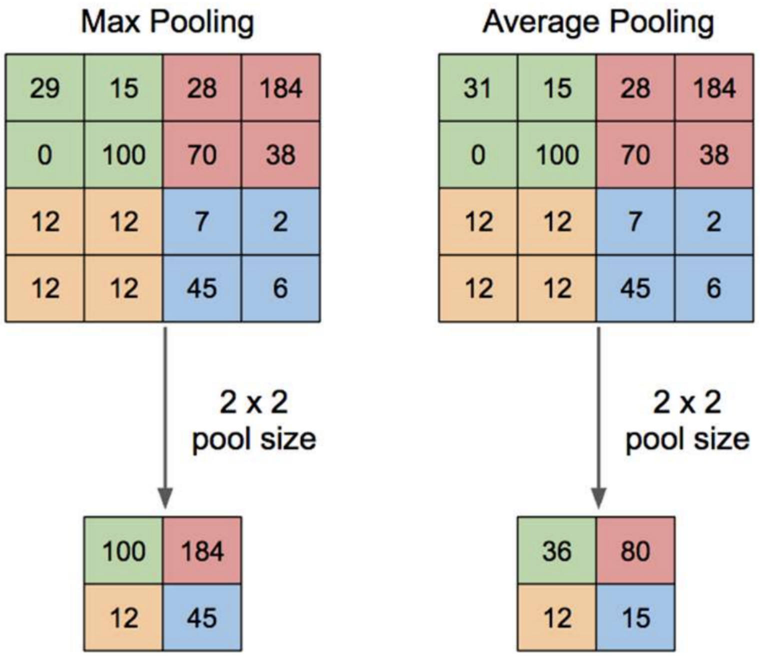


Figure 3.4: Image showing the effects of pooling on a dataset, with the two most common pooling types mentioned - Max and Average Pooling. Max pooling can lead to faster training, whereas average pooling retains more information thus keeping translational invariance, but suffers from longer training time. Image taken from [137].

form

$$s(t) = \int x(a)w(t-a)da \tag{3.13}$$

where $x(a)$ is the input of features, and $w(t-a)$ the kernel function applied to the input data. The output $s(t)$ is generally referred to as the feature map [103].

Hence these convolutional operations apply local transformations to a local section of the input. These are then applied to each local section on the input data, with the total section then forming the feature map. These can be seen figure 3.3, where a (2×2) kernel is applied to an input of size (4×3), with the local section of size (2×2) applied on the input data.

Changing the sizes of the kernel function leads to different effects, which are necessary to be taken into account for different datasets. An example of this can be that

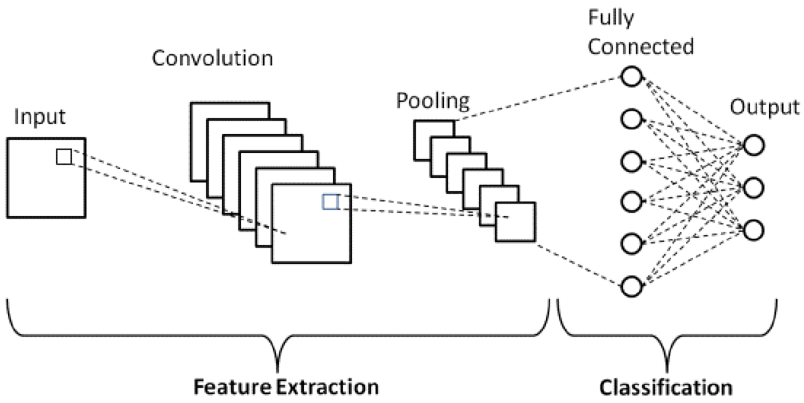


Figure 3.5: Image showing an example of a convolutional neural network architecture, with a single convolutional and pooling layer, followed by a fully connected layer leading to an output. Image taken from [138].

changing the kernel size can lead to feature reduction of the data.

Pooling layers are used in convolutional neural networks to reduce the size of the previous layer to a single value by mapping sections of the input space. An example can be seen in figure 3.4, which shows the most common types of pooling; Max and Average Pooling. By using a pooling layer, the model complexity can be reduced, with max-pooling often allowing for faster training of the network, as well as allowing for translational invariance. However, by using average pooling, the region of interest within the section is taken into account, thus not losing much information. However, this does lead to networks with average pooling layers taking longer to train.

Summary

Machine learning algorithms try to learn generalisations and patterns found within an input dataset. They do this by optimising a set of weights by reducing a loss function. A brief overview of machine learning algorithms was given in this chapter to set the foundations for the methods and architectures used for the LZ direct detection dark matter experiment. This ability to learn generalisations allows machine learning to be tested across a wide variety of diagnostic and analysis methods in LZ. The next section will describe the data processing techniques deployed in the

LZ experiment.

3.5 LZ Data Generation

In order to be able to apply any machine learning algorithms to data, an understanding of the processes implemented to generate that data is required. LZ uses its own simulations software package, which was designed and built simultaneously with the physical design and construction of the detector. This section will briefly describe the LZ simulations package stack, which is built utilising closely the GEANT4 package used extensively in High Energy Physics Experiments.

3.5.1 Simulations

Before any data acquisition can take place, an extensive simulations campaign must be run in order to understand the response of detector apparatuses to background and signal events. Monte Carlo simulations can be used to test theories against data thus giving a deeper meaning of the physics involved.

By using simulations, the background rates involved in the LZ experiment can be fully understood, thus giving the ability to set sensitivity limits and WIMP-discovery potentials. Simulations also allow for the tuning and optimisation of detector components and operations. Taking into account background rates specific to the materials used in construction, a detector optimised in detecting low background dark matter signals can be built.

The LZ simulation stack works by simulating an initial energy deposit in LXe and simulating its trajectory and energy depositions until an output that is designed to be identical in data type to what real data would look like. The initial energy deposit is run using an in-house code base developed on top of GEANT4 [139] called BACCARAT (Basically A Component Centric Analog Response To Anything) [140]. This package is primarily used to generate and track particles within the detector and to identify and record the points of interaction. It is also designed to be more user friendly for easier operations.

By using a component centric approach, BACCARAT allows for more accurate modelling of materials that are used in construction. This is vital in low background

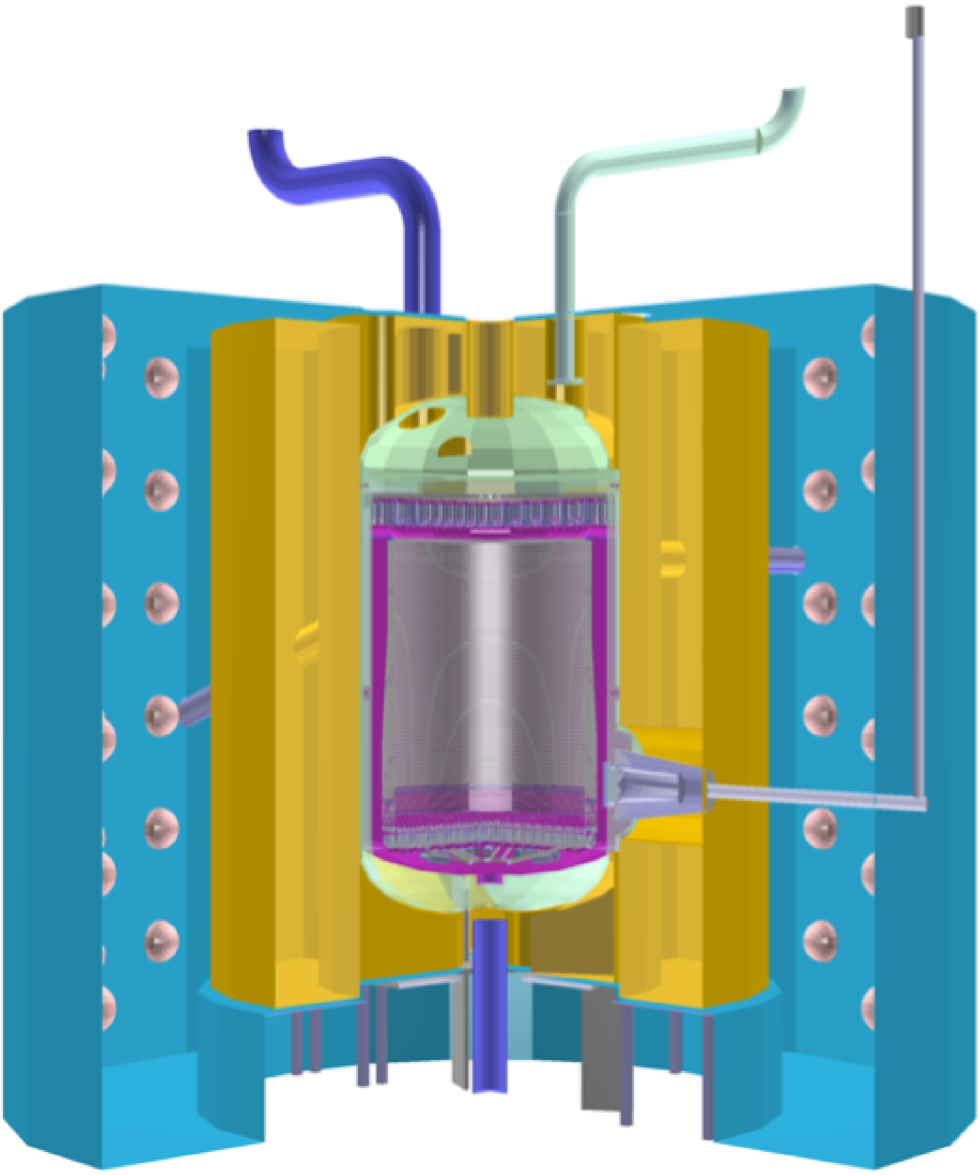


Figure 3.6: The LZ detector geometry constructed by BACCARAT. Image shows a blue water tank with PMTs, with the Outer Detector shown in yellow. The ICV and the OCV are shown in green, with the TPC in purple.

experiments, where accurate detector component modelling is needed to mimic impurities found within the materials. It uses C++ for fast computations, and can allow for component-specific background values and recording levels to optimise data storage.

Implementing, and being built on top of GEANT4, BACCARAT also includes physics modules within the GEANT4 toolkit such as *G4EMLivermorePhysics*. This allows for effects such as those induced by low energy electrons $\sim$ 10 eV to be included.

After running BACCARAT, the output can be fed into two different simulation chains (see figure 3.7), depending on the intended use. A fast chain exists which allows for quick simulations generally used on a smaller scale. Hence the output of BACCARAT is given to NEST (Noble Element Simulation Technique) [141]. This software package uses light and charge yield to convert the energy depositions to S1 and S2 signals. However, the fast chain is unable to give PMT information such as photon hits, or the times of interactions.

In order to simulate what a real-world event would look like, the full chain in the LZ simulation stack is utilised. This is where the output of BACCARAT is fed into the Detector Electronics Response (DER) package. This package models the PMT and electronics response to photons reaching the photocathode. It can model PMT specific features such as the quantum efficiency, afterpulsing and noise, and the electronics used in the front and back-end of the PMTs, thus enabling the generation of realistic waveforms.

This output is then saved in a data format identical to the format which will be used when recording real data, mainly opened using the ROOT analysis package. This is a package which utilises C++ in order to analyse large amounts of data [142].

However, saving the waveforms data requires extra storage space, hence the waveforms are then analysed using the LZ Analysis Package (LZap). This deconstructs the waveforms into pulse and event-level information (instead of the time-series data available in the waveforms) which are then used in analyses. This reduction creates

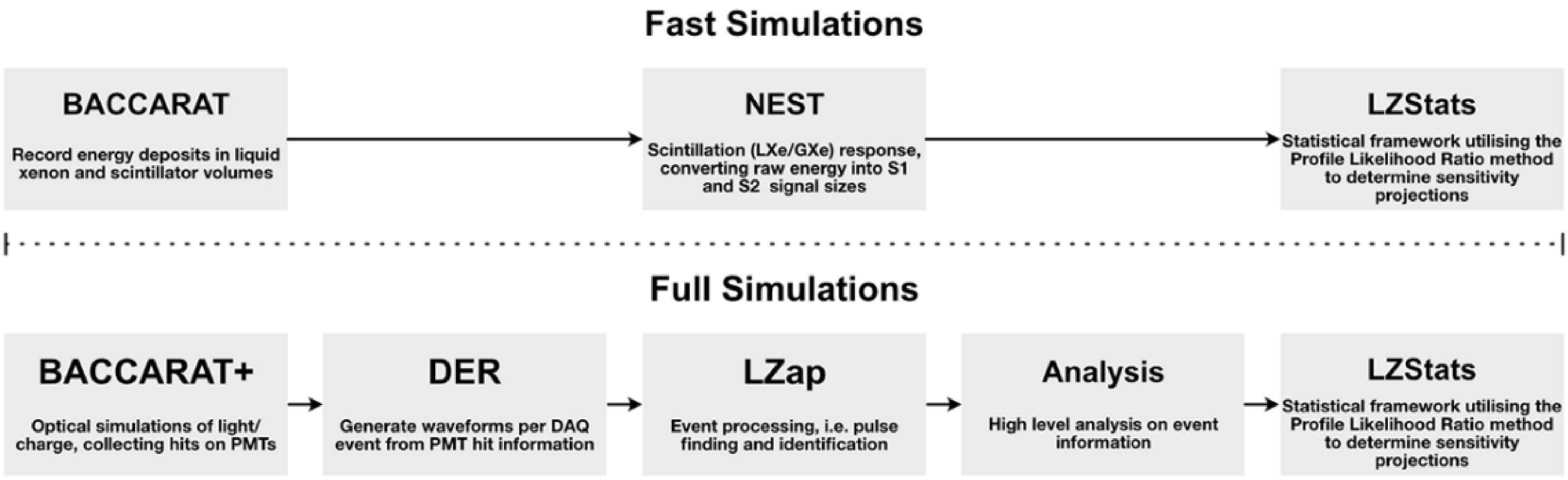

Fast Simulations
BACCARAT
Record energy deposits in liquid xenon and scintillator volumes
NEST
Scintillation (LXe/GXe) response, converting raw energy into S1 and S2 signal sizes
LZStats
Statistical framework utilising the Profile Likelihood Ratio method to determine sensitivity projections
Full Simulations
BACCARAT+
Optical simulations of light/ charge, collecting hits on PMTs
DER
Generate waveforms per DAQ event from PMT hit information
LZap
Event processing, i.e. pulse finding and identification
Analysis
High level analysis on event information
LZStats
Statistical framework utilising the Profile Likelihood Ratio method to determine sensitivity projections

Figure 3.7: Image showing the LZ simulation package stack, with the corresponding packages used for the fast and full simulation separated.

Table 3.1: Parameters of the LZ detector used in simulations [91]. *phd* refers to photons detected, *ph* referes to photon, SE to single electron, and PDE refers to photon detection efficiency.

Parameter	Value
Electric Field	310 V/cm
Electron Lifetime	850 μs
Electron Extraction Efficiency	95%
Average PMT Efficiency	27%
Average PDE in liquid (g_1)	0.119 phd/ph
Average PDE in Gas ($g_{1,gas}$)	0.102 phd/ph
Single Electron Size	83 phd
Effective charge gain (g_2)	79.2 phd/e
S1 coincidence level	3 ph
Singe phe trigger efficiency	95 %
PTFE reflectivity in LXe	97.7 %
PTFE reflectivity in GVe	85 %

Reduced Quantities (RQs) files, which are saved in `.root` file format, utilising the tree and branch structure available.

3.5.2 Detector Parameters

For BACCARAT to simulate accurately, the detector parameters seen in table 3.1 were used. After initialising an event to be simulated, the photons would travel within the LXe, reflecting off the PTFE walls before being detected by a PMT. A g_1 value was set to represent the average successful recording of a single photon by taking into account the PMT quantum efficiencies and reflectivity. g_2 represents the photons that are detected from a raw S2 signal. It takes into account the liquid-gas boundary and the extraction efficiency, as well as the lost electron correction this produces. This is different to the single electron size, which is the value of the number of photons extracted for a single electron.

3.5.3 Sensitivity Calculations

A Profile Likelihood Ratio (PLR) method is used to calculate the LZ sensitivity [143] [144]. The median 90% confidence level upper limit on a specific WIMP mass is found for a background-only hypothesis. NEST is used to create probability

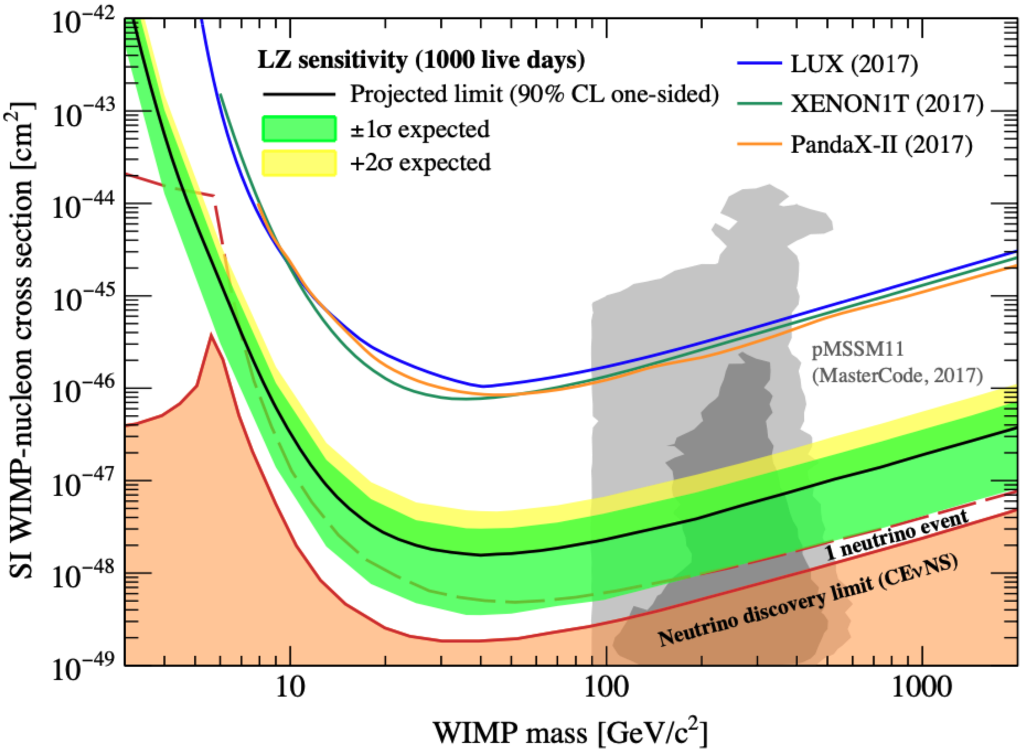


Figure 3.8: The projected sensitivity of LZ to WIMP-nucleon elastic scattering for a 1000 day run, and with a 5.6 tonne fiducial mass [91]. The sensitivities of LUX (blue line), XENON1T (green line) and PandaX-II (orange line) are also displayed - with the latest PandaX-II result not displayed due to still undergoing peer-review.

density functions (PDFs) parameterised in S1_c and S2_c (which are corrected values for the S1 and S2 signal sizes), which are then fed into the PLR. BACCARAT is used to simulate the expected background rates caused by material and surface radioactivity. Hence, for a 40 GeV/c^2 WIMP mass, the LZ sensitivity is calculated to be 1.7×10^{-48} cm^2 [91] for a 1000-day run with a 5.6-tonne fiducial volume mass, and can be seen in figure 3.8.

3.6 Data Flow

LZ aims to detect low energy events, with the focus on looking at events below 40 keV because of exponentially falling WIMP signal models. To compare the amount of data storage that each event takes, the predecessor to LZ - the Large Underground Xenon experiment (LUX), can be examined, where each event took

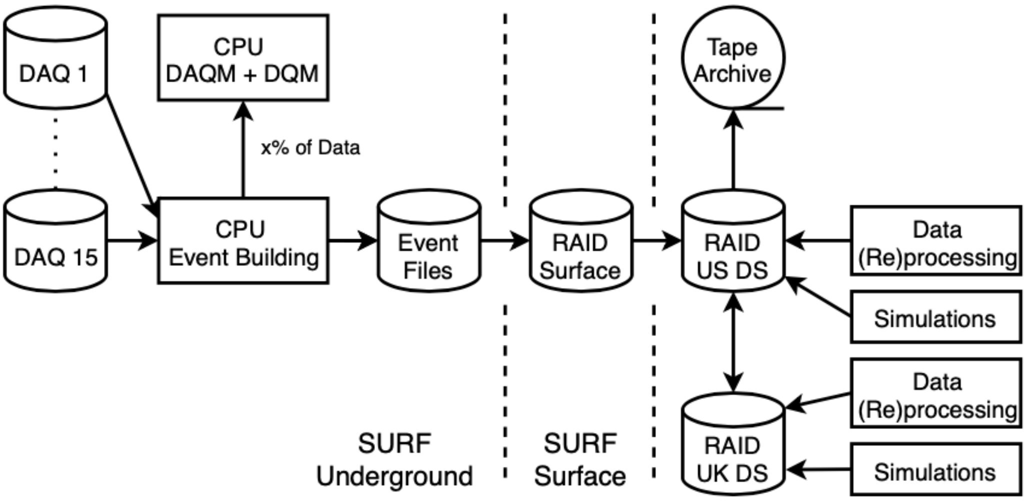


Figure 3.9: Image showing the LZ data flow schematic [145].

203 kB of data storage (1.7 kB/channel). Since LZ has ~ 4 times the number of PMTs, each event in LZ will take ~ 1.6 mB of data from the TPC only [73]. Hence for a 1000 day science run, LZ expects to collect a total of 2 PB of data (equivalent to 10^{15} bytes), with an average of 2.8 TB of data per day.

This section will highlight the main data engineering principles undertaken by LZ to ensure data taking happens efficiently and continuously.

Figure 3.9 shows a schematic of the data flow implemented by LZ. Five Event Builders take input from the 15 Data Collector disks (DAQ1-15) to assemble events by extracting relevant information. These are stored on local disks underground before being transferred to the surface.

The data is then transferred to two Data Centres for storage and analysis. These are the US Data Centre (USDC) located at UC Berkeley and NERSC; and the UKDC, located at Imperial College London.

A portion of the data is analysed via the Data Quality Monitor (DQM). These are separate servers located at the Davis Laboratory which monitor the performance of the detector. The DQM will monitor detector parameters such as trigger rates and the hit distributions of the PMTs. This allows for instant detector monitoring if significant deviations were to occur. All the data is also stored on physical tapes for archiving as a final redundancy.

3.7 Summary

An overview of the theory of Machine Learning techniques used was presented, and these will underpin the remainder of the thesis. The data upon which these machine learning techniques will be applied to was also discussed, with an overview of the simulation program BACCARAT given, as well as the LZ computing stack. Finally, an overview of the data flow that will be used by LZ was given, highlighting the final data storage capacity expected after Science Run 1 - the first science dataset that will release results in late 2021/early 2022.

Chapter 4

Radon backgrounds in LZ

He is the One who created seven heavens, one above the other. You will never see any imperfection in the creation of the Most Compassionate. So look again - do you see any flaws? Then look again and again — your sight will return frustrated and weary.

Quran [67:3-4]

Radon was discovered in 1900 by Friedrich Ernst Dorn, who showed a radioactive gas emanating from radium while studying its decay chain [146]. At normal room temperature, radon is a colourless, tasteless and odourless radioactive gas, with all of its isotopes being radioactive. Due to advancements in being able to reduce radiogenic backgrounds in LXe within direct detection dark matter experiments thanks to high precision radio-assays of components [147] [148] and 3D vertex reconstruction, backgrounds attributed to radon are one of the biggest contributors to the WIMP search Regions of Interest (ROI). This chapter will focus on measuring radon within the LZ experiment; highlighting methods to measure its activity using classical methods, as well as using Machine Learning.

4.1 Overview

Radon is a naturally occurring noble element present in the air. It has an atomic number of 86 and is the 6th noble gas in the periodic table. It is a monoatomic and inert gas that is naturally found in nature. Due to being a noble element, radon has long diffusion lengths within solids, thus making it difficult to remove chemically, hence mitigation of radon is done by employing physical means.

Radon has relatively large melting and boiling points at 202K and 211.5K, and a density of 9.7 $\mathrm{kg/m^3}$. Radon is produced via the decay of its parent radium, and its activity can vary significantly due to differences in the environment. Atmospheric activities of radon are approximately 10 $\mathrm{Bq/m^3}$, with indoor activities reaching up to $\sim$ 50 $\mathrm{Bq/m^3}$ [149] [150].

Radon related backgrounds were analysed using the 3rd Mock Data Challenge (MDC3) dataset produced by the LZ collaboration, with the challenge running between July - October 2019. The dataset was the final Monte Carlo simulation based challenge produced by the LZ collaboration and is the closest resemblance to real data being produced during the first Science Run, due to start in late 2021.

Background activities for radon and its daughters were not made public to the collaboration, hence allowing for a blind analysis of the data and testing of algorithms in preparation for data collection and rapid turnaround of results. This blinding was to ensure no bias was present within the analysis of the dataset when quantifying backgrounds.

This chapter begins with an overview of the radon decay chain, the backgrounds produced from its daughters and their effects on the WIMP search ROI. A tagging algorithm to identify radon rates using the Bismuth-Polonium decay chain will be discussed, as well as results for implementing it on the MDC3 WIMP search dataset. Analysis of the cuts used for the BiPo tagging will be shown, as well as results calculating the radon activity. The use of Machine Learning as another possible way to find radon related backgrounds will also be introduced and discussed.

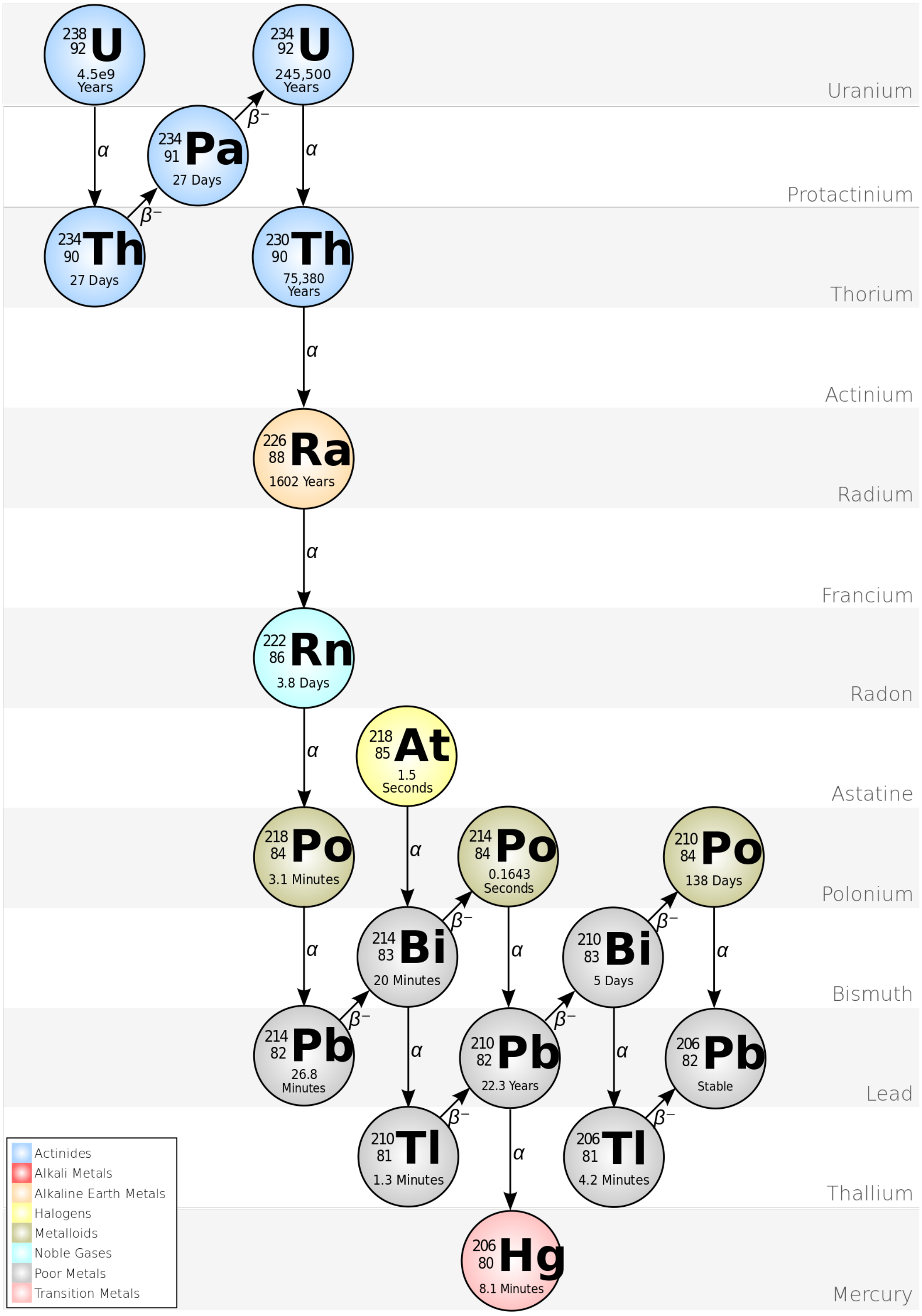


Figure 4.1: Image showing the ^{238}U decay chain. Image adapted from [151].

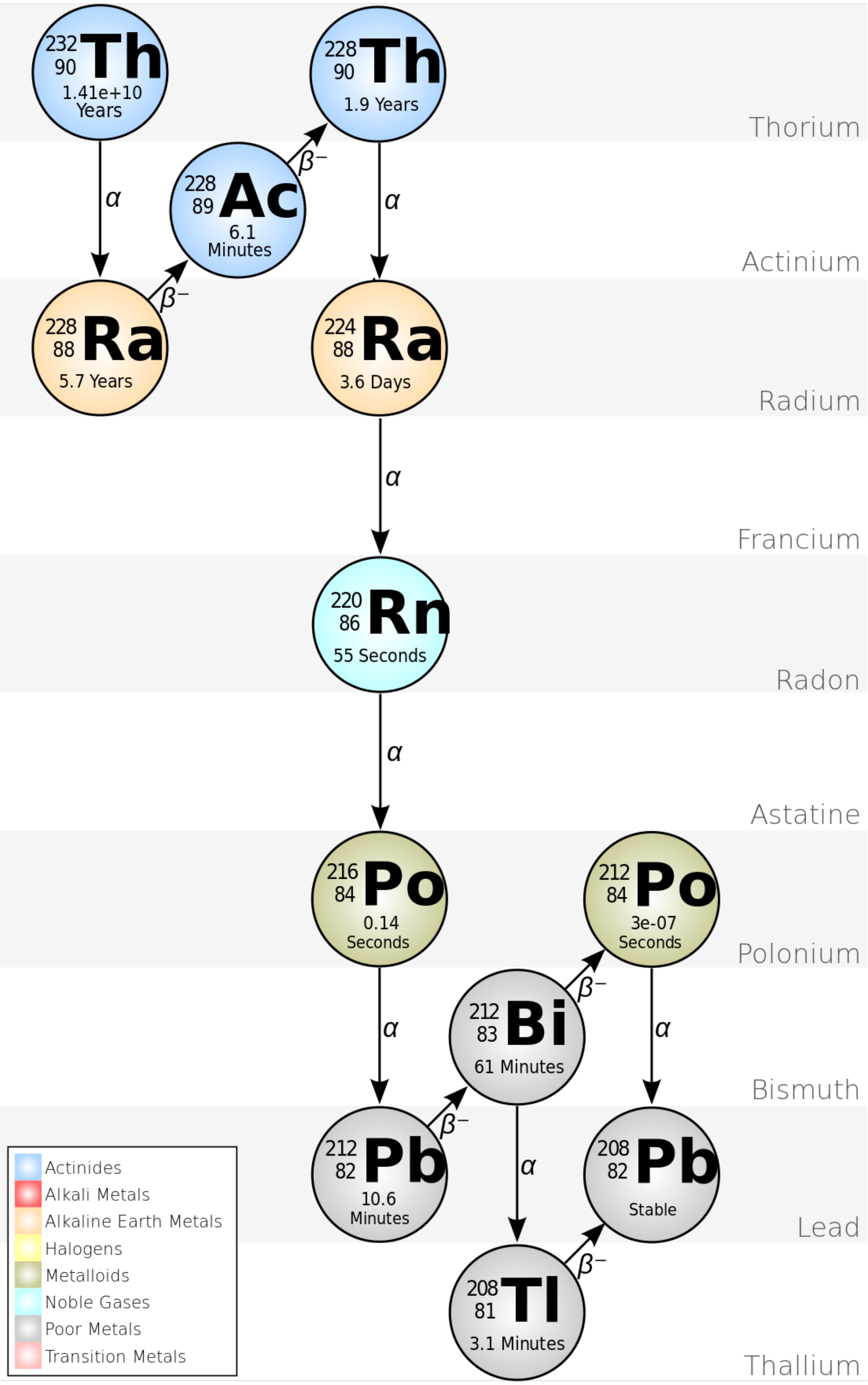


Figure 4.2: Image showing the ^{232}Th decay chain. Image adapted from [151].

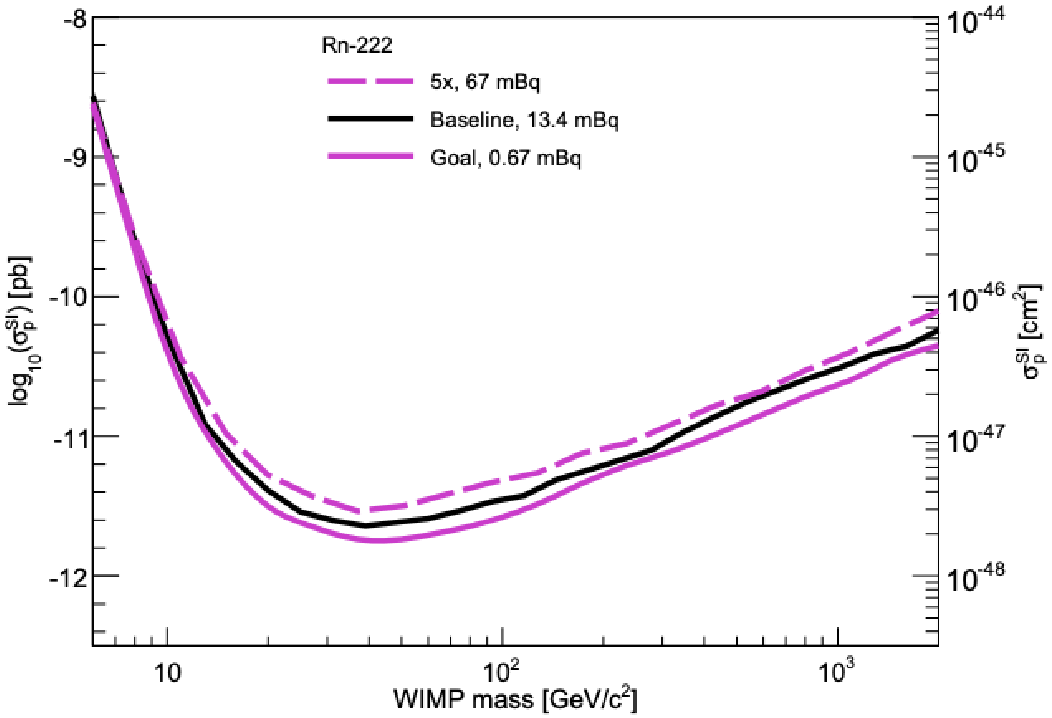


Figure 4.3: LZ sensitivity projection to spin-independent WIMPs for different concentrations of radon. A goal concentration of 0.67 mBq is shown along with a purple line, whereas the reduced (dashed purple line) gives a “worst-case” scenario [73].

4.2 Radon Physics

^{222}Rn is produced in the tail of the ^{238}U decay chain, as shown in figure 4.1, and decays via alpha and beta decays until reaching the stable Lead isotope. Due to the half-lives of all the daughters ranging vastly from milliseconds to thousands of years, secular equilibrium is achieved where the quantity of ^{222}Rn remains constant because its production rate is equal to its decay rate. This occurs when there is a difference in the relative half-life of parent and daughter nuclei. Therefore, by measuring the activity of one part of the chain, the activity of the rest of the chain can be inferred. This method is used when determining the activity of ^{222}Rn, as it allows for the measurements of ^{222}Rn daughters.

The main method of radon entering the TPC and causing a background is due to the presence of primordial ^{238}U. Thus, materials exposed to air which have ^{222}Rn present can cause isotopic plate-out on the surface [152]. ^{210}Pb, which has a half-

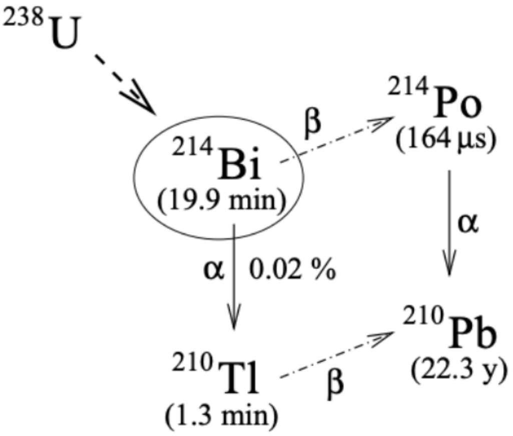


Figure 4.4: Image showing the branching ratios of the BiPo-214 decay.

life $\tau_{1/2} = 22.3y$, is the main isotope within the ^{238}U chain which causes a breaking in this secular equilibrium. Dust molecules also present a large source of radon background, with a cumulative dust count of 1g [91] being set. The effect of radon backgrounds on the WIMP search sensitivity can be seen in figure 4.3.

Materials that are in direct contact with gaseous or LXe are surfaces that contribute the largest to the backgrounds caused by radon. With a half-life of $\tau = 3.8$ days, ^{222}Rn is expected to mix homogeneously with the active volume.

The alpha decays present within the radon decay chain are of high energy, with Q-values exceeding 5.5 MeV. This energy is above the WIMP search ROI, however, these can often lead to other backgrounds from the recoiling nucleus. These alphas are easily detected within the data due to their large energy deposits corresponding to large S1 signals. The beta decays present affect the WIMP Search ROI more, with backgrounds contributing to the ER background - these will be further discussed in the following sections.

The main reason for such a high background is the naked-beta decay which occurs from the ground-state to ground-state of ^{214}Pb - which is found within the ^{222}Rn chain, to ^{214}Bi at a branching ratio of 9.2%, as seen in figure 4.4. A naked-beta decay is when the ^{214}Pb decays without emitting a gamma. When the emission of the gamma occurs, the Skin or Outer Detector (OD) is able to detect this gamma,

Table 4.1: Table showing the details of isotopes in the ^{222}Rn decay chain until ^{2214}Po.

Isotope	Decay	Q-Value [MeV]	$\tau_{1/2}$	Daughter
^{226}Ra	α	4.87	1602 yr	^{222}Rn
^{222}Rn	α	5.59	3.82 days	^{218}Po
^{218}Po	α	6.12	3.07 min	^{214}Pb
^{214}Pb	β	1.02	26.9 min	^{214}Bi
^{214}Bi	β	3.27	19.8 min	^{214}Po
^{214}Po	α	7.83	162.3 μs	^{210}Pb

hence veto the event. However, during naked beta decay, the decay would produce a uniform ER background, with a β-spectrum reaching 1019 keV. A summary of the decay of each isotope in the ^{222}Rn sub-chain is found in table 4.1.

4.2.1 Radon projection in LZ

LZ carried out large scale measurements of individual material components at various facilities to find the contribution that each component would give to the total radon background present in LZ. The ICV was also measured at various stages of the construction, with a final assay made after the completion of the ICV.

At cold temperatures where LZ is expected to operate (175.8 K), radon diffusion from materials in contact with the LXe is suppressed. This, coupled with how emanation rates are also affected by the type of material - with porous materials such as ceramics and plastics yielding larger emanation rates; result in uncertainties of the suppression rates at cold temperatures. This leads to a projected radon activity of 11.0 ± 1.0 mBq under less-conservative cold suppression rates [148], and with the deployment of the radon removal system to also lower radon levels.

4.2.2 Investigating the shape of the radon decay spectrum in LZ

The decay of ^{222}Rn observed in LZ looks like figure 4.5, which shows the electron recoil energy [keV] and rate. The plot shows all the beta decays within the ^{222}Rn chain which causes a background. The reason for the shape is due to electromagnetic interactions between an excited nucleus and the inner shell electrons after a

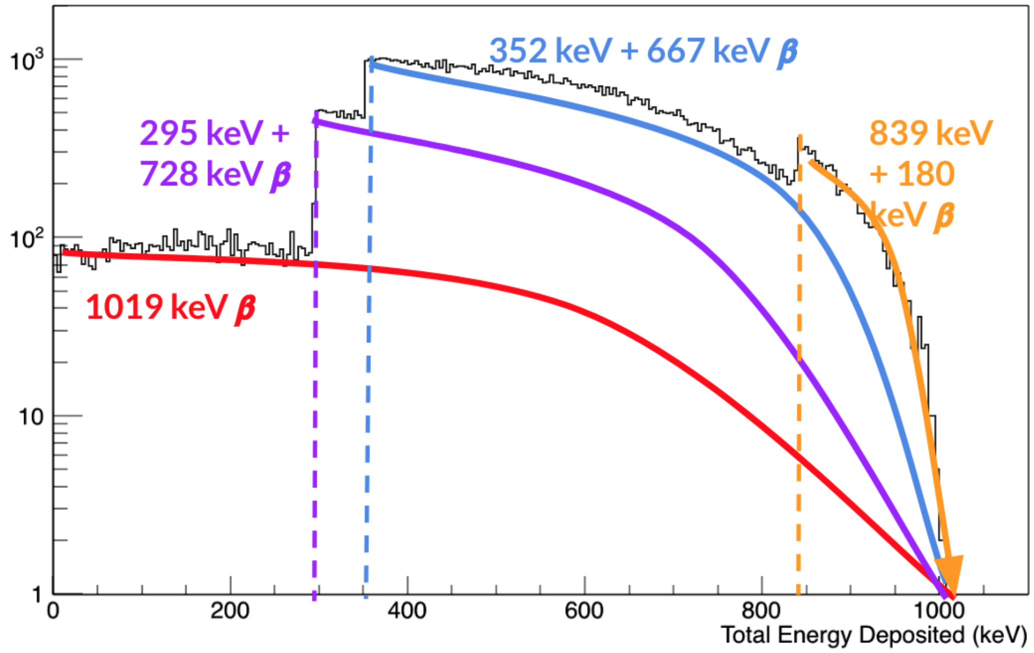


Figure 4.5: Plot showing the simulated decay spectrum of ^{222}Rn using BACCARAT (black line), overlaid with energies caused by Internal Conversion [153].

beta decay has occurred. This interaction can cause the inner electrons to be ejected, which then causes a hole to appear. This is then filled by electrons from the shell above, with the process being repeated by subsequent holes being produced within inner shells. This process leads to x-ray and auger electrons being released.

The beta decay spectrum of ^{214}Pb is the main component of background for LZ. There are 3 main branching ratios for the decay of ^{214}Pb to ^{214}Bi. The first branch is called the naked-beta decay. This occurs 9.2% of the time, and is the decay that affects the WIMP search ROI the most. This is as this decay does not have an associated γ-ray, hence vetoing this event is difficult.

The next decay which affects the shape of the ^{222}Rn decay has a branching ratio of 41.1%, where a step at 295 keV occurs, followed by the remaining energy being shared for beta decay. Another jump in the rate seen in figure 4.5 occurs at energy 325 keV, and has a branching ratio of 46.5%. Finally, an increase occurs at 839 keV, with the remaining 180 keV going towards beta decay. This has a branching ratio of 2.8%.

4.2.3 Investigating the Radon Decay Chain

In order to fully understand the relationship between the progenies of ^{222}Rn and its daughters, it can be helpful to consider the following notation to help keep track of all the atoms and their decay constants:

$$^{222}Rn \rightarrow ^{218}Po \rightarrow ^{214}Pb \rightarrow ^{214}Bi \rightarrow ^{214}Po \tag{4.1}$$

$$N_0, \lambda_0 \rightarrow N_1, \lambda_1 \rightarrow N_2, \lambda_2 \rightarrow N_3, \lambda_3 \rightarrow N_4, \lambda_4 \tag{4.2}$$

Thus, the number of atoms of a given daughter of ^{222}Rn can be modelled by looking at the decay of its direct parent isotope. This can be represented as

$$\frac{dN_i}{dt} = \lambda_{i-1}N_{i-1} - \lambda_i N_i \tag{4.3}$$

The number of atoms for each daughter isotope and their number N_i can be calculated as time t evolves, by iteratively solving equation 4.3 starting with $i = 0$ representing ^{222}Rn, and taking initial starting conditions i.e. $\mathrm{t = 0, N_0 = xBq, N_{1-4} = 0}$.

Figure 4.6 indicates that the activities of all progenies of ^{222}Rn until ^{214}Po reach equilibrium within ∼4.5 hours after an initial activity of ^{222}Rn. Hence by measuring the activities of ^{222}Rn daughters, the initial activity of ^{222}Rn can be inferred, provided that equilibrium has been achieved. By measuring the activities of a Bismuth-Polonium (BiPo) decay, which has a known event topology within the LZ detector, the activity of ^{222}Rn can be measured.

4.3 Bismuth-Polonium (BiPo) Decay

As radon diffuses into the TPC from the materials, it decays into ^{218}Po and ^{214}Pb ions via alpha decay. The ^{214}Pb decays via beta decay to form ^{214}Bi ion, which

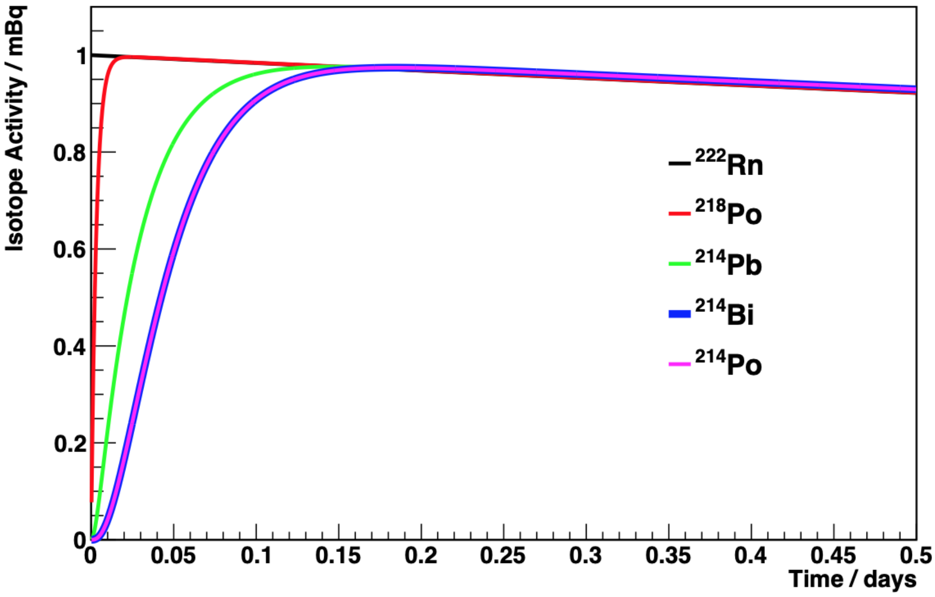


Figure 4.6: Plot showing the change in activities of ^{222}Rn and its daughters, with an initial ^{222}Rn activity of 1mBq. Plot adapted from [154].

then decays again quickly via beta decay to form ^{214}Po. These ions drift towards the cathode. The half-life of the ^{214}Po is 164.3 μs, and releases an alpha particle of energy 7.7 MeV to form the stable ^{210}Pb, which has a half-life of 22 years. This section will discuss how to infer ^{214}Pb activity, using measurements of ^{214}Bi and ^{214}Po and the fact that they can be seen together in a single event due to the short half-life.

4.3.1 Monte Carlo dataset

The datasets used for this analysis were the MDC3 background dataset, between dates = [2018-04-06 to 2018-04-19] hence comprising of 14 days of simulated WIMP search data. Initial coincidence cuts were applied to the dataset PMT$_{\text{coincidence}} > 2$, hence removing any pulses which had a PMT coincidence of less than 2 PMTs. When calculating the activity, the live times for each individual file was found, with the total livetime $t_{livetime} = 1137156$ s used. This was due to some of the files being resulting in errors during the processing step of the MDC3 simulation.

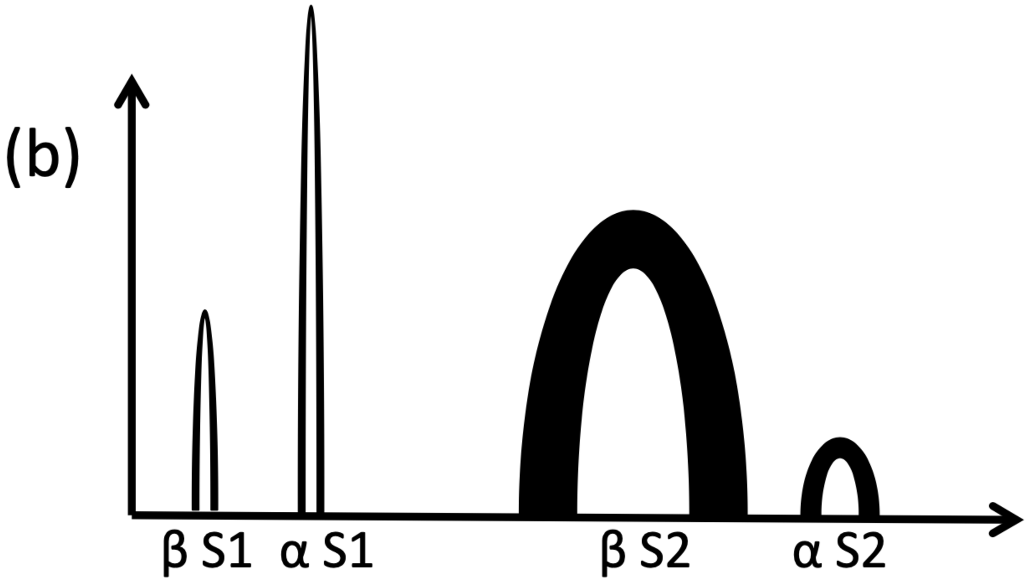


Figure 4.7: Image cartoon showing a ‘perfect’ BiPo event signal within LZ. Pulse heights and widths are not to scale, and only representative to relative sizes.

4.3.2 BiPo Tagging Algorithm

To tag the BiPo event present within the ^{222}Rn chain, characteristics of the decay were used as the basis for the tagging algorithm. A BiPo decay is defined as a beta decay followed by a relatively quick alpha decay of fixed energy Q = 7.7 MeV, as can be seen in the schematic in figure 4.7. It was assumed that all the ^{214}Bi would decay into ^{214}Po since there is a branching ratio of above 99%. To tag a BiPo, all S1 pulses per event were put into pairs with the time difference t_{diff} between each pair found. Since the ^{214}Po has a half-life $t_{1/2} = 164.3$ μs, S1 pairs which had a time difference less than 164 μs were used. The application of this can be seen in appendix A on line 473.

When determining the time difference to choose between the two S1 pairs, different time separations were considered, with the aim to find out how many BiPo events would be included within the time separation. The percentage of particles with a time difference t_{diff} between two S1s can be calculated as:

$$\frac{\int_0^{t_{diff}} \lambda e^{-\lambda t} dt}{\int_0^{\infty} \lambda e^{-\lambda t} dt} = 1 - e^{-\lambda t_{diff}} \tag{4.4}$$

Table 4.2: Table showing the percentage of BiPo events remaining for different t_{diff} values.

$t_{diff}[\mu s]$	Percentage within ^{214}Po	Percentage within ^{212}Po
0.5	0.21	68.5
1	0.42	90.1
5	2.1	99.9
10	4.1	99.9
100	34.4	100.0
164	50.0	100.0
200	57.1	100.0
300	71.8	100.0

where $\lambda = ln(2)/t_{1/2}$ and is the activity of the ^{214}Po.

Equation 4.4 can be used to calculate the percentage of BiPo events remaining within a $0 - t_{diff}$ between the two S1s of a BiPo, as can be seen in table 4.2.

Table 4.2 shows that keeping a large time difference between two S1s increases the percentage of BiPo events observed. However, by keeping a large time difference, the possibility of BiPo events leaking into another event time window is increased, as well as adding events that may not be BiPos.

BiPos found in the ^{220}Rn chain also have the same decay profile as a BiPo within the ^{222}Rn chain, however the ^{212}Po has a shorter half-life $t_{1/2} = 300ns$. Table 4.2 also shows the percentage of particles remaining within t_{diff} between two S1s for BiPo-212. After $1\mu s$, more than 70% of ^{212}Po have decayed. Hence, the time difference between ^{214}Po events was changed to be $[1 - 164]\mu s$.

The false acceptance rate (FAR) and false rejection rate (FRR) for applying a timing cut of $t_{diff} = [1 - 164]\mu s$ can be calculated when finding BiPos. The false acceptance rate is the amount of false events (i.e. not BiPos) which would accidentally pass through this timing cut. Since ^{212}Po has a short half-life (300 ns), a BiPo decay within the ^{212}Po chain would not be able to pass through the upper limit of this cut ($164\mu s$). However, more than 9.9% of ^{212}Po would remain after the lower limit of the timing cut has been applied ($1\mu s$). Hence, a false acceptance rate of 9.9% is associated with the timing cut $t_{diff} = [1 - 164]\mu s$.

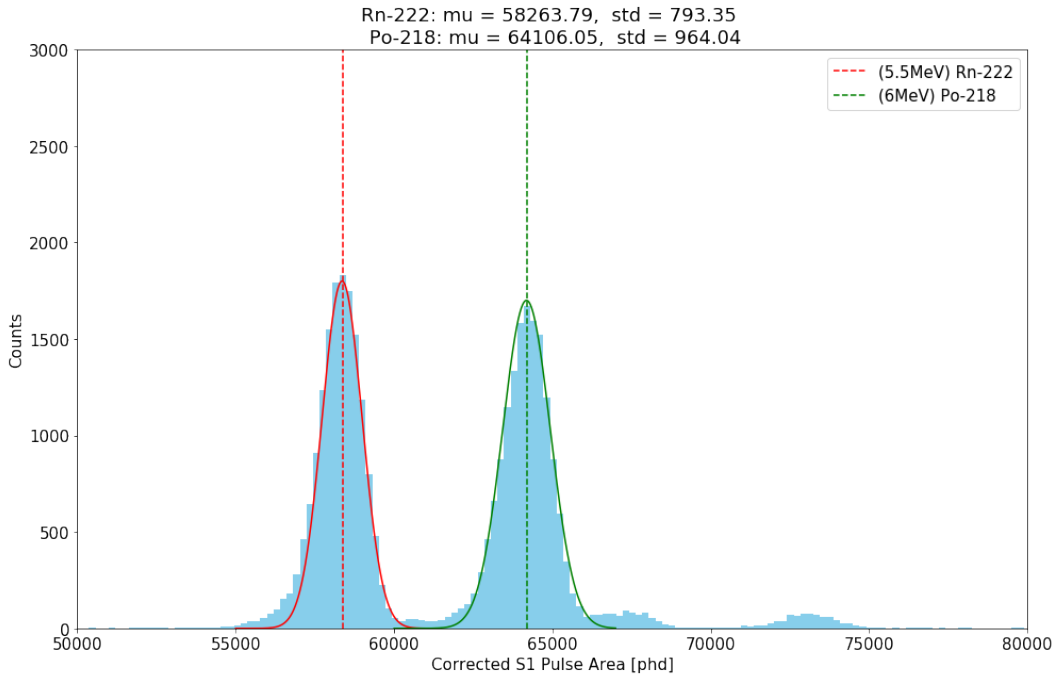


Figure 4.8: Histogram showing Pulse Areas for the known alpha decays of ^{222}Rn and ^{218}Po, each decaying with energy 5.5 MeV and 6 MeV respectively. Vertical red and green lines show peak S1 pulse areas in detected photons (phd). Gaussian fits are applied to the peaks, with a mean (mu) and standard deviation (std) calculated, and used to calculate the uncertainties in figure 4.9.

The false rejection rate is the rate at which correct BiPo events would be rejected by a cut. Since the half-life of ^{214}Po is $164.3\mu s$, the upper limit of the timing cut would remove 50% of actual BiPo events. The lower limit on the timing cut would also remove 0.42% of correct BiPo events from passing through the cut. Hence, this would give a FRR of 50.4% for the timing cut.

The second S1 pulse in each S1 pair would correspond to the α-decay of the ^{214}Po, hence having a known energy of $\mathrm{Q} = 7.7$ MeV. To estimate the pulse area of this S1, known pulse areas of two α-decays corresponding to the ^{222}Rn and ^{218}Po, as seen in figure 4.8 were used. The pulse area was then extrapolated by assuming a linear relationship between the pulse area and energy of the α-decay of the form

$$\mathrm{S1[phd]} = \mathrm{LY} \times \mathrm{Energy[MeV]} \tag{4.5}$$

Hence, the S1 pulse area of an α-decay of the ^{214}Po was estimated to be $85000 \pm$

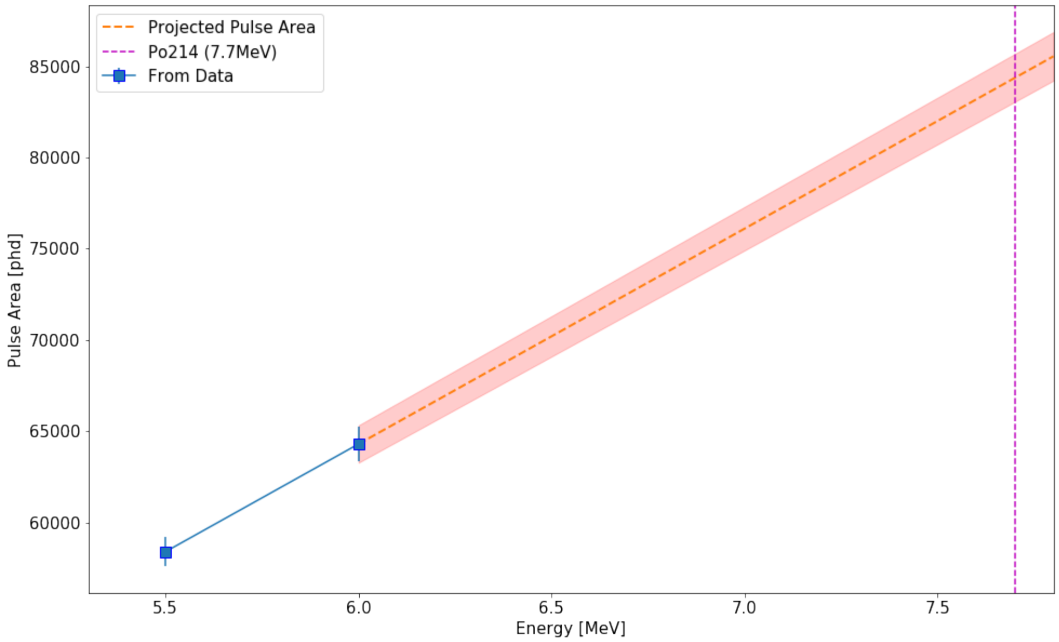


Figure 4.9: Plot showing pulse areas of ^{222}Rn and ^{218}Po (blue dots) taken from figure 4.8, with the error bars being the width of the fitted Gaussians. These are then extrapolated to find the Pulse Area in detected photons (phd) for the 7.7 MeV α-decay of ^{214}Po. Red shaded region shows the $\pm\sigma$ uncertainty in the projected values of the pulse area of ^{214}Po. This gives a S1 pulse area for ^{214}Po of 85000 $\pm$ 960 phd.

960 detected photons [phd]. Since the second S1 would correspond to the alpha of the ^{214}Po of energy 7.7 MeV, its pulse area was used to limit which areas would correspond to the ^{214}Po. Hence, a pulse area limit of 85000 ± 960 was placed on the second S1 pulse pair.

The first S1 pulse in a BiPo would correspond to the ^{214}Bi β-decay, therefore there was a range of energies that the associated β particle would have up to $Q = 3.28$ MeV. However, this is the upper bound, with the energy of the ^{214}Bi being shared between the beta particle and the antineutrino produced in the decay.

4.3.3 Cuts to S1 pulses

Top-Bottom Asymmetry (TBA) was also used to distinguish between different α-decays. TBA is the ratio between the light collected the top and bottom PMT arrays, and is defined as:

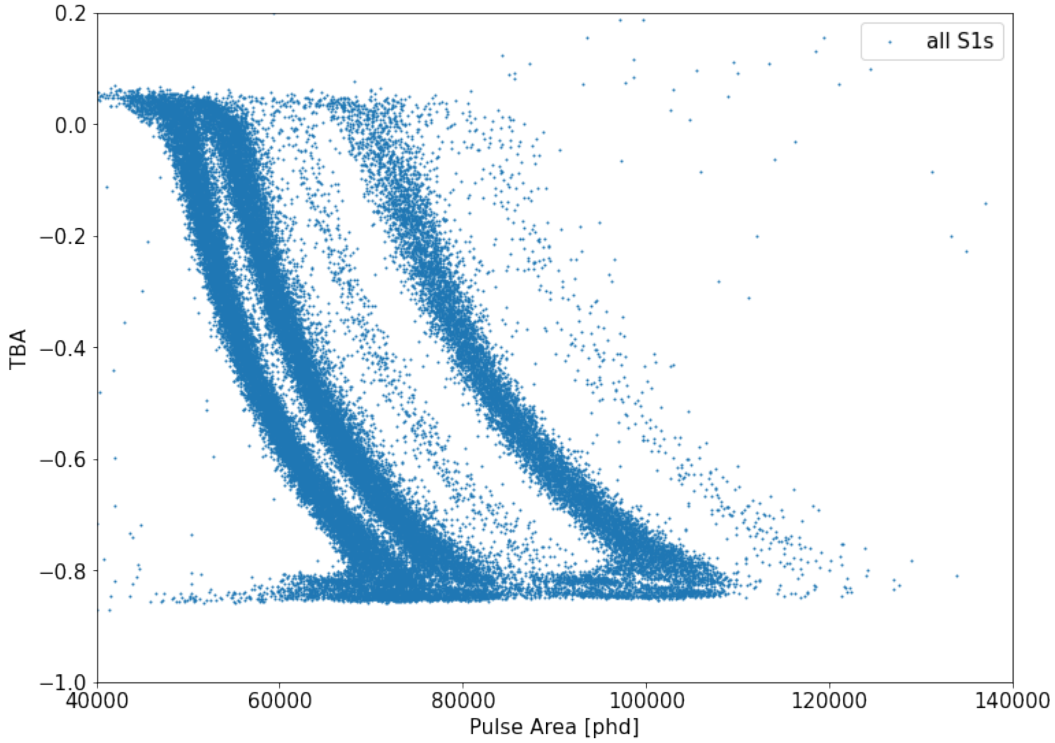


Figure 4.10: Plot showing Pulse Area vs TBA for all S1 pulses. Each vertical band represents an α-decay of ^{222}Rn daughters, with the first band due to the 5.5 MeV α-decay of ^{222}Rn, the second band representing the 6 MeV α-decay of ^{218}Po, and the third band representing 7.7 MeV α-decay of ^{214}Po.

$$\text{TBA} = \frac{\text{Top} - \text{Bottom}}{\text{Top} + \text{Bottom}} \tag{4.6}$$

where Top and Bottom is the light collected in the respective PMT arrays. Since high energy α-decay would produce large amounts of light, TBA can be used to estimate where in the TPC an α-decay took place. A large Top value would imply an α-decay near the top of the TPC, which would correspond to a TBA approaching 1. A TBA near -1 would imply an event near the bottom of the TPC.

However, since LZ has a gas-liquid interface near the top of the TPC, there would be internal reflection of the light produced from an S1 pulse, hence TBA values would be expected to be nearer to -1 for high energy α-decays, as more light would be collected at the Bottom PMT arrays. This can be seen in figure 4.11 which shows pulse area vs TBA. The majority of the S1 α-decays have negative TBA

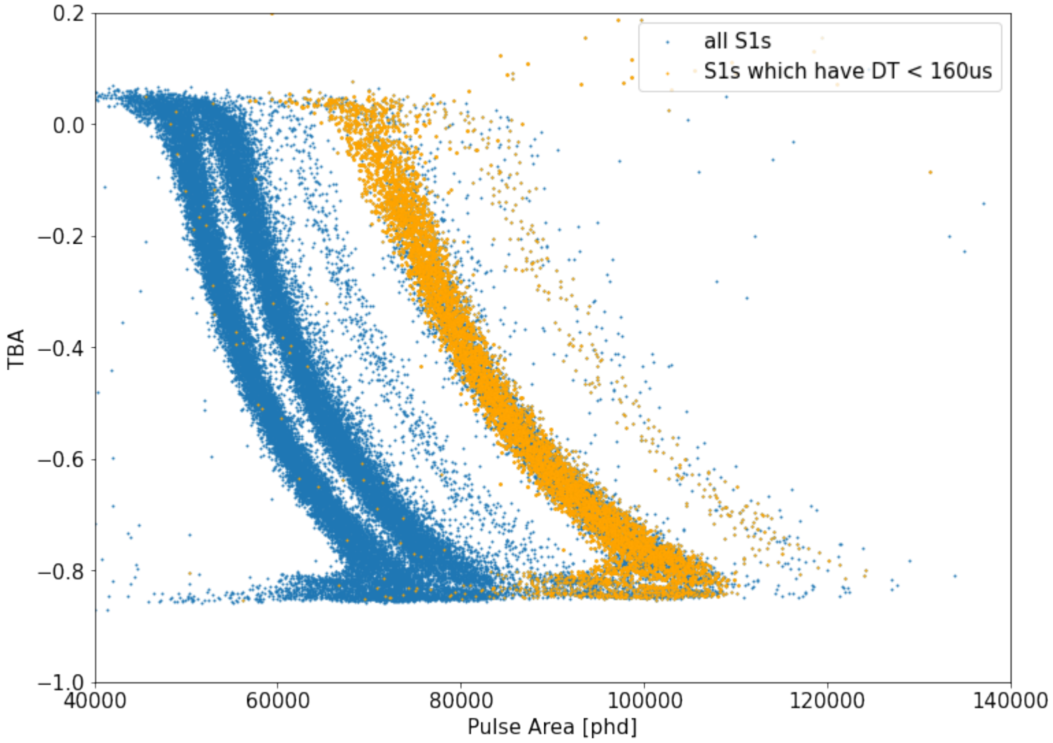


Figure 4.11: Plot showing Pulse Area vs TBA. S1 pulses within an event are paired together, with the time difference t_{diff} calculated. Orange points show events in for which t_{diff} is within $[1, 164\ \mu\text{s}]$, thus corresponding to a BiPo decay.

values.

The vertical bands seen in 4.11 were identified to be the α-decays found within the ^{222}Rn decay chain, with the orange band showing all S1 pulse areas of the second S1 between all the S1-pairs with t_{diff} within $[1, 164\mu s]$. Hence, only S1 pairs in which the second S1 pulse would be within the orange region seen in figure 4.11 would go onto the next stage of the BiPo tagging algorithm.

By looking at figure 4.11, the bands exist as bands of finite pulse areas due to the definite energies of the α-decays within each ^{222}Rn isotope. The pulse areas shown by the orange bands are the α-decays of ^{214}Po and ^{212}Po, each belonging to the radon and Thoron chains respectively. These two bands within the TBA vs pulse area plot would be expected to occur due to the unique event topology of a BiPo event. The relative number densities of the ^{214}Po and ^{212}Po bands in figure 4.11 (shown more clearly as the red and green bands in figure 4.15) are also indicative of the number of BiPo events expected to pass the initial t_{diff} cut.

Although α-decays would be expected to have a definite pulse area value due to the α energies being definite, a variation within the S1 values is observed relative to the TBA values, as seen in figure 4.11. The majority of S1 events in the bulk of the Liquid Xenon are totally internally reflected and collected in the bottom PMT arrays. There is a dependence of the Light Collection Efficiency (LCE) with depth for S1 interactions within the Liquid Xenon. The bottom PMT array would see the majority of events due to the strong internal reflection of the VUV light occurring at the liquid surface. Hence, S1s occurring at the bottom of the TPC would have a higher detection efficiency compared to events occurring near the top [155].

4.3.4 Cuts to S2 pulses

A limit was also places on the S2s for each event passing through the initial S1 cut. This was done by placing a limit on the minimum peak amplitude that the S2 must have. Peak amplitude was chosen instead of pulse area for the S2 pulses, as it was often found that electron trains, which occur after a large S2 pulse, get misclassified as S2s. These misclassified S2s are large in time and short in amplitude, hence would still have an S2 value comparable to an S2 associated with a BiPo.

Figure 4.12 shows the number of S2 pulses per event after applying the initial limits on the S1 pairs from section 4.3.3. The histogram shows that there are a large number of S2s per event, hence only using the number of S2s as a cut would be insufficient.

However, the majority of these S2s would be misclassified electron-trains being classified as S2s, hence by placing a bound of peak amplitude $\text{pA} > 500\text{phd/s}$, figure 4.13 shows the remaining number of S2s per event. These are S2s which have a peak amplitude greater than 500 phd/s only.

Since the first S2 in a BiPo topology would correspond to the β-decay of the ^{214}Bi, it would have a range of values, as the energy released is shared between the β particle and anti-neutrino. Thus, for BiPo decays with a lower β energy, the associated S2

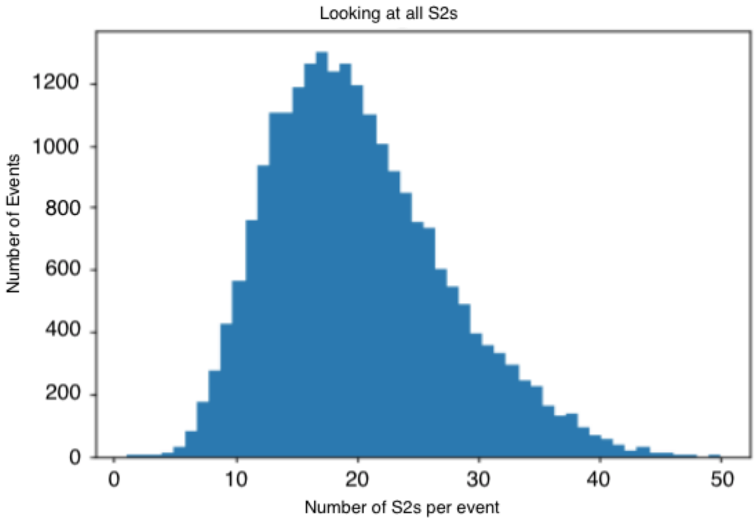


Figure 4.12: Plot showing the number of S2s per event.

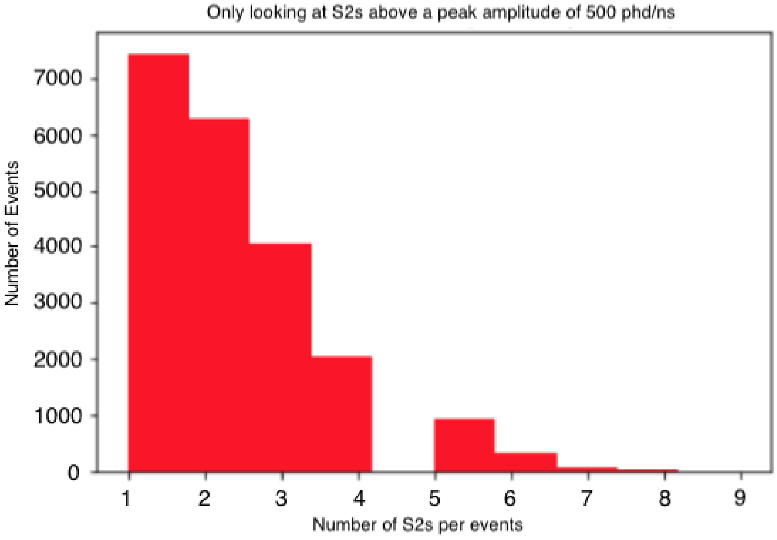


Figure 4.13: Plot showing the number of S2s per event.

Table 4.3: Table showing the number of events remaining after applying each successive cut.

Cut applied	Events remaining
No cut	1778204
t_{diff} only	34393
S2 cut only	13718
Total	**11008**

would also be smaller. Hence this could contribute to the number of events with only 1 S2 above 500 phd/s, as seen in figure 4.13.

4.3.5 Example of BiPo Event

Pulses that have been tagged by the BiPo algorithm can be seen in the LZ event-viewer in figure 4.14. Figure 4.14 shows an example of a clean BiPo event. The S1s are followed by two high energy S2s, with the time between the S1 and S2 pulses being the amount of time taken for the drift of the electrons within the TPC.

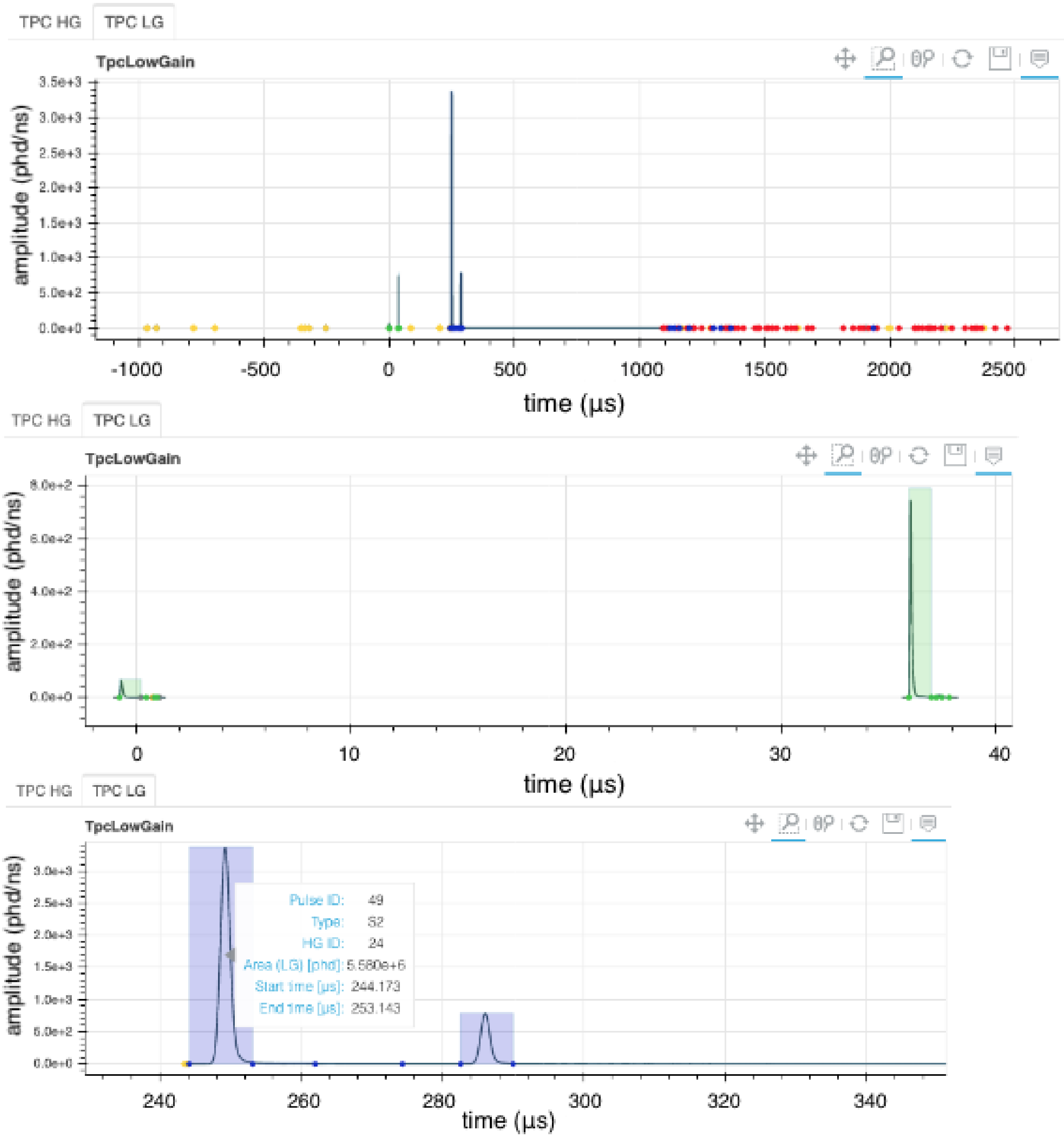
The time between the S1 pulses indicates that time taken for the ^{214}Po to decay, whilst the time taken for the first S2 pulse to arrive shows the time taken for the electrons to drift from the initial β-decay.

4.3.6 Analysis and Discussion

After applying all the cuts mentioned in this section, the total number of BiPo events found was $\mathrm{N_{BiPo}} = 11008$. By using table 4.2, this number would only account for 49.8% of actual BiPo events, hence the total number of BiPo events can be estimated to be double $\mathrm{N_{BiPo}}$, thus giving $\mathrm{N_{tot}} = 22018$. This gave an activity $^{222}\mathrm{Rn} = 3.58\ \mu\mathrm{Bq/kg} \pm 0.04\ \mu\mathrm{Bq/kg}$.

To account for the uncertainty in the measurement, the Poisson error was used, which gave an error of $\mathrm{e} = \pm 0.04\ \mu\mathrm{Bq/kg}$. Error due to the pulse classifier was also taken into account, as the LZap classifier has a pulse classifier that is 99% accurate. This gave an additional error of $\mathrm{e} = \pm 0.01\ \mu\mathrm{Bq/kg}$.

By using a cuts based method to analyse the decay of a BiPo event, an estimate for



Pulse ID: 49
Type: S2
HG ID: 24
Area (LG) [phd]: 5.580e+6
Start time [µs]: 244.173
End time [µs]: 253.143

Figure 4.14: Event trace showing a tagged BiPo pulse using the LZ Event Viewer. Top shows a whole BiPo event from the start of the trigger to the end, thus showing the length of a typical LZ event recorded to data. Middle plot shows the S1 events only of the BiPo, with the time different between S1 pulses being $\sim 36\ \mu s$. Bottom plot shows the S2 events corresponding to the BiPo event, with information regarding the first S2 pulse being highlighted. The event follows the 'perfect' BiPo schematic shown in figure 4.7.

the activity of ^{222}Rn was found to be $^{222}\mathrm{Rn} = 3.58\ \mu\mathrm{Bq/kg} \pm 0.04\ \mu\mathrm{Bq/kg}$. The region which this cut based method extracted (shown in orange in figure 4.11) is defined to be the BiPo region, hence all events within this region are defined to be a BiPo events.

Estimates for the false acceptance and rejection rates can be found by looking at the cuts applied when defining this region, with the $t_{diff} = [1, 164\mu s]$ cut giving an upper estimate for the FAR of 9.9%, and a FRR of 50.4%. However, the upper limit of the timing cut was based on one half-life of ^{214}Po of 164 μs, which corresponds to exactly half of the estimated number of decays that would be expected. Since the number of expected BiPo events is double the number of BiPo events found, the effect of the upper limit on the FRR will be negligible, thus giving a FRR only dependent on the lower $t_{diff} = 1\mu s$, with a value of 0.42%.

4.4 Machine Learning for BiPo Tagging

Machine Learning is an area within Computer Science that allows for computers to learn generalities from data, which can then be applied to new datasets. Although the 1950s saw the advent of the Turing Test - a test designed to test the cognitive abilities of a Machine, as well as the first Artificial Neural Network being developed, research - and thus excitement, has increased over the last few decades due to the increasing computational resources available to help train algorithms.

This section will apply some fundamental Machine Learning classifiers to the LZ dataset to help find BiPo events within the dataset.

4.4.1 Decision Tree Classifier

A Decision Tree classifier was used to train on 2000 BiPo events (shown in red in figure 4.15) from the ^{222}Rn chain. This training dataset was chosen at random from within the BiPo-214 region, which had a total of $\sim$ 6000 events, thus corresponding to $\sim$ 33%. 2000 non-BiPo events were also chosen to be within the training set, and were chosen to be at random from outside the BiPo region of interest.

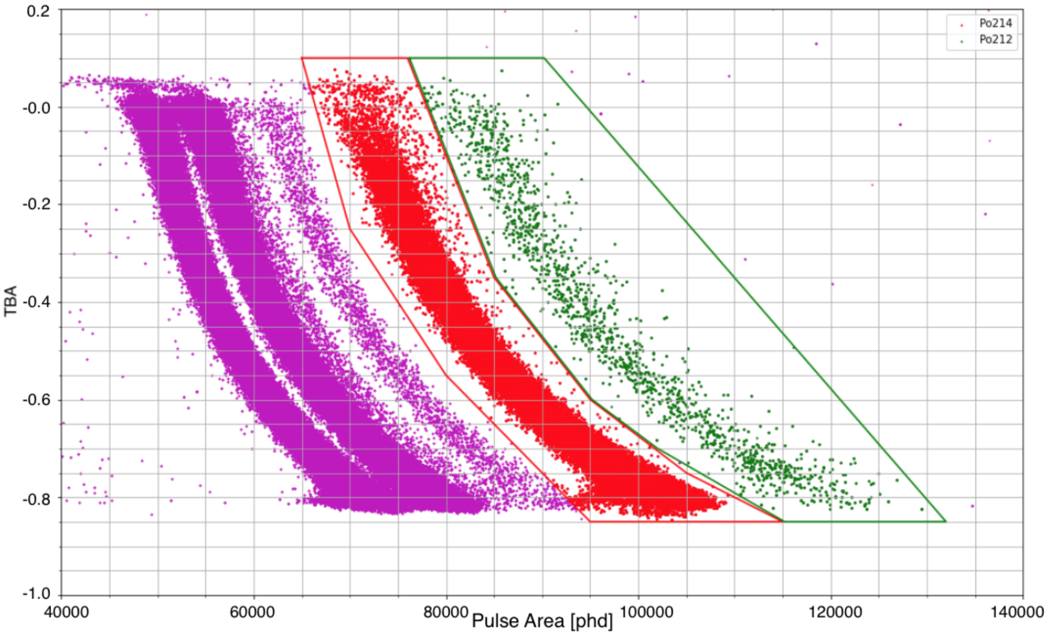


Figure 4.15: Scatter plot showing the different locations of the BiPo events, with red showing BiPo-214, and Green showing BiPo-212 events.

A Decision Tree classifier, discussed in chapter 3 of this thesis, was initially chosen as a classifier due to its versatility when being used with numerical and categorical data types. The rules that govern a Decision Tree classifier are easier to generate and understand, and the classifier requires little data manipulation when being used. Python `3.7.1` and `scikit-learn` [156] was used to implement the Decision Tree algorithm (and later the Random Forest algorithm) when developing a machine learning algorithm for BiPo analysis.

The training data was in tabular format, with Reduced Quantities (RQs) chosen which would represent a BiPo event as a whole. 19 RQs were chosen as features for the training set, and represented the features of the 2 S1s, as well as the time difference between them (see table 4.4 for list of RQs used). Any RQs which were event-specific, and not generalised for a BiPo event, were removed i.e. start time of S1s were not included, as these are event specific. However, the time difference between two S1 events was chosen as a feature, as these would be latent features for all BiPo events.

The trained classifier was tested on a subset of unseen data initially to optimise its

Table 4.4: Table describing the total number of Reduced Quantities (RQs) used as features for the inputs to the Machine Learning algorithms. A short description is also given regarding what they represent. `aft5` is time (ns) at which a pulse reaches 5% of the total area relative to the start of the pulse.

RQ name	Description
area_1	Pulse Area of first S1 pulse
area_2	Pulse Area of second S1 pulse
TBA_1	TBA of first S1 pulse
TBA_2	TBA of second S1 pulse
promptFaction_1	Prompt faction of first S1 pulse. Prompt faction is the ratio of the area of a pulse within the first 50 ns to the total area of the pulse.
promptFaction_2	Prompt faction of second S1 pulse.
peakAmp_1	Peak amplitude of the first S1 pulse
peakAmp_2	Peak amplitude of the second S1 pulse
peak_time_1	Time taken to reach peak amplitude for first S1 pulse
peak_time_2	Time taken to reach peak amplitude for second S1 pulse
pulseArea50_1	Area in fixed integration window ranging from `aft5` to 50 ns after `aft5` for first S1 pulse
pulseArea50_2	Area in fixed integration window ranging from `aft5` to 50 ns after `aft5` for second S1 pulse
pulseArea100_1	Area in fixed integration window ranging from `aft5` to 100 ns after `aft5` for first S1 pulse
pulseArea100_2	Area in fixed integration window ranging from `aft5` to 100 ns after `aft5` for second S1 pulse
rmsWidth_1	Width of pulse for first S1 pulse
rmsWidth_2	Width of pulse for second S1 pulse
FWHM_1	Full width half maximum for first S1 pulse
FWHM_2	Full width half maximum for second S1 pulse
time_diff_(ns)	time between both S1 pulses

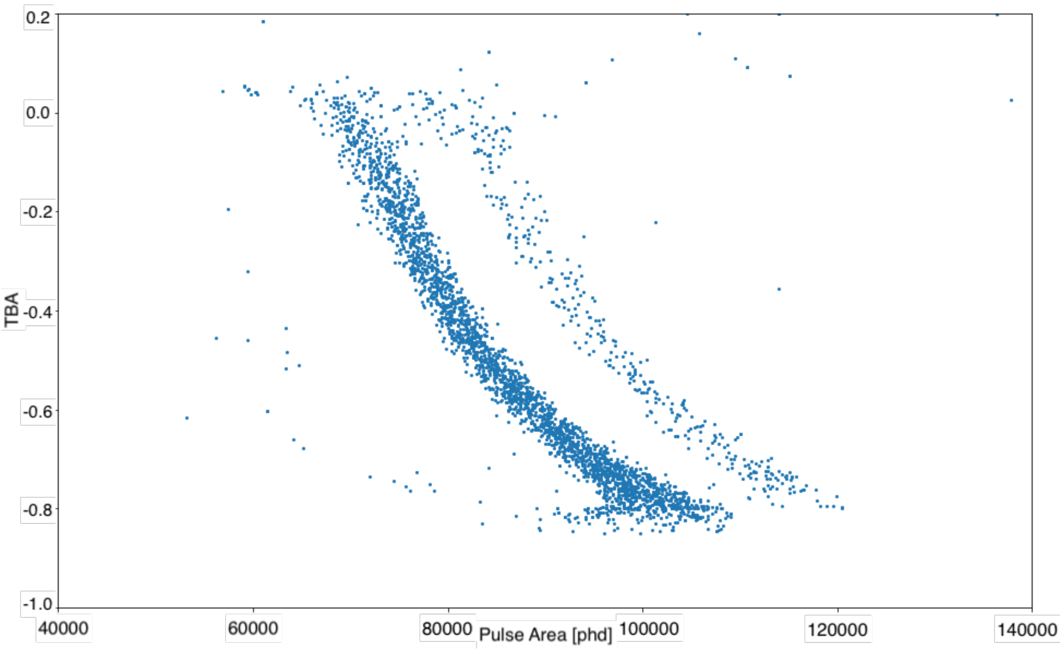


Figure 4.16: Scatter plot showing what the ML classifier classified as BiPo events. These events were not in the training set.

accuracy. By tuning the hyperparameters (such as the depth of the tree) of the Decision Tree, a classification accuracy of 98.1% was achieved. Accuracy is defined as correctly identifying whether an event was a BiPo or not, from the test dataset.

Figure 4.16 shows the results from the Decision Tree classifier with input test data being the remaining events. The results show what the classifier classified as BiPo events. These were events that were not within the training set.

The results show that the classifier was able to recover the remaining 4000 events from within the BiPo-214 spectrum, with only 74 events being misclassified as not a BiPo. Figure 4.16 also shows that the classifier was able to recover events from within the BiPo-212 band. Although these were not included from within the training set, the overall time-series event topology of a BiPo-212 is similar to a BiPo-214 event topology. This would suggest that the classifier was able to generalise the unique qualities of BiPo events. These generalities would also be similar for a BiPo-212 event, hence indicating that the classifier did not overfit the data.

The results also show events that are not within either of the red and green bands

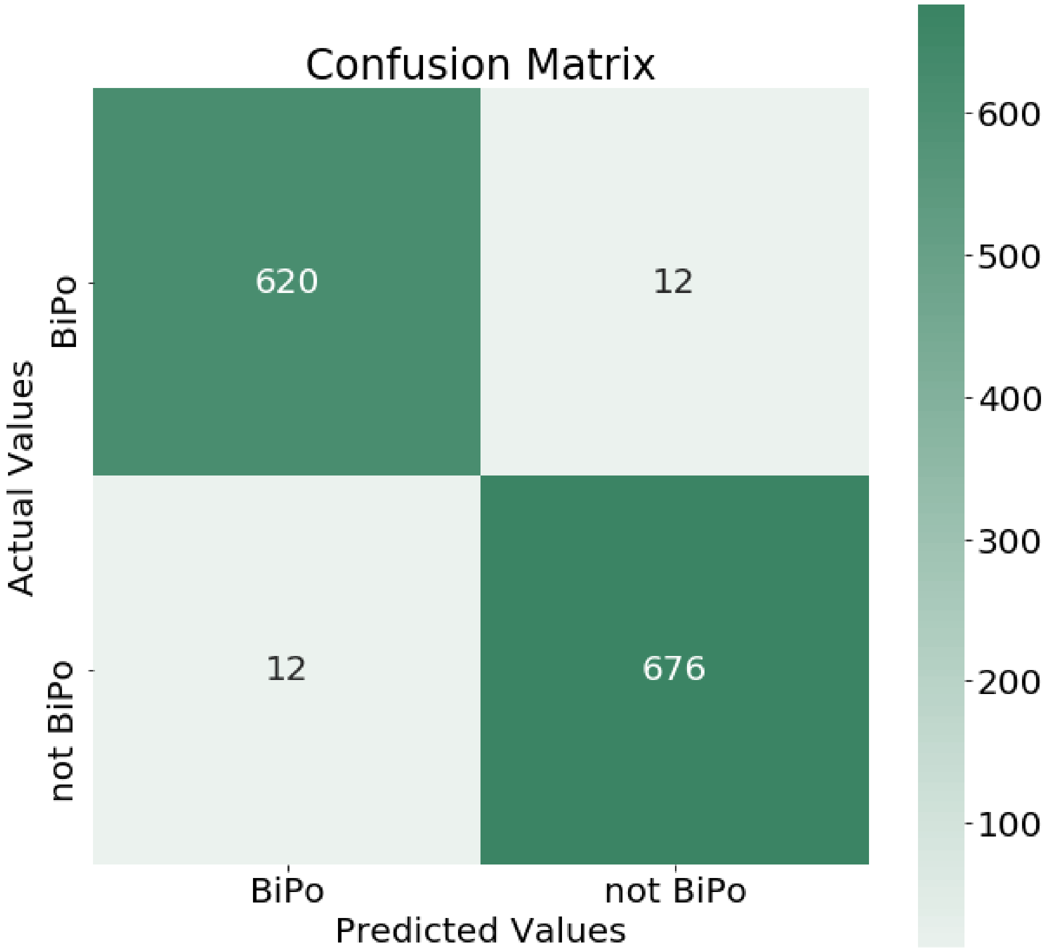


Figure 4.17: Image showing the confusion matrices for the Decision Tree classifiers. A confusion matrix is a specific table layout to allow for the visualisation of the classifications of an algorithm. Each box (starting from the top left, and going in a clockwise manner) represents the True Positive (TP), False Negative (FN), True Negative (TN) and False Positive (FP) of a classification output [136].

shown in figure 4.15. These are events that the ML classifier has classified to be BiPo events, but were completely missed out using a classical analysis. In total, 57 extra events were classified as BiPos by the Decision Tree classifier, but which were not within either of the ^{214}Po or ^{212}Po bands. This gave a total rate for $^{222}\text{Rn} = 3.60\ \mu\text{Bq/kg}\ \pm 0.04\ \mu\text{Bq/kg}$ using the Decision Tree classifier.

The confusion matrix can be used to calculate the FAR and FRR [136] of the Decision Tree classifier using the equations

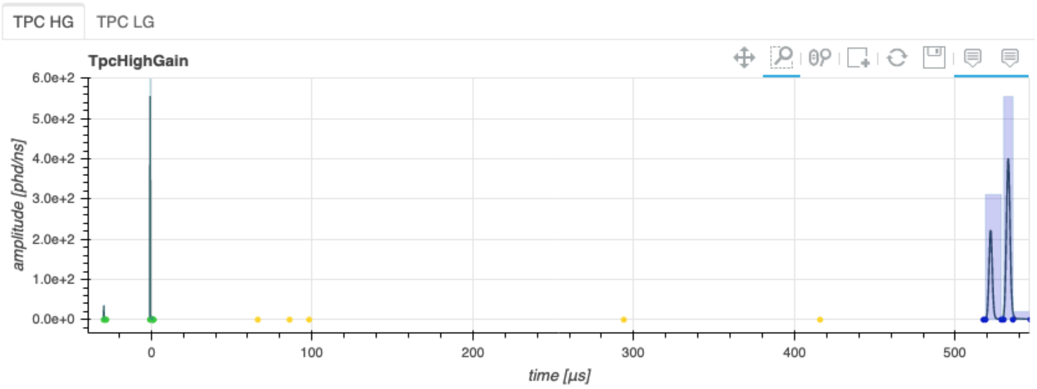


Figure 4.18: Event Waveform showing a classified BiPo event using ML which was not within the original ROI. The characteristic signature of a BiPo can be seen, with two S1s followed by two S2s.

$$\text{FAR} = \frac{\text{False Positive}}{\text{False Positive + True Negative}} \tag{4.7}$$

$$\text{FRR} = \frac{\text{False Negative}}{\text{False Negative + True Positive}} \tag{4.8}$$

Thus, by using the confusion matrix as seen in figure 4.17, a FAR of 1.74% and a FRR of 1.89% is found.

Figure 4.18 shows an example of such an event. From the waveform spectrum, a BiPo event is seen with 2 S1s which have a time difference of less than 164 μs, and the second S1 corresponding to the alpha of the ^{214}Po. These are followed by 2 S2s of high energy.

4.4.2 Random Forest Classifier

As well as a Decision Tree classifier a Random Forest classifier was also trained on the same data set to improve the classification accuracy, as well as compare between the classifiers. This is as Decision Trees can also suffer from overfitting, hence when being given new unseen data, the classifier may not be able to classify properly. This results when the classifier has not generalised the features unique to a BiPo, hence not being able to apply those generalisations to different datasets. An

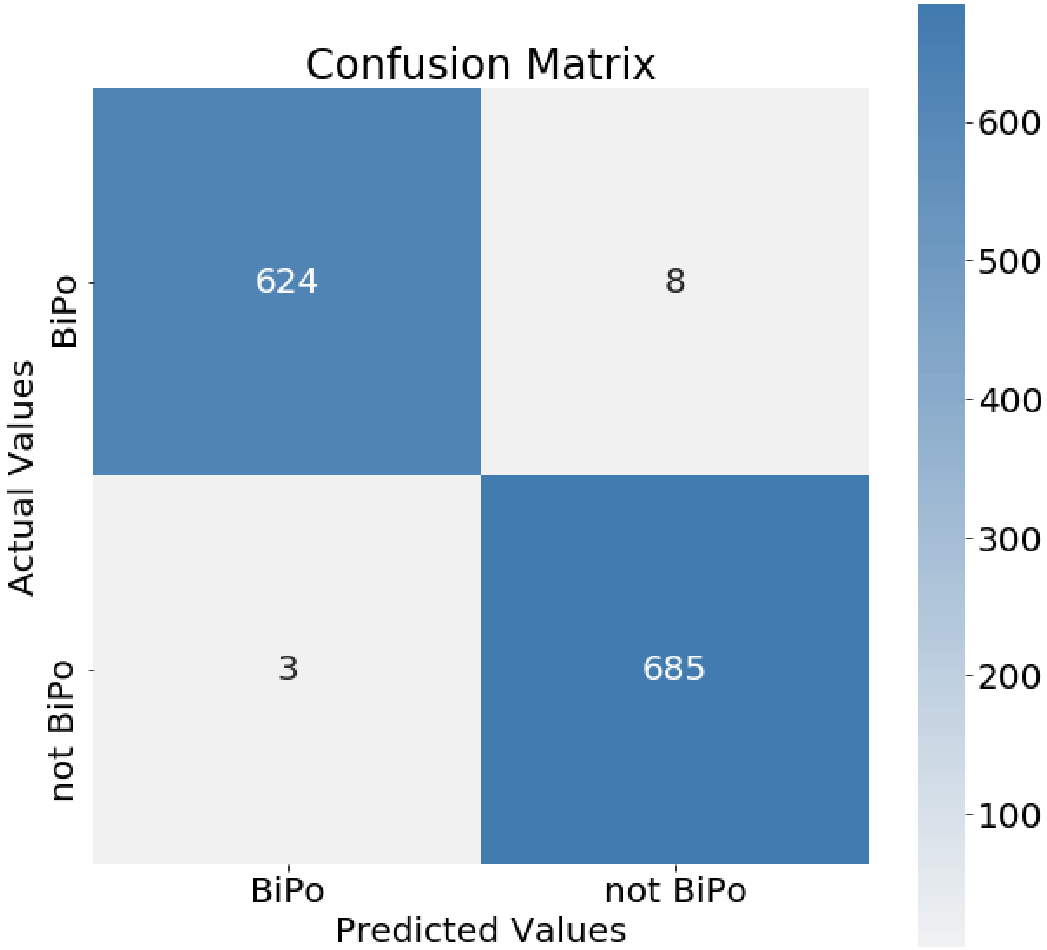


Figure 4.19: Confusion matrix for the Random Forest classifier tested on a subset of the test dataset.

example of this would be when a BiPo event occurs near the end of an event. In an overfitted Decision Tree, the event would not be classified as a BiPo, as the majority of BiPo events occur near the start of an event.

A Random Forest is a collection of Decision Trees being applied to the data, thus is an example of Ensemble Learning. Ensemble Learning is when many classifiers are applied to a dataset for a classification task, with the aim of averaging the results, thus leading to a more accurate classification.

By applying a Random Forest classifier to the same train and test dataset as that used for the Decision Tree Classifier, an accuracy of 99.2% was found. This is to be expected, as due to Random Forest being comprised of many Decision Trees, thus results would be better compared to a normal Decision Tree. The accuracy was calculated by comparing with the MCTruth information, with the confusion matrix for both ML classifiers shown in figures 4.19 and 4.17. Hence, by using the Random Forest classifier, a total rate for ^{222}Rn $= 3.62\ \mu\text{Bq/kg}\ \pm 0.03\ \mu\text{Bq/kg}$ was found. The smaller error is due to the higher accuracy of the Random Forest classifier being able to determine BiPo events, as compared to the Decision Tree classifier.

By using equations 4.7 and 4.19, a FAR = 0.42% and FRR = 1.27% is found for the random forest classifier. Comparing these rates to the decision tree classifier, the random forest classifier performs better, with a lower FRR and FAR. Both classifiers also have a lower FAR compared to the classical method, with the random forest classifier also having similar false rejection rates compared to the classical method.

4.4.3 Feature Importance

Figure 4.20 and 4.21 show the relative importance that the classifiers gave to each of the RQ features during the training process. Although both figures give importance to `pulseArea_50`, there are more features used within the Random Forest Classifier than the Decision Tree classifier.

4.21 shows that the pulse areas of the first and second S1 pulses were the most important features during the training process, with `area_2, pulseArea50_2 and pulseArea100_2` being the most important features. Although there is some degeneracy, as they all correspond to the pulse area of the second S1 pulse, they all have different definitions, using a combination of them would allow for a more accurate description of the pulse area.

Figure 4.21 also indicates that peak amplitude and rms-width are also important features when classifying BiPo events. These reduced quantities are not used in the initial classification process mentioned in section 4.3.3, however, are given importance for the Random Forest classifier. The importance of `time_diff` in both classifiers is lower than initially thought, as it would be assumed that the time dif-

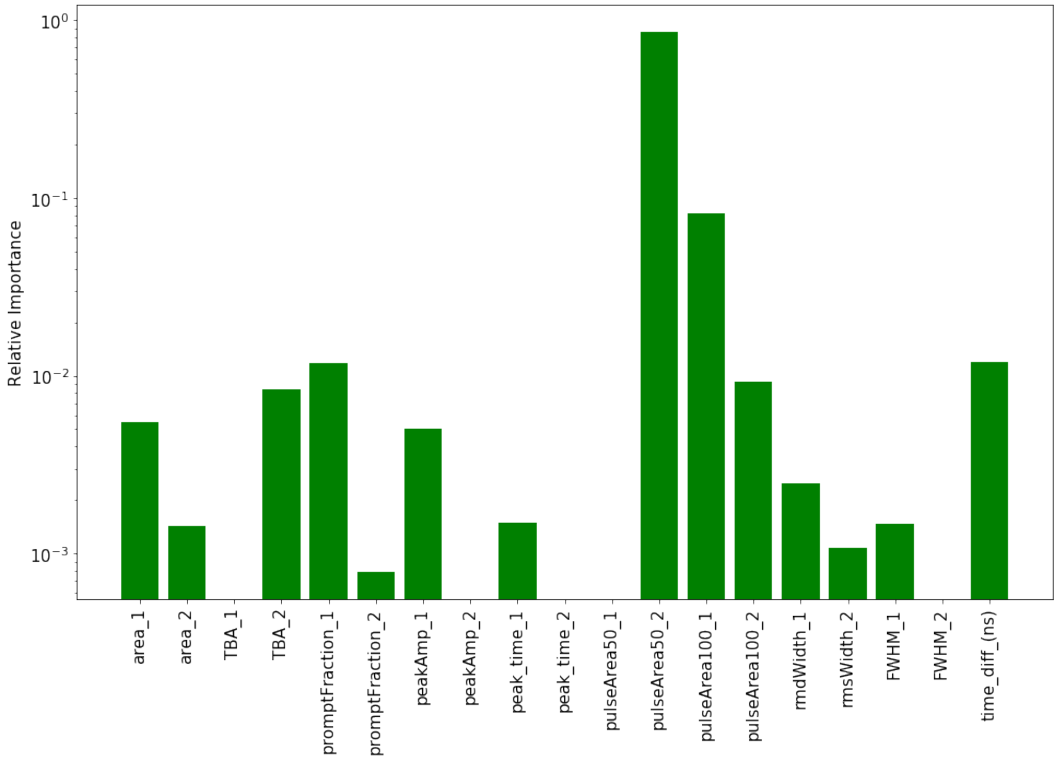


Figure 4.20: Plot showing the relative importance that the Decision Tree Classifier gave to each input features for Classification. A suffix '1' refers to information regarding the 1st S1 pulse, associated with the β-decay of the ^{214}Bi, and a suffix '2' refers to the S1 pulse associated with the α-decay of the ^{214}Po.

ference between the two S1 pulses would be the main discerning factor in BiPo and non-BiPo events.

4.4.4 Other Features

By looking at the relative feature importance in figure 4.21, the peak amplitude of the second S1 peak can be compared to the pulse area for the dataset. Figure 4.22 shows the pulse area vs peak amplitude for all the S1 points in the MDC3 dataset. The plot indicates that there are distinct regions within the phase space, with the location of the α-decays all occupying a similar region of space.

Since an α-decay is of high energy, the S1 pulse would have a large energy, hence this would result in a larger amplitude.

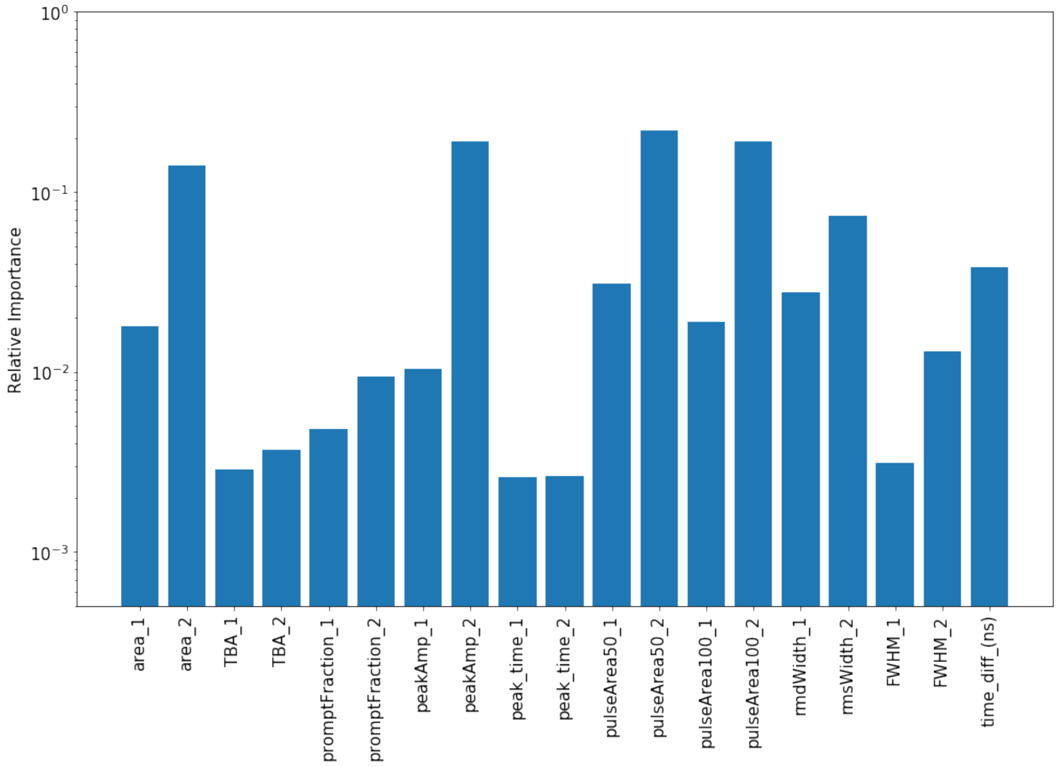


Figure 4.21: Plot showing the relative importance that the Random Forest Classifier gave to each input features for Classification. A suffix '1' refers to information regarding the 1st S1 pulse, associated with the β-decay of the ^{214}Bi, and a suffix '2' refers to the S1 pulse associated with the α-decay of the ^{214}Po.

Table 4.5: Table showing ^{222}Rn activities from the different methods.

Classifier Type	^{222}Rn Activity
Actual	3.65 μBq/kg ± 0.03 μBq/kg
Classical	3.58 μBq/kg ± 0.04 μBq/kg
Decision Tree (ML)	3.60 μBq/kg ± 0.04 μBq/kg
Random Forest (ML)	3.62 μBq/kg ± 0.03 μBq/kg

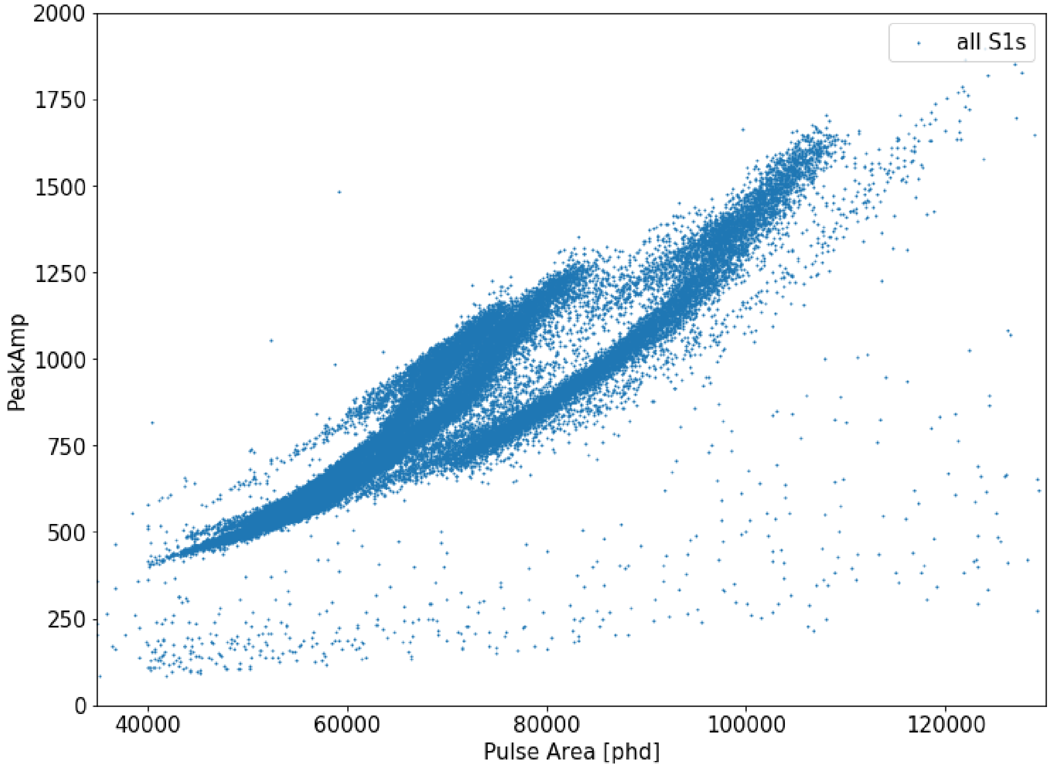


Figure 4.22: Pulse Area vs Peak Amplitude for all points.

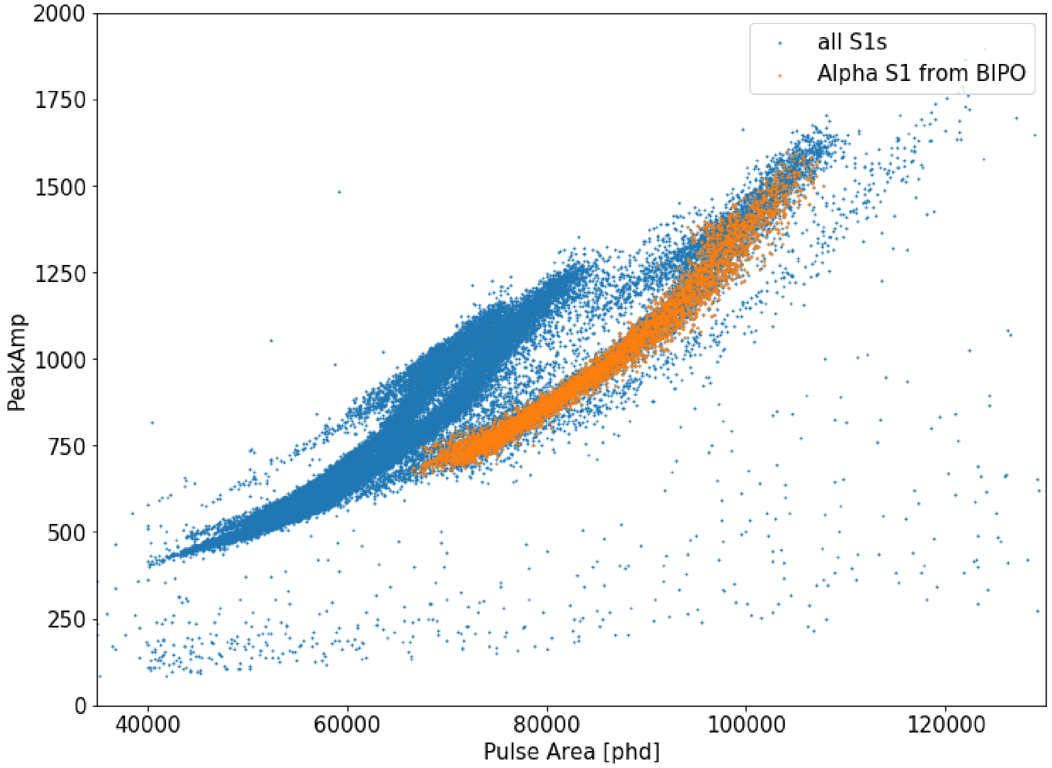


Figure 4.23: Pulse Area vs Peak Amplitude [phd] for all points, with orange points showing the S1 pulses relating to the ^{214}Po α decay. The ^{214}Po α decay has a higher ratio of area and amplitude due to the higher energy α decay (7.7 MeV) compared to the other decays within the ^{222}Rn chain. This results in a larger area, as more energy is deposited, but also a larger amplitude, recorded by the PMTs.

4.5 Conclusion

The study conducted in this chapter demonstrated two methods of tagging BiPo events to measure the activity of ^{222}Rn within the LZ detector, with the results shown in table 4.5. The first method employed classical techniques to find BiPo events. These methods applied known cuts to pulse area and TBA space to help characterize the unique decays present within a BiPo. BiPos are defined as being from within a certain region of interest (ROI) within the pulse area vs TBA phase space. Although these bands, as seen in figure 4.15 and figure 4.11, would represent the majority of BiPo decays, there is still the possibility of BiPo events being characterised to occur outside of this region. This could be due to the assumption of the initial timing cut being made between the events, with only the half-life being considered. Due to this cut, events that have a timing separation greater than this would not be seen.

The second method was to use Machine Learning to help classify BiPo events. Decision Tree and Random Forest classifiers were implemented, with a training set consisting of 2000 BiPo and 2000 non-BiPo events. By using more features than the classical method for tagging BiPo events, the classifiers - specifically the Random Forest Classifier, was able to tag events with an accuracy of over 99%. The classifier was also able to run over all the data set to then find more BiPo events not within the defined regions within pulse area vs TBA space, as used for the classical BiPo tagging algorithm. This demonstrates readiness to rapidly retrieve a ^{222}Rn rate from real data.

Due to the nature of WIMP Searches, and how less than 8 events per year are expected [91], being able to correctly identify any backgrounds which would be within the ROI is vital to increasing the sensitivity to WIMPs. Since the ^{222}Rn chain has a naked beta decay due to ^{214}Pb which contributes much of the backgrounds within the ROI, being able to tag such events would help in the sensitivity to WIMP searches.

By using Machine Learning, BiPo events that were not within the specified region defined classically were found. By tagging these events, the number of expected backgrounds within the ^{222}Rn chain can be quantified, and thus increase the sensitivity to WIMPs.

To improve on the Machine Learning method of tagging BiPos, more input features could be used. The removal of degeneracy between features could also yield a higher accuracy, as the classifier would then be able to give importance to more distinguishing features. The use of a larger training dataset could have been implemented to help the classifier learn more generalities from a BiPo.

Neural Networks could also be implemented as a way to analyse events on an RQ level, as well as working on the time series data of each event. This would mean using Long Short-Term Memory (LSTM) neural networks [157] to help classify events using the time-series event data.

Chapter 5

Position Reconstruction using ML

No leaf falls without His knowledge, nor is there a single grain in the darkness of the earth, or anything - fresh or withered, that is not written in a clear Record.

Quran [6:59]

One benefit of using a dual-phase LXe TPC is the ability to reconstruct the position of events in three dimensions. By being able to carry out this reconstruction, LZ can determine the physical origin of an event, which can then be used to determine the probability of an event being a WIMP-like particle or a background.

The z-position of an event can be calculated by measuring the time difference between an S1 and S2 signal. The *xy*-position of an event is determined by analysing the hit pattern of an S2 pulse incident on the top PMT array. However, due to difficulty in calibrations, coupled with reflections with the PTFE walls, *xy*-position reconstruction can be difficult to achieve, with events near the walls being more difficult to reconstruct.

5.1 Overview

This chapter will look into the Mercury algorithm [158], the current method of *xy*-position reconstruction employed by LZ, and initially developed for LUX. It will then discuss the limitations of such algorithms for events originating at the edge of the TPC. Results from a new method of *xy*-position reconstruction using convolutional neural networks will be discussed, with illustrative comparisons made between the leakage of events at the walls being reconstructed within the fiducial volume, being discussed. It must be noted that representative values for the parameters chosen by the author are used for the Mercury/LZap results presented here for ease of comparison with ML techniques. For exact prediction for LZ, the reader is directed to read [73].

5.2 Motivation

Liquid xenon time projection chambers benefit from the self-shielding feature that comes from using liquid xenon as a detection medium. By having a large atomic number (Z = 54) and a large density ($\sim$ 3 g/cm^3), liquid xenon is able to attenuate radiogenic backgrounds within the outer edge of the liquid volume. This thus creates a low background *fiducial* region within the centre of the detector, which is ideal for low energy recoils caused by WIMPs. Figure 5.1 shows this self-shielding effect within LZ in action, with a radioactively quiet region located in the centre of the detector shown in white.

Hence, accurate position reconstruction is needed to ensure that only events within the inner fiducial volume are analysed when calculating the sensitivity to WIMP masses, with *xy*-position reconstruction being vital in knowing where in the TPC an event has occurred. A fiducial volume of radial width $=$ 68 cm and height z = $\sim$ 134 cm has been taken for LZ, giving a fiducial volume of $\sim$ 5.6 tonnes.

The layout of the PMTs has been specifically designed to optimise the light collection efficiency, with the S1 signal being primarily detected in the bottom array. Hence the layout of the bottom array is mainly in a hexagonal array, with the PMT

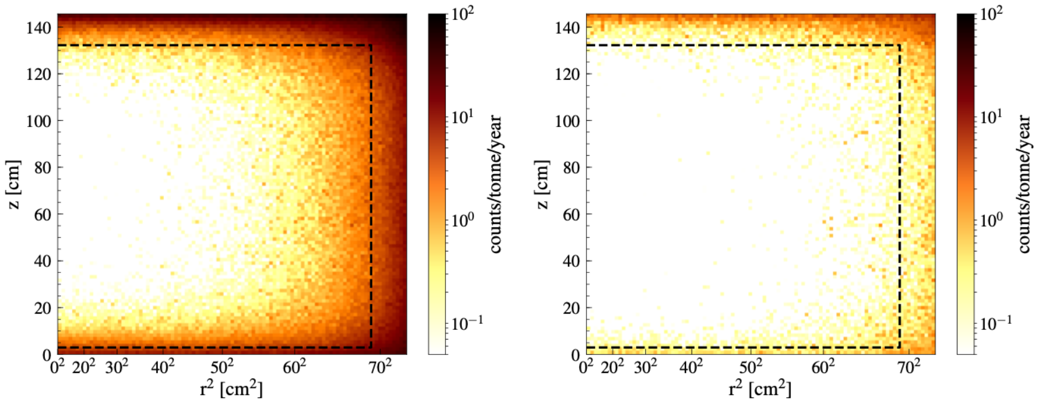


Figure 5.1: Image showing the effect of the fiducial volume in reducing the number of events within the centre of the LXe TPC. Left shows the NR event scatters with no vetoing; and right shows the same input data with Outer Detector (OD) and Xenon Skin vetoing taking place. The integrated counts for each are 1.03 counts / 1000 days with veto system in place, and 10.4 counts / 1000 days with no veto system in place. Taken from [91].

faces covering $\sim$ 54% of the bottom array area [73], with the remaining 46% being covered by PTFE. The top array is mainly used for *xy*-position reconstruction and collects the majority of the S2 light released in an interaction. This can be seen in figure 5.10 which shows an image of the bottom array.

The PMT placement in the top array is thus in a hexagonal pattern, with a circular ring of PMTs along the outer edge to improve the position reconstruction along the walls. There is also an overhang of the PMTs along the outer edge for the TPCs to further improve *xy*-position reconstruction along the edges of the TPC. Figure 5.9 shows an image of the top array.

However, even with the current layout of PMTs within the top array to optimise the *xy*-position reconstruction accuracy, this chapter will present a new method which will improve on the current position reconstruction, with an aim to improve the *xy*-position reconstruction specifically at the TPC edges (called walls).

5.3 The Mercury Method

Mercury was used in LUX as a new way to reconstruct the *xy*-position of events within the TPC. It is a statistical method in which Light Response Functions (LRFs) are found for each PMT. These are functions that characterise the response of the PMTs as a function of the position of the emission of the S2 light emitted at an event vertex. It relies on the maximum likelihood test to find the best output parameters to predict the response of each PMT caused by interactions at an initial arbitrary distance from the PMT.

The difference between Mercury and other conventional methods used is how the LRFs are obtained. In Mercury, the LRFs are obtained iteratively, by fitting to calibration data until the LRFs converge to below a threshold.

The LRFs are functions developed for each PMT i, and is defined as the probability that a photon from an S2-signal which is detected by the PMTs, is detected by PMT i. These can be different for each PMT, as the PMTs within the top PMT array are subject to different levels of reflectivity from the PTFE walls, with PMTs near the edge detecting more reflected light compared to those within the centre.

To calibrate Mercury for position reconstruction, ^{83m}Kr was used in LUX, with subsequent simulations developed and used for LZ. ^{83m}Kr isotope has a half-life of 1.83 hours which occur in two transitions of energy 32.1 keV and 9.4 keV, and can be seen in figure 5.2. This gives an S2 signal between 4000 phd and 20000 phd, which is also distributed along many PMTs in the top array. The signal is dependent on the depth of the ^{83m}Kr in the TPC, with decays lower in the TPC giving a smaller S2 signal. The decays also ensure that the PMTs do not get saturated, with the spread of the decays giving S2 signals below 10000 for any one PMT.

For accurate LRF, the distribution of the ^{83m}Kr must be uniform within the TPC. This was achieved for LUX [159], with simulations developed for LZ to ensure a uniform distribution of ^{83m}Kr events within the LZ TPC.

5.4 1D Convolutional Neural Network

To test possible improvements on the accuracy of the position reconstruction given by Mercury, convolutional neural networks were implemented. The aim was to

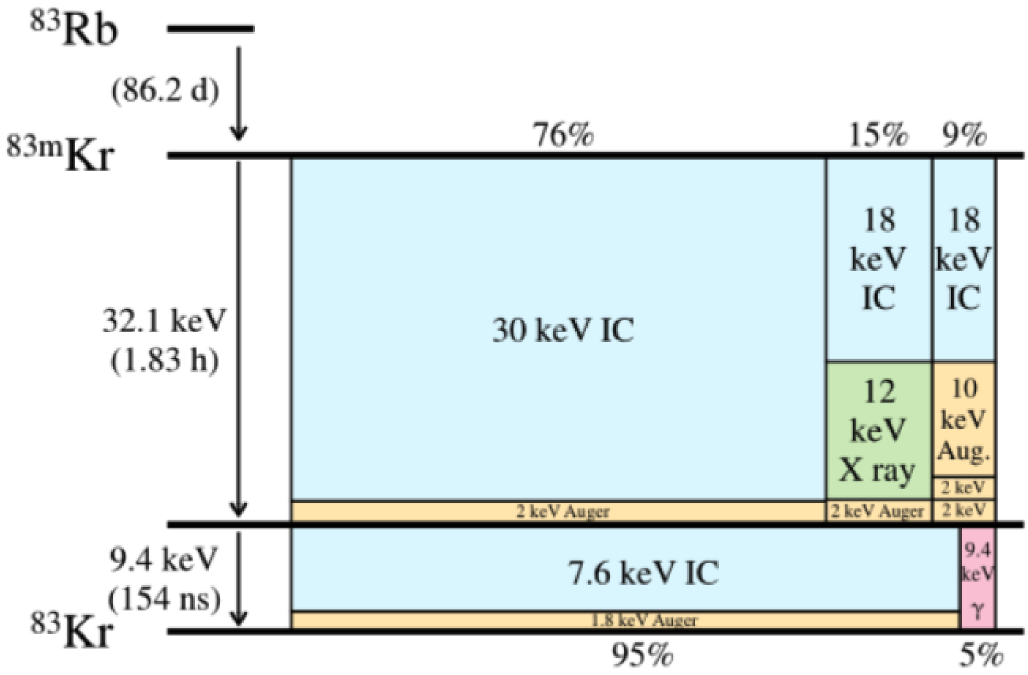


Figure 5.2: Image showing decay of Kr83m, used in the calibration of Mercury. Blue boxes indicate a decay due to Internal Conversion (IC), yellow boxes indicate a decay via Auger electrons [160]. The emission of Auger and IC electrons make up the majority of the decay method for Kr83m. Green and red boxes indicate a decay via the emission of X-ray and γ-ray photons respectively. Size of each box horizontally is proportional to its probability to decay. Vertical axis is represents the energy of the decay. Image adapted from [161].

train on a simulated dataset generated during the MDC3 data generation, with the training set specifically containing wall events of different origins. This can be seen in figure 5.3, which shows a subset of the training dataset locations. The training set contained $\sim$ 300000 events which were used for training.

Both BACCARAT and DER were used in the full chain to simulate the events to simulate the response of the detector running live. Events were analysed using LZap - the LZ analysis package which contains the Mercury method currently implemented. When running LZap to achieve the output given by Mercury, two Reconstructed Quantities (RQ) were available: one giving the position without a correction applied, and one with a correction. The difference is due to a mapping that is applied to map the uncorrected Mercury predictions to variations in the Electric-

Field within the detector. Hence the corrected positions represent the true position predicted by Mercury and is used for comparison thereof.

Convolutional neural networks were chosen due to their ability to generalise spatial features within the input data. This would not be possible if a feed-forward fully dense neural network was chosen, as the relationship between PMT locations in real space would not be modelled. Figure 5.4 shows an example 1D input to the neural network.

The data was initially represented as a 1D input as it has lower dimensionality compared to a 2D input. This was to examine whether these could be used in theory to reconstruct the position. Further work on a 2D input is demonstrated in the next few sections. The size of the input data was thus 253, with each dimension in data space representing a single PMT input. The PMTs were normalised to ensure uniformity throughout the training process.

5.4.1 Model Architecture

Many models were trained to try and determine the best architecture which would give an accurate output to the *xy*-position, whilst also taking into account computational limitations as well as training time. Four layers of convolutional neurons and max-pooling layers were used to initially work onto the dataset. These were then flattened, with a feed-forward network used to take the output from the convolutional layers, with three dense layers being used. The output layer had two neurons to give values of the x and y predicted position.

5.4.2 Comparing on Test data

The LZ TPC wall is defined to be at a distance r = 72 cm in radius from the TPC centre, with the current fiducial volume edge being at r = 68cm. To compare the predictions of the neural network with mercury, the original location of the event must be known. This can be found by using the MCTruth RQ location of the event, which is given as an output during the simulations and not seen during training or

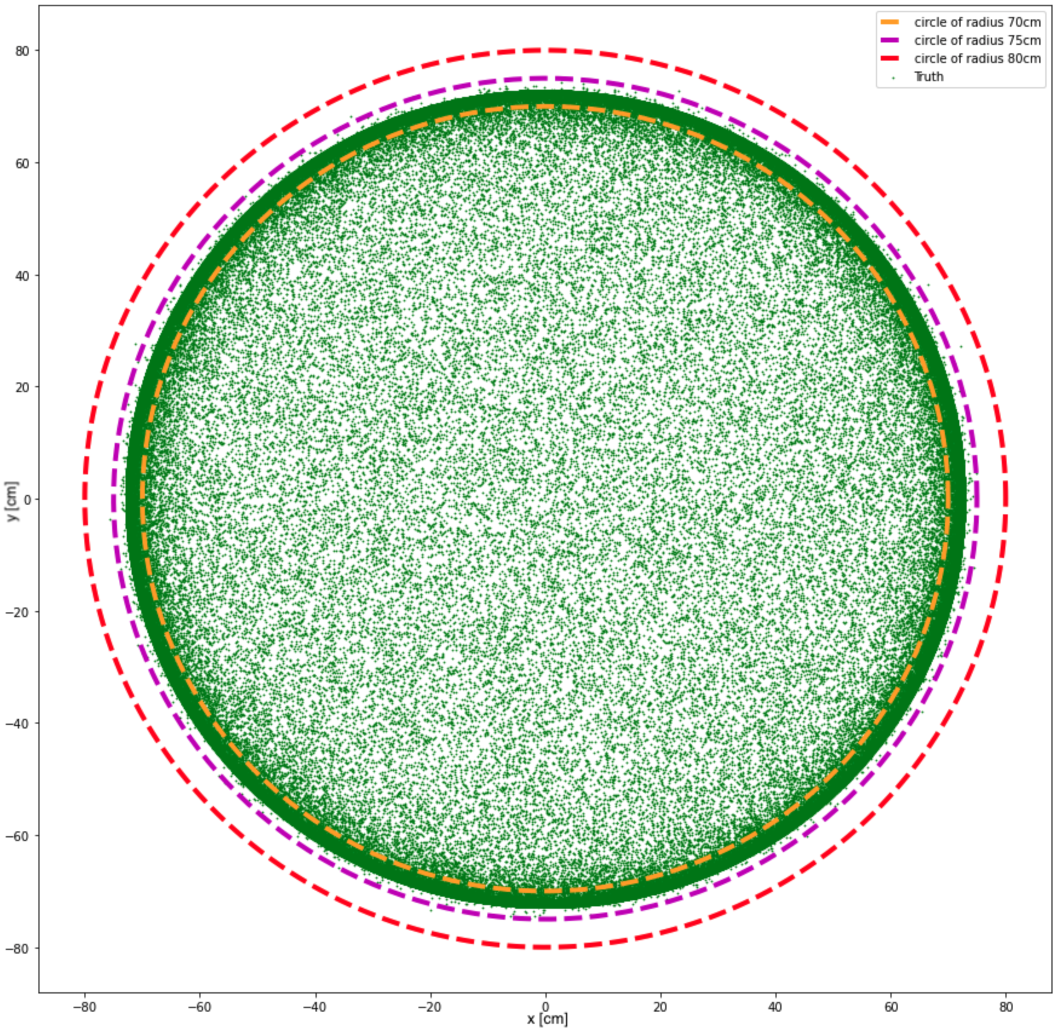


Figure 5.3: Image showing the input training data to both the Machine Learning methods. Axes are in cm, with a top-down view of the events in the xy space. Image shows 1/10th of the total number of events within training dataset.

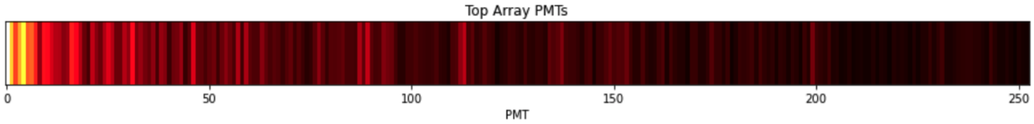


Figure 5.4: Image showing the 1D input to the 1D ML CNN of the Top array. Each PMT is numbered from the centre, with numbers going out in concentric circles towards the outer edge of the TPC. This means that events near the edge of the TPC may not necessarily be shown together using a 1D input.

by Mercury.

The Euclidean distance between the true position and predicted positions for ML and Mercury (called LZap in plots) were taken, with a larger distance between the truth and predicted position indicating a larger error in that position.

Figure 5.5 shows the average difference between predicted and truth position for Mercury (labelled LZap, and seen in orange) and machine learning (seen in blue). At smaller radii from the centre, Mercury performs better compared to the 1D convolutional network (labelled as ML). This is to be expected, as the number of points within the training dataset set was significantly lower for smaller radii, hence the neural network was not able to learn parameters needed for accurate position reconstruction at smaller radii.

However, as the radius increases, the average difference between a predicted position for ML decreases compares to Mercury, with Mercury (LZap) giving a larger difference in predicted position from truth position the further the distance is from the centre. This could be due to how nearer the edges, a full hit-pattern is not seen, with only edge PMTs giving responses. Hence, this could be one reason for the large difference in error between Mercury and the Truth position.

Wall events within Test dataset

One major source of backgrounds within LZ are wall events caused by charged radon-progeny which has undergone plate-out onto the surface of the PTFE detector walls. These radon progeny decay from the walls, with either the α-decay or the ^{210}Po daughter either decaying towards the centre of the TPC. This can lead to misreconstruction of the position of the event.

To compare only wall events within the test data, a cut of test events originating between $72cm < r < 73cm$ was taken, with the aim to compare the predicted position of both Mercury and ML. Figure 5.6a shows a plot of reconstructed position for ML for a section of the TPC. The reconstructed positions are closer to the original position of the wall, as compared to figure 5.6b, which shows a lot more events being

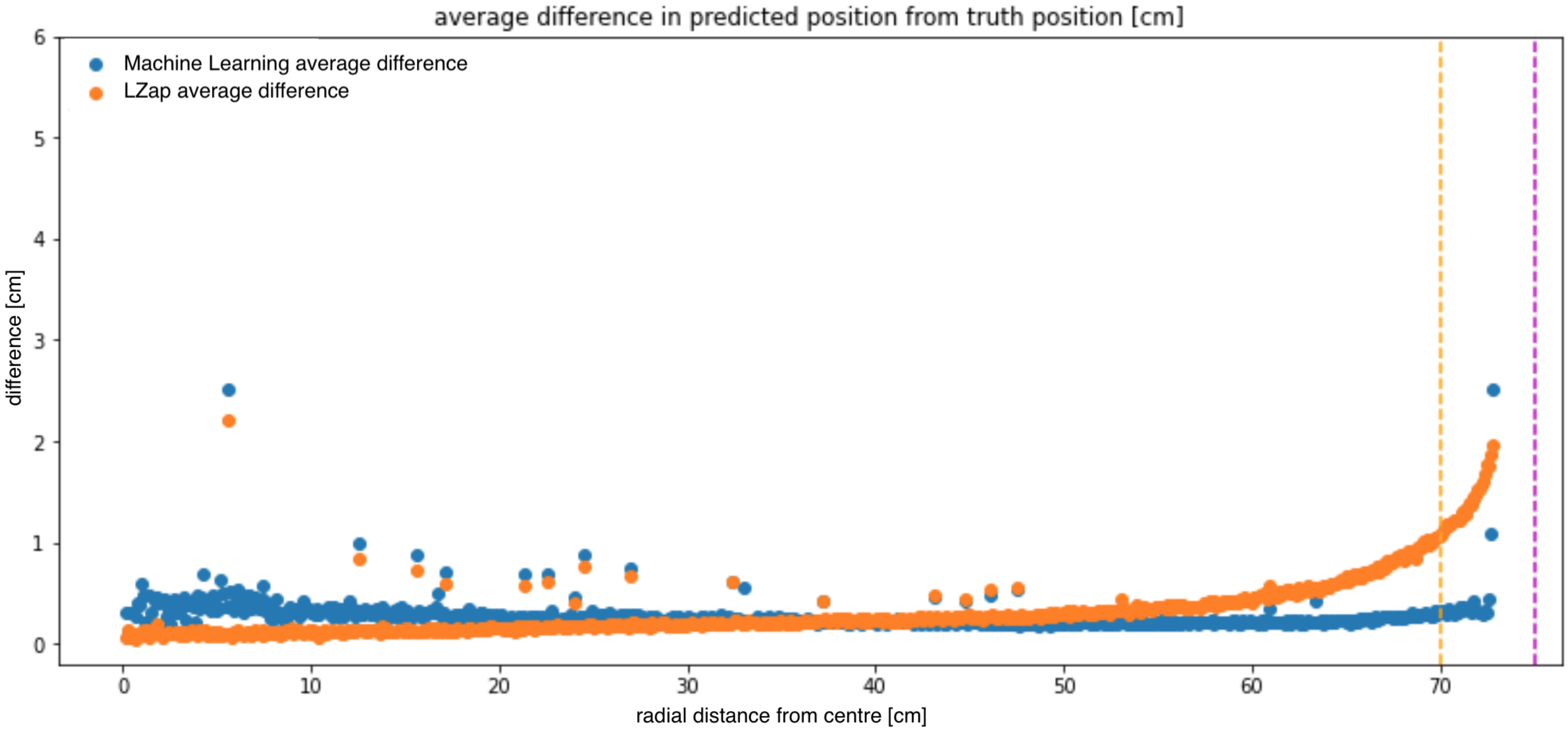
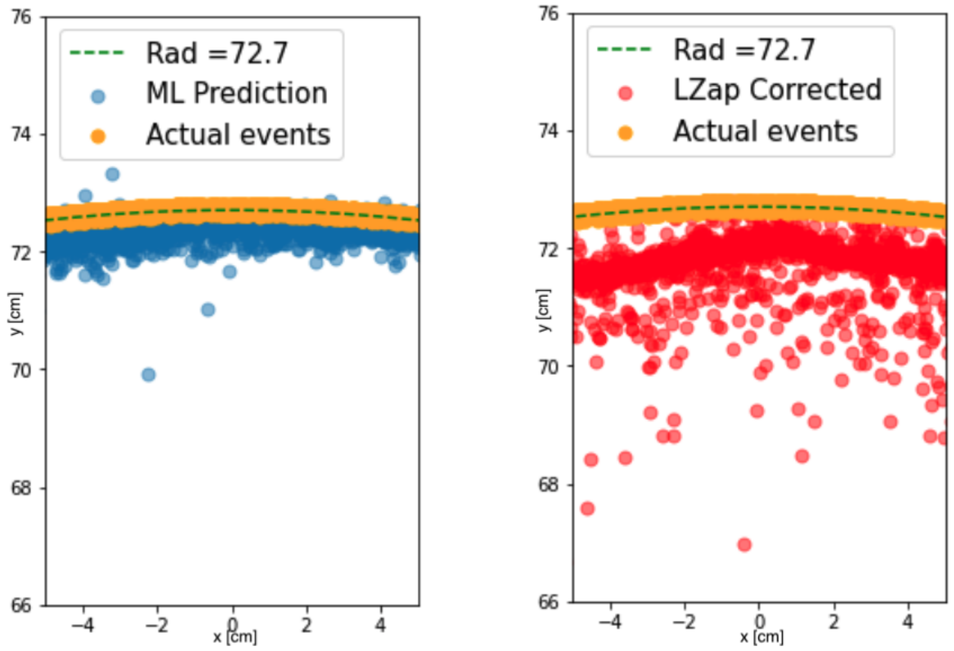


Figure 5.5: Image showing the average difference between the reconstructed and actual position for Mercury (orange) and ML (blue). Plot shows that above ∼ 40cm in radial distance from the centre, ML has a lower error in average distance between reconstructed and actual position compared to Mercury.



(a) 1D ML reconstruction on test dataset. **(b)** Mercury reconstruction on test dataset.

Figure 5.6: Image showing the events within the test dataset used for training (Orange) and reconstructed position due to ML (Blue) and Mercury (Red). Test dataset contained over 25000 data points.

reconstructed within the TPC.

The leakage of these events can also be found (see section 5.4.3 for more detail on leakage) in figure 5.7. The plot shows that the convolutional neural network has a better leakage compared to LZap, with the number of events that leak into the fiducial volume an order of magnitude smaller for the machine learning position reconstruction compared to LZap.

However, the test dataset contained S2 pulses of varying sizes, most of which would be below the threshold of S2 acceptance for WIMP searches ROI. These smaller sized S2s would be more difficult to reconstruct their position, with more leaking into the fiducial volume. To find a better way of comparing the leakages of ML and Mercury, a standard size of S2 pulses would need to be simulated at the wall. This will be discussed in the next section.

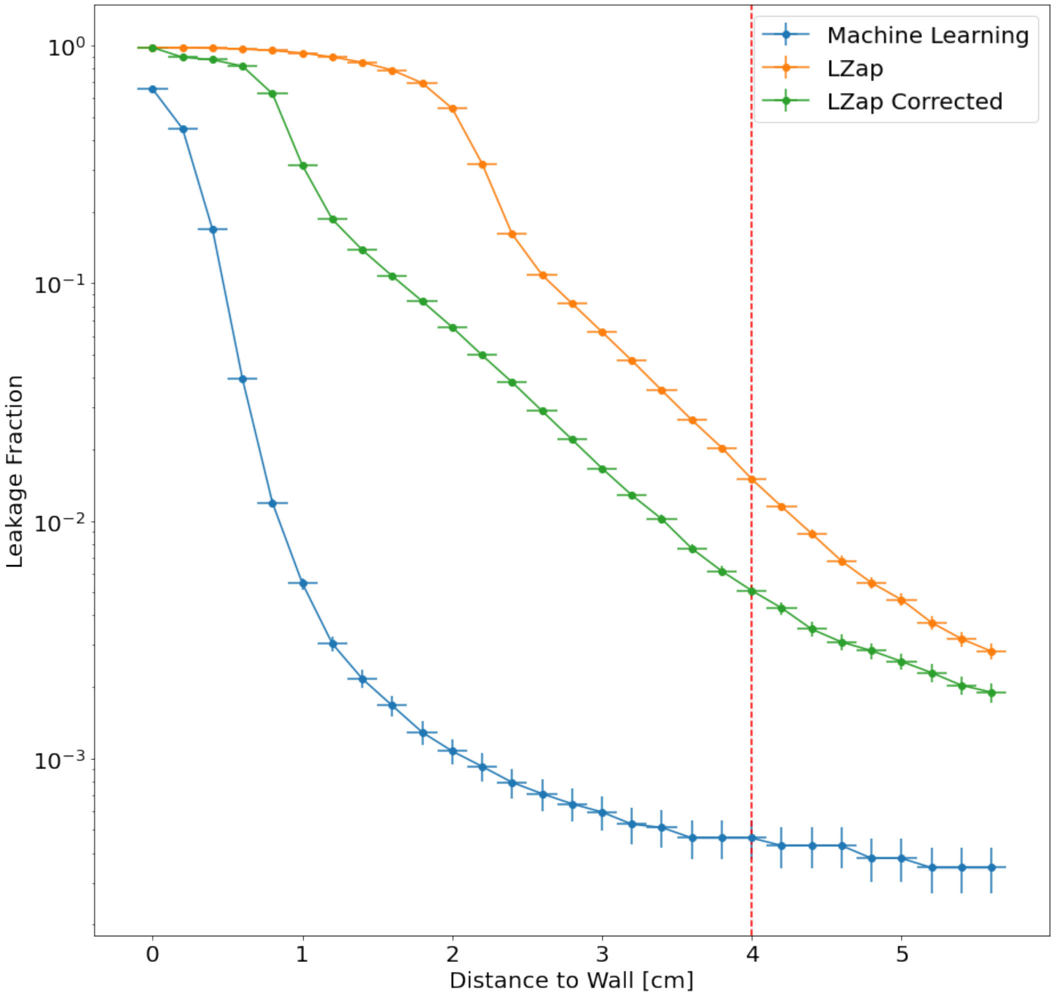


Figure 5.7: Image showing the leakage of ML and LZap and LZap corrected (Mercury) on the test dataset. The LZap corrected takes into account the drift of electrons towards the centre of the TPC. It is shown separately in this image, but in further images where only LZap in mentioned, the correction is applied. Plot with linear y-axis is shown in Appendix B.1.

5.4.3 Leakage

Another way to characterise the performance of a position reconstruction algorithm is by testing the number of events that originate at the walls of the TPC, but are reconstructed at a different position from the wall. A standard set of events that is used is the signal caused by 5 thermal electrons situated just below the LXe level, and originating at the wall of the TPC. The signal caused by these 5 thermal electrons would produce an S2 signal size of 400 detected photons (phd). In total, more than 10^6 events were simulated using BACCARAT of such events, with the aim to find to compare the leakages between the 1D convolutional network and Mercury (also called LZap in plots).

The reconstructed events given by the position reconstruction of these thermal electrons can then give a leakage fraction - the fraction of events that leak into the fiducial volume which originated from the walls. The fiducial volume is currently defined as being 4cm from the walls of the TPC.

Since the PMT array is radially symmetric, the leakage can be found by only simulating the thermal electrons from a single point at the wall. For convenience, the origin of the points is given at $[x,y,z] = [0,72,145.5]$. Since only the top array is used for xy-reconstruction, the bottom array PMTs are not considered.

Figure 5.8 shows the leakages of ML and Mercury for the signal caused by 5 thermal electrons. The plot shows that the leakage produced by ML is lower compared to mercury by over 3 orders of magnitude for a distance of 2 cm from the wall, indicating that a lower fraction of events are reconstructed within the fiducial volume for ML-1D compared to Mercury.

Figure 5.8 indicates that by using the 1D ML position reconstruction for wall events, a leakage of lower than 10^{-5} after 2 cm from the TPC wall can be achieved. This would give a fiducial volume of ~ 6.2 tonnes, which is 600 kg larger than the current 5.6 tonnes fiducial volume. However, a better leakage may be possible if a 2D input was used instead. This will be discussed further in the next section.

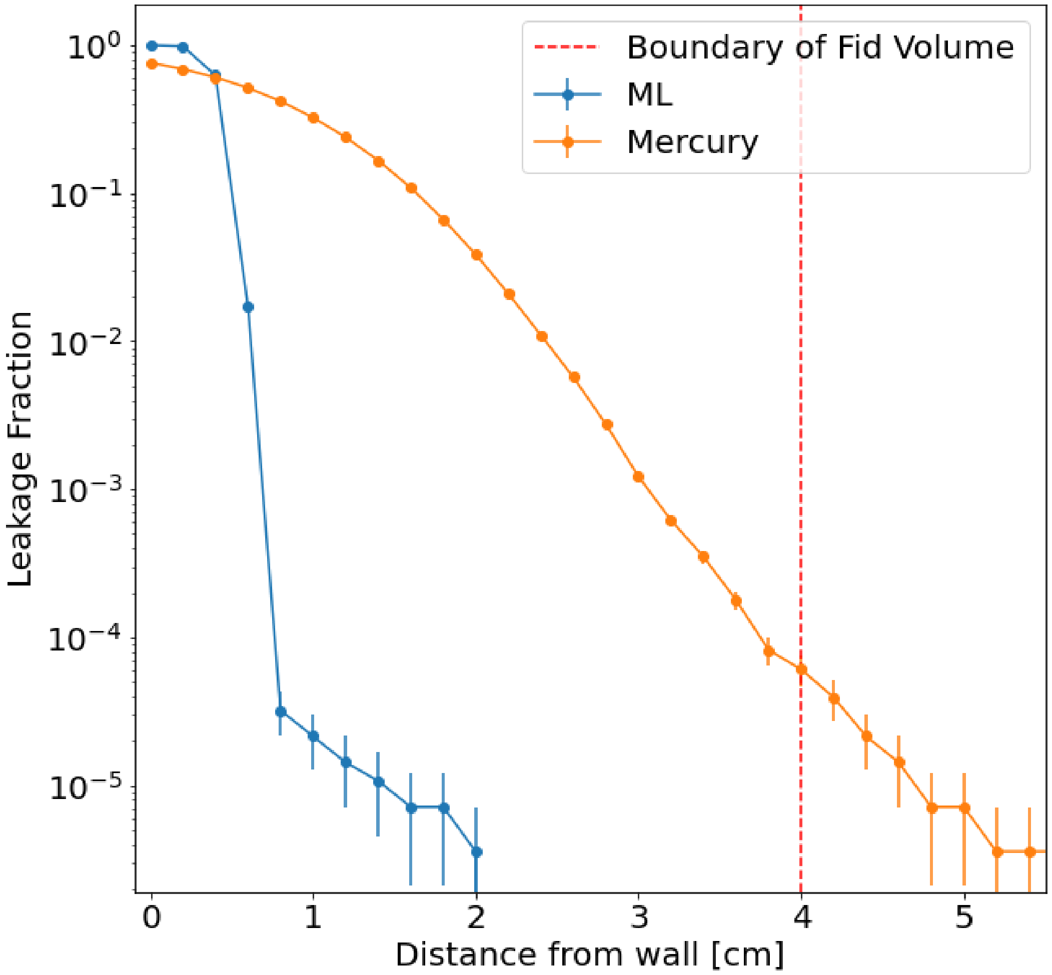


Figure 5.8: Plot showing leakages of ML-1D and Mercury on 10^6 events each consisting of 5 thermal electrons. The events were simulated using BACCARAT, and were given an initial location at the top edge of the detector $[x,y,z] = [0,72,145.5]$. Plot shows that the leakage fraction using ML of low energy thermal events is less than 10^{-5} compared to Mercury.

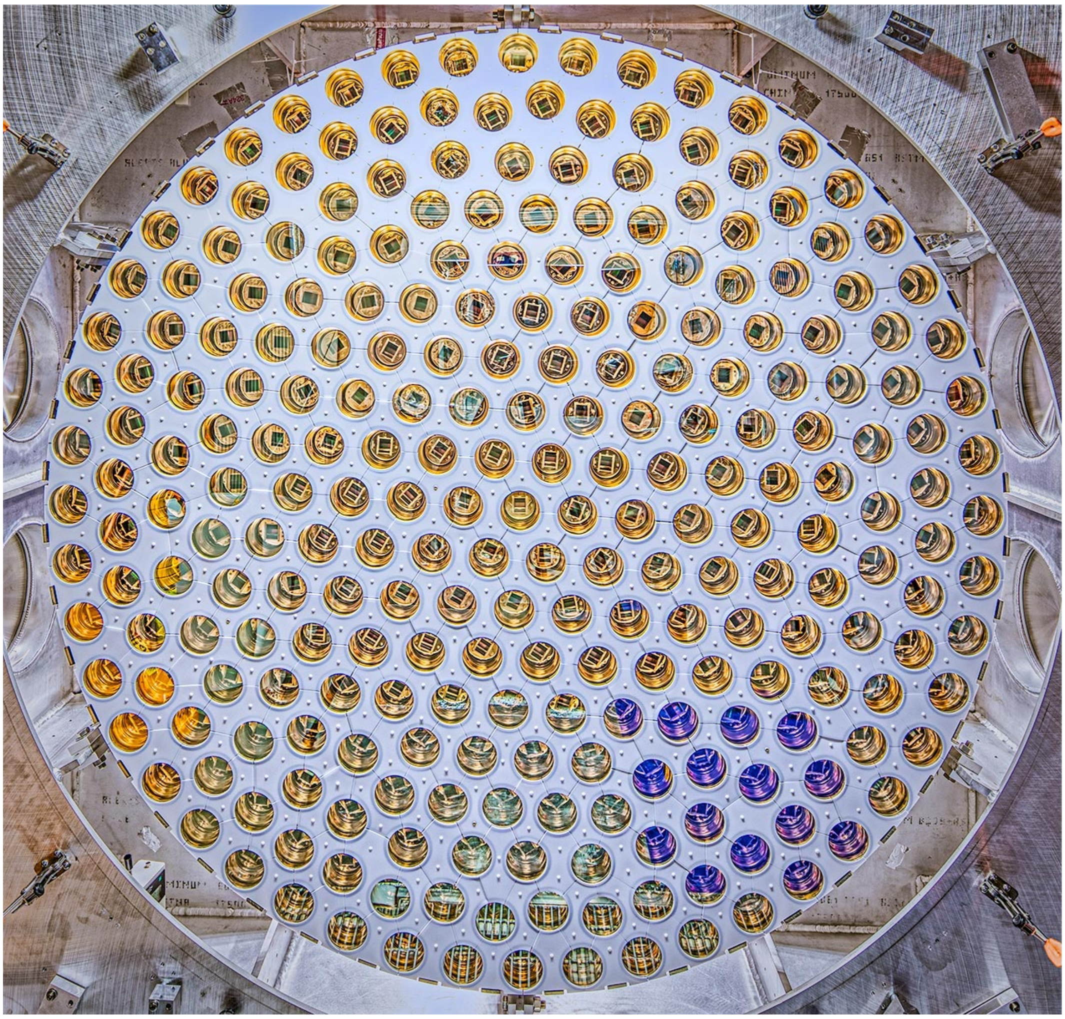


Figure 5.9: Image showing the top array PMT layout in LZ. The Top array employs a circular structure of PMT layout, with the outer ring of PMTs placed further than the TPC walls. This is to improve position reconstruction and light collection for events nearer to the walls of the TPC. In total, there are 253 PMTs in the Top array.

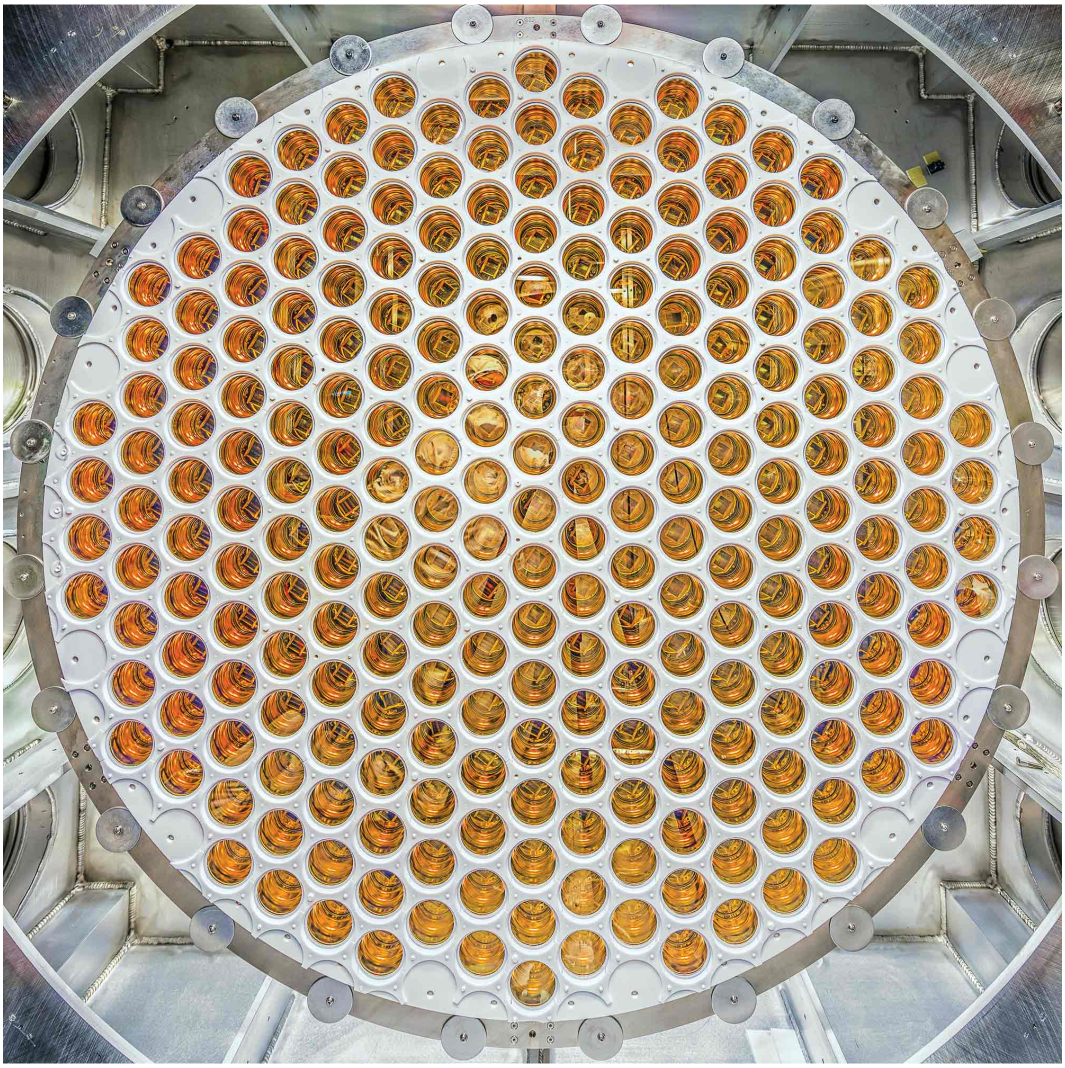


Figure 5.10: Image showing the bottom array PMT layout in LZ. The Bottom array places PMTs in a hexagonal honeycomb-structure to allow for better light collection of the bottom array. In total, there are 241 PMTs in the bottom array.

5.5 2D Convolutional Neural Network

A 2D input to the CNN was also constructed to see whether a greater improvement could be achieved for position reconstruction for wall events, and can be seen in figure 5.11. Since the input to a 2D CNN must be a $(n \times n)$ square matrix, the PMT position representations were adjusted in order to allow for the input to be a (20×20) input, with zero-padding between some of the PMTs channels.

5.5.1 Model Architecture

The 2D CNN was trained on the same dataset, with many model architectures tested. The best model contained 3 sets of convolutional layers with network sizes $[250, 500, 750]$, with pooling layers after every layer. This was then followed by 3 fully connected dense layers of sizes $[1000, 256, 64]$, and an output layer of 2 neurons, each corresponding to the x and y dimension.

It was found that more neurons, and thus more layers were needed in the 2D input compared to the 1D input. This could be due to the increased complexity involved in analysing 2D input data, as the data would have a higher dimensionality compared to the 1D input data. Padding between the PMT channels to create the (20×20) input also added sparsity to the data, which may have contributed to the greater need for complexity.

Max pooling was used instead of Average pooling, even though the latter is predicted to be better for tasks that require generalisations in positional space. This was due to the sparsity introduced within the input data by zero-padding, hence averaging samples within the data would mean a loss in the intensity contributed by each PMT.

Challenges in Training

However, training the 2D CNN was more challenging compared to the 1D input. This was caused mainly by the increase in the number of parameters that were trainable in the 2D CNN. By increasing the number of layers and neurons, the total

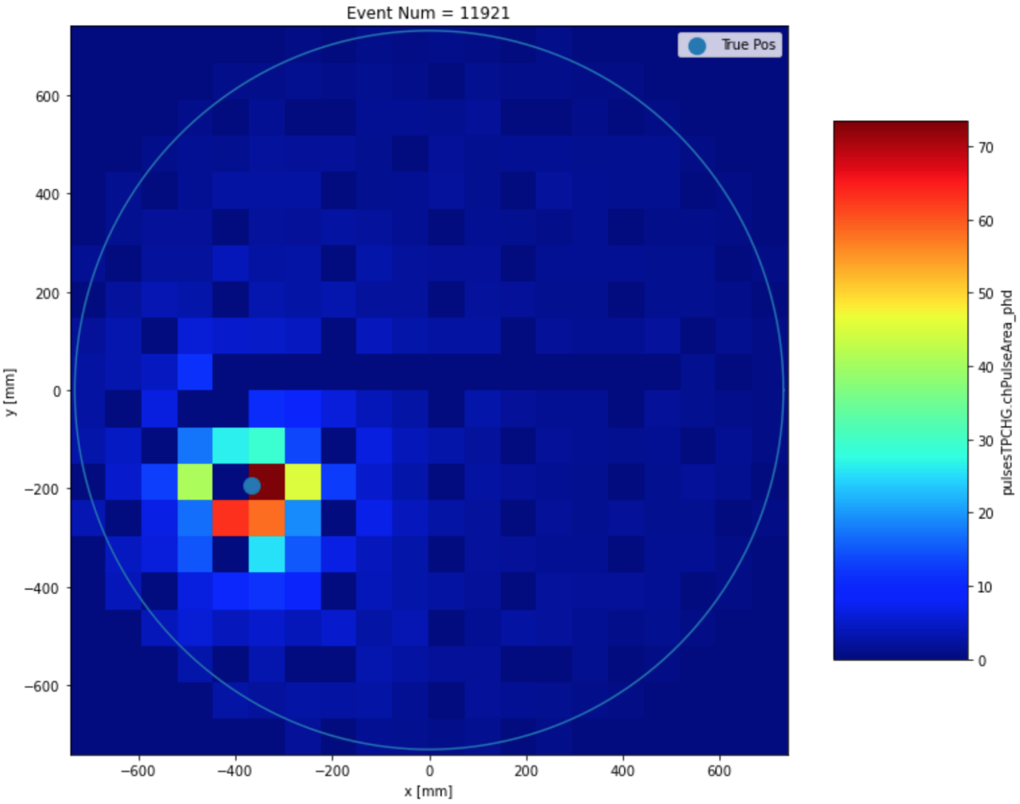


Figure 5.11: Image showing the input data for 2D Convolutional neural network. Since the input of the Neural Network is in square matrices, the PMTs were arranged in 2D space, with zero padding involved to pad between the PMTs. This also reduced the dimensionality of the dataset, as well as allowing for the 2D position to be taken into account by the Neural Network. All pulse area values were summed for each event, with the input data being normalised before being input to the Neural Network.

number of parameters to optimise was over 5 million. This led to an increase in the training time required to minimise the loss function.

5.5.2 Leakage

The leakage of the 2D CNN was also calculated and can be seen in figure 5.13. It shows that the leakage for the 2D CNN falls below 10^{-5} after only 1 cm from the TPC wall. This is much lower compared to the leakage produced by Mercury, which is seen in orange in figure 5.13. This can be used to determine a new wall boundary for the fiducial volume which can be used when using the 2D position reconstruction using Machine Learning.

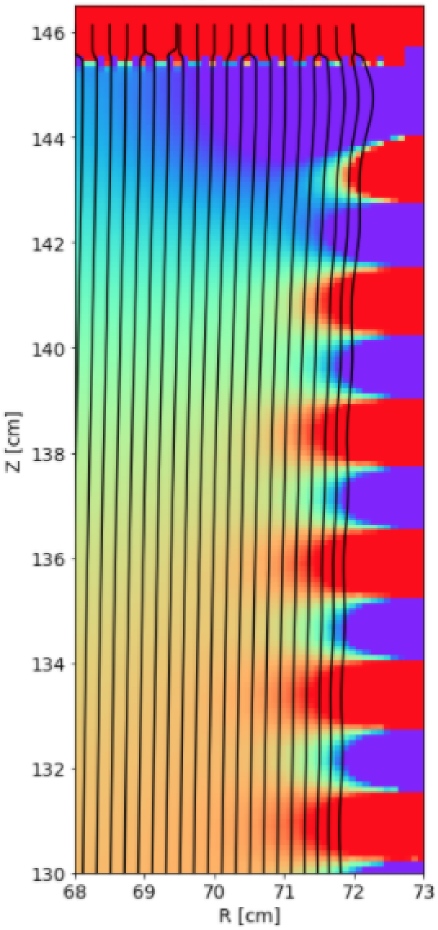


Figure 5.12: Plot shows the non-uniformity of the Electric field at the top of the TPC near the walls. Black lines represent the electric field lines in the absence of diffusion. Taken from [73].

However, due to the non-uniformity of the electric field near the wall, as seen in figure 5.12, a maximum distance of only 2 cm from the TPC wall can be taken. This is as nearer to the edge of the TPC, the electric field is non-uniform [73], hence leading to the movement of ions in a non-uniform way. This non-uniformity was not included in simulations of events using BACCARAT for MDC3, which assumed that the electric field was uniform throughout the TPC.

Therefore, by taking a new fiducial volume which is a distance of 2 cm from the wall, this would give a new fiducial volume mass of 6.2 tonnes of LXe.

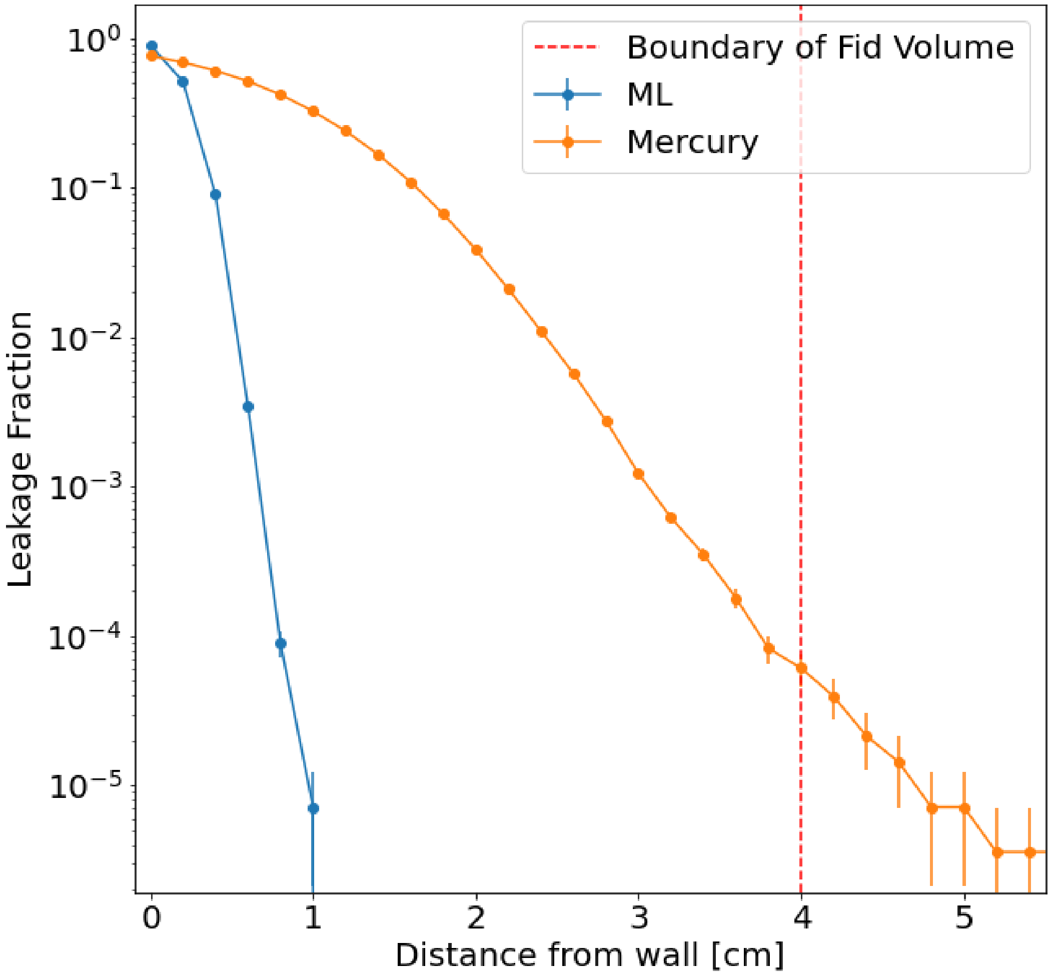


Figure 5.13: Plot showing the leakage of Mercury (orange) and ML (blue) for the 2D input ML model. The leakage shows that mercury only achieves a leakage fraction of 10^{-4} whereas the ML model achieves the same leakage fraction at below 1cm from the TPC wall. This is significantly better compared to Mercury, as well as the 1D ML input seen in figure 5.8. Both models in this figure were tested on 10^{-6} events consisting of 5 thermal electrons originating at $[x,y,z] = [0,72,145.5]$.

Table 5.1: Total ER and NR backgrounds with 99.5% ER discrimination and 50% NR efficiency for both the current fiducial volume of 5.6 tonnes and a proposed increase of 6.2 tonnes of fiducal if using ML 2D position reconstruction for wall events.

Recoil	5.6T (4cm wall)	6.2T (2cm wall)
ER	1195	1351
NR	1.06	1.27
Total	6.51	7.39

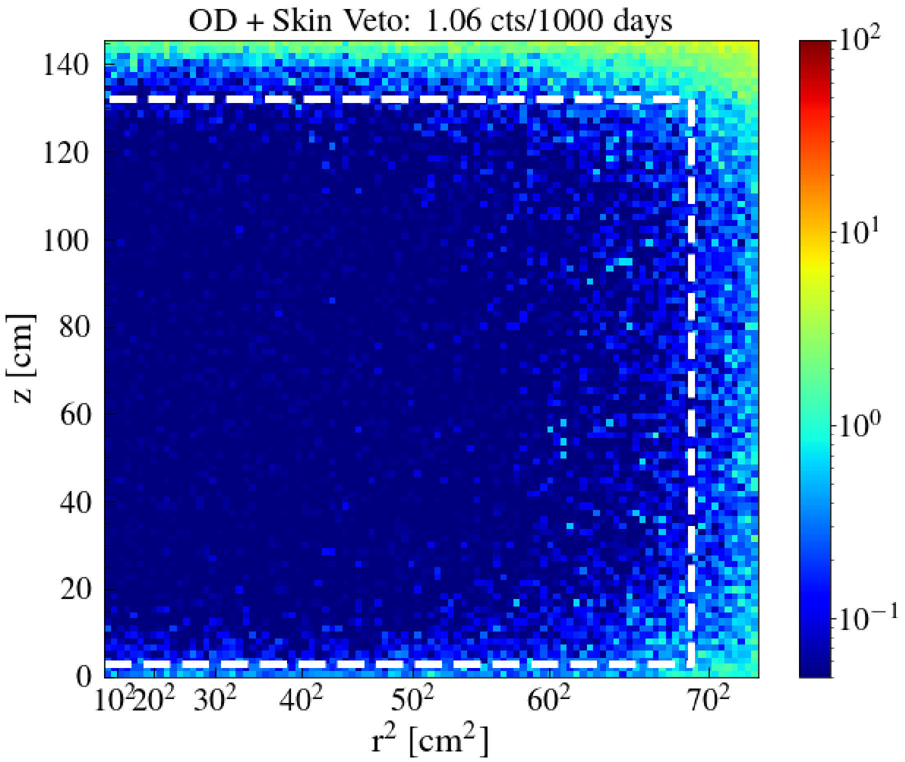
5.6 Background studies

When changing the distance of the fiducial volume wall location, there would be an increase in backgrounds which have to be taken into account. This can be seen by looking at figure 5.14, which shows the projected NR backgrounds. By increasing the size of the wall, there would be a larger component of backgrounds within the fiducial volume, with 1.27 NR events / 1000 days expected for a wall distance of 2 cm, and 1.06 NR events / 1000 days expected for a wall distance of 4 cm.

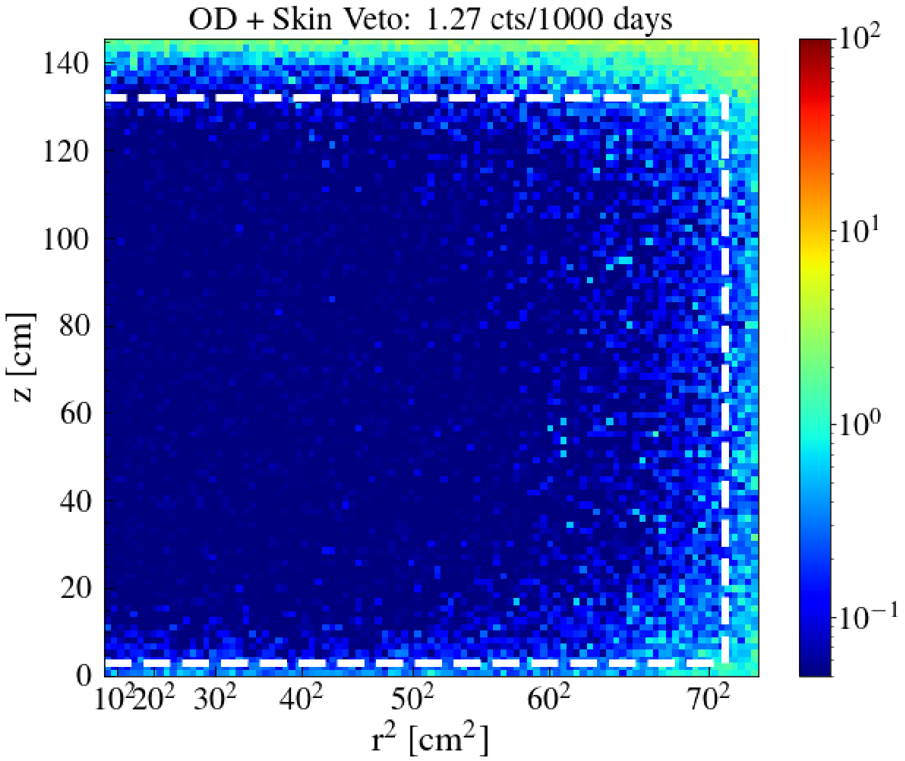
Changing the wall distance also affects the ER backgrounds, which are more prominent compared to the NR. This is due to the fact that having a smaller wall distance (hence larger fiducial volume), more gamma and wall events would be within the fiducial volume. This can be seen in figure 5.15, where the boundary of the fiducial volume for a 2cm wall includes more ER events compared to the boundary of the wall with 4 cm. Hence, the expected counts would be 1351 ER expected / 1000 days for 2 cm wall, compared to 1195 ER expected / 1000 days for 4 cm wall.

Assuming a 99.5% ER/NR discrimination, and an NR efficiency of 50%, the total number of expected events for a 5.6 tonne fiducial volume is 6.51 per 1000 day run. For a 6.2 tonne fiducial volume, 7.39 events are expected per 1000 day run.

By taking the ratio between the expected total ER and NR events between 5.6 tonnes and 6.2 tonnes in fiducial volume, the backgrounds for each component can be taken and scaled. Hence, these can then be used to feed into the PLR method (see next section) in developing background PDFs that can be input into the PLR method to determine the impact of having a larger fiducial volume on the LZ sensitivity.



(a) Plot showing the NR backgrounds for 4cm wall.



(b) Plot showing the NR backgrounds for 2cm wall.

Figure 5.14: Figure comparing the number of backgrounds within the different fiducial volume sizes of 2 cm and 4 cm from the TPC wall. The input data was the MDC3 dataset.

Figure 5.15 can also be used to show the origin of the ER events, with all the ER events located at the edges of the wall (high R^2) value, with no ER events expected at the bottom of the TPC (low Z), and even fewer events expected at the top (high Z). This shows that the origin of the ER background events originates at the walls. This is different for NR background events, as seen in figure 5.14, where the number of NR background events occurs around the edges of the fiducial volume.

5.7 LZ Sensitivity

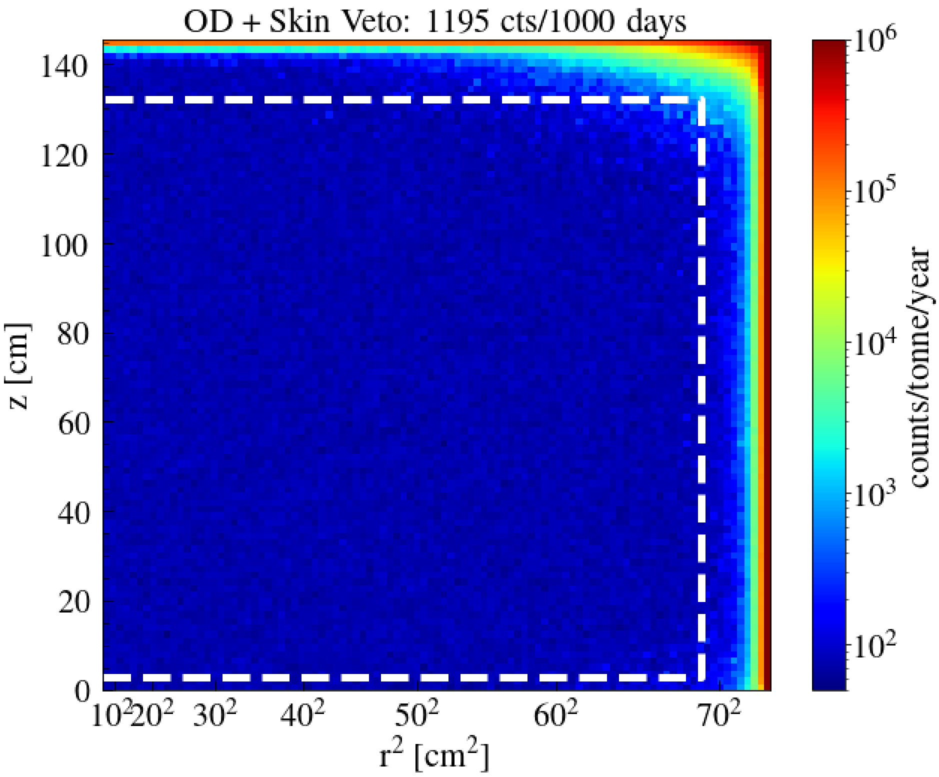
The expected outcome of the LZ experiment is the discovery of dark matter particles by searching for an excess of events that is higher compared to the number of events expected from the background model; or to verify that the background model implemented is accurate and place a statistical limit to exclude model parameters.

A null-hypothesis H_0 and an alternative hypothesis H_1 are evaluated, with H_0 being assumed to be true unless the observed data seen by the experiment requires the rejection of the null hypothesis H_0 and acceptance of the alternative hypothesis H_1.

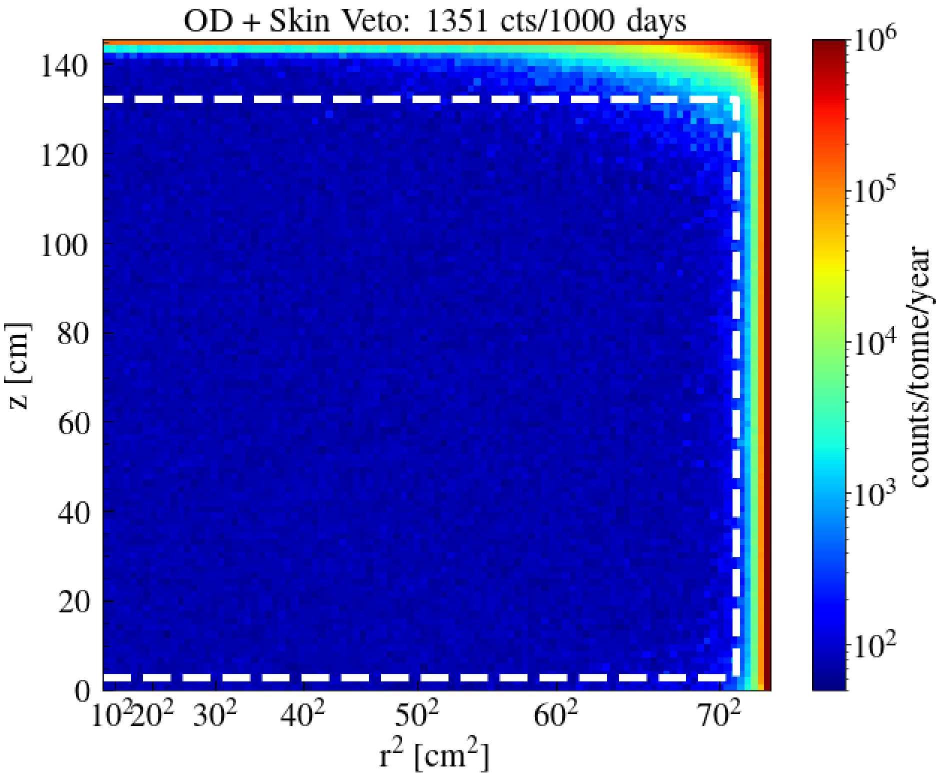
In order to discover dark matter, signals that are in excess of the expected background are examined. The H_0 in this case is then the background-only hypothesis, with H_1 being a background + signal hypothesis that is only accepted if the data suggests a rejection of the initial null hypothesis, as it is incompatible with the observed data.

If the observed data seen by the experiment fails to reject the null hypothesis, then the observed data can still be used to determine which areas of the parameter space is excluded by the observed data. This is the limit-setting scenario and is used to give the sensitivity of the experiment as well as help determine the theoretical parameter space still possible for a discovery.

To determine the discovery potential and LZ Sensitivity to WIMPs, the Profile Like-



(a) Plot showing the ER backgrounds for 4cm wall.



(b) Plot showing the ER backgrounds for 2cm wall.

Figure 5.15: Figure comparing the number of backgrounds within the different fiducial volume sizes of 2cm and 4cm from the TPC wall. The input data was the MDC3 dataset.

lihood Ratio (PLR) test statistic is used [144]. This uses an unbinned maximum likelihood and allows comparisons to be made on an event by event level of the observed data to a given model. The PLR method was implemented using the RooStats package, which is used commonly within High Energy Physics communities. The LZ experiment built upon the RooStats package to develop the LZStats package. The following section will highlight briefly the foundations of the PLR test statistic, and how it can be used to obtain the sensitivity of LZ to 5.6 and 6.2 tonnes of fiducial mass LXe.

5.7.1 The PLR Method

The LZ experiment will record large amounts of data, almost $\sim$ 3Pb during a 1000-day science run to help in the discovery of WIMPs. Hence, cuts are needed to restrict the WIMP search, with the four cuts used for the WIMP search detailed as follows:

- Region of Interest (ROI) cut - this is to constrain the energy window for a 40 GeV/c^2 WIMP. This is equivalent to 1.5-6 keV for ER events, and 6-30 keV for NRs.
- Single Scatter cut - where an event must only scatter once, with the energy deposition taking place at a specific vertex in xyz coordinate space within the TPC
- Veto cut - with both skin and outer-detector vetos being applied.
- Fiducial Volume - events must take place within the fiducial volume of the detector.

The region of interest (ROI) can be used to place a limit on the size of S1 and S2 signals, hence only S1s with size $0 < S1_c 80$ detected photons (phd) are accepted; and for S2s, a signal size of greater than 420 phd is required.

Each event that passes a cut is then parameterised by a vector of observed quantities $\mathbf{x} = [S1_c, S2_c, x, y, z, t]$, however, the spatial quantities are uniform due to the fiducialisation of the TPC, with the events also assumed to be time-independent, hence the variables used in the PLR are only the S1 and S2 signal sizes, therefore $\mathbf{x} = [S1_c, S2_c]$, with the subscript showing corrected quantities.

Different hypotheses can be represented, each of which representing different models under background and signal probabilities, using probability density functions (PDFs). These would represent a particular event for a specific signal or background component, and can be expressed as $f_c(x|\theta)$, where the subscript c represents different components of the background or signal models. By summing each of these components, the total probability model for N components can be expressed as

$$f(x|\theta) = \sum_{c=1}^{N} (\frac{\mu_c(\theta)}{\mu(\theta)}) f_c(x|\theta) \tag{5.1}$$

where $\mu_c(\theta)$ represents the expected number of events, and $\mu(\theta)$ is the total number of events summed over the components.

Given a dataset $\mathcal{D} = x_{i}{}_{i=1}^{n}$ with n events, where each event is drawn independently from the same underlying distribution, the joint probability distribution can be calculated, as it would be equal to the product of each individual probability distribution of each event. This is then equal to

$$\begin{aligned} f(\mathcal{D}|\theta) &= Pois(n|\mu(\theta)) \prod_{i=1}^{n} f(x_i|\theta) && (5.2) \\ &= \left[\frac{\mu(\theta)^n}{n!} e^{-\mu(\theta)} \right] \prod_{i=1}^{n} f(x_i|\theta) && (5.3) \end{aligned}$$

The PLR test-statistic

The likelihood function $\mathcal{L}(\theta)$ is the same as equation 5.2, and can be used to determine what combination of model parameters are needed to maximise the probability of obtaining the observed dataset $\mathcal{D}_{\text{obs}}$. Hence, $\mathcal{L}\theta$ allows a way to determine the level of agreement between a given hypothesis and the observed data obtained. This is done by creating a test-statistic $t(\mathcal{D})$, which is a scalar function used to discriminate hypotheses. For each hypothesis H_0 and H_1, a test-statistic distribution $\mathrm{f(t|H_0)}$ and $\mathrm{f(t|H_1)}$ is calculated.

The profile likelihood ratio is a widely used test statistic of the form

$$\lambda(\mu) = \frac{\mathcal{L}(\mu,\hat{\hat{\nu}})}{\mathcal{L}(\hat{\mu},\hat{\nu})} \tag{5.4}$$

where μ is the given parameter of interest, and ν represents the collection of nuisance parameters [144] [143]. These nuisance parameters are added into the final form of the likelihood function, and are approximated as Gaussian constraints, as they contain some uncertainty on the estimate.

The double hat $\hat{\hat{\nu}}$ is the conditional maximum-likelihood estimator, and represents the value of ν that maximises the log-likelihood for a fixed μ. The single hat $\hat{\nu}$ in the denominator is the maximum likelihood estimator (MLE) for μ and is the function that maximises the likelihood function.

Hence, when comparing different hypotheses, a test statistic of the form

$$t_\mu = -2log(\lambda(\mu)) \tag{5.5}$$

which can take values between $[0,\infty]$, with large values of t_μ indicating incompatibility between the hypotheses.

By taking the p-value, which is defined as the probability of obtaining the observed dataset under the assumption that the null hypothesis is true, a hypothesis test can be defined to determine the statistical relevance of accepting the null hypothesis.

$$
\begin{aligned}
p &= P(t > t_{obs}|H_0) && (5.6)\\
&= \int_{t_{obs}}^{\infty} f(t|H_0)dt && (5.7)
\end{aligned}
$$

By taking a pre-determined size of a critical region α, the null hypothesis would be rejected if the observed test statistic $\mathrm{t_{obs}}$ is within this region, hence if $\mathrm{p} < \alpha$.

A confidence level (CL) can be defined as $\mathrm{CL} = (1-\alpha)$, and represents the confidence level of the test, with typical values of α either 5% or 10%. In the remaining of this thesis, and generally in direct detection dark matter experiments, a confidence level of 90% is taken (hence $\alpha = 10\%$).

5.7.2 Impact of larger fiducial volume on LZ Sensitivity

By taking a larger fiducial volume with a 2 cm wall distance, thus corresponding to a fiducial mass of 6.2 tonnes LXe, the sensitivity can be calculated using the Profile Likelihood Ratio (PLR) method, and is shown in figure 5.16. The plot shows that by increasing the fiducial volume by 600 kg of LXe for a 1000 day run, there is an increase in the sensitivity to WIMPs, with the largest change being seen between 18 to 45 GeV/c^2 WIMP masses. For a WIMP mass of 40 GeV/c^2, a 6.2 tonne fiducial mass gave a sensitivity of 1.41×10^{-48} cm^2, compared to 1.76×10^{-48} cm^2 for 5.6 tonnes LXe.

This improvement in the sensitivity by utilising a position reconstruction method based upon machine learning shows the potential that Artificial Intelligence (AI) can have on dark matter physics experiments. By changing only the method of position reconstruction, an increase in the fiducial volume of over 11% can be found, which thus leads to an increase in the sensitivity by $\sim 20\%$ for a 40 $\mathrm{GeV/c^2}$ WIMP mass.

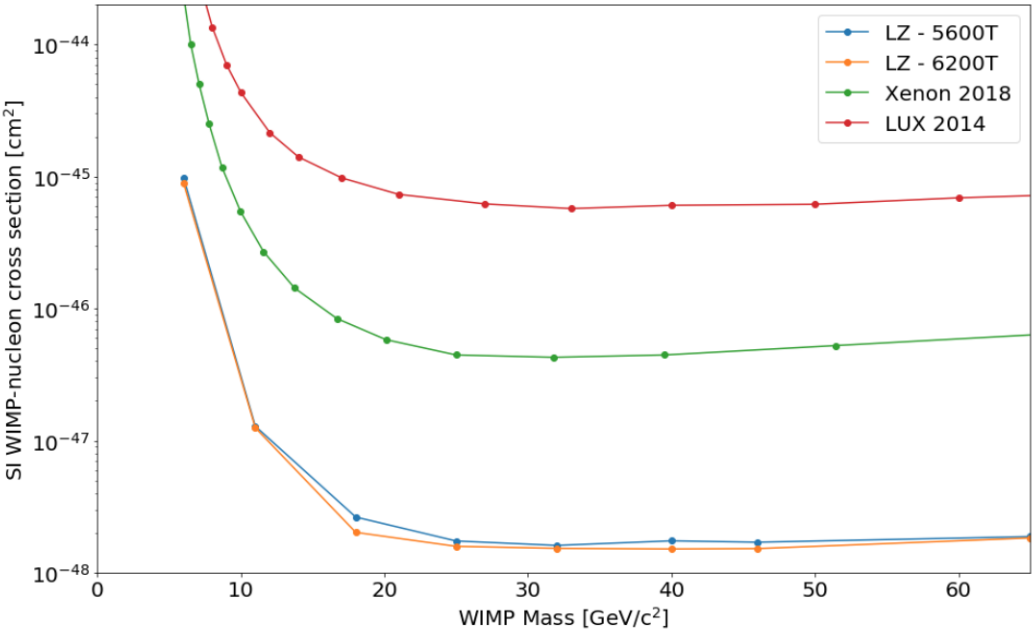


Figure 5.16: Plot showing the projected sensitivity for LZ - 6.2 tonnes, LZ - 5.6 tonnes, Xenon 1T and LUX 2014 to SI WIMP-nucleon elastic scattering for 1000 days.

5.8 Discussion

Changing the size of the fiducial volume, and how that impacts the LZ sensitivity to WIMPs was shown in this section. Having an accurate position reconstruction process, specifically for wall events, is essential to reduce wall backgrounds and improve on the LZ sensitivity to WIMP masses. By increasing the wall distance of the fiducial volume from 4 cm to 2 cm using the 2D ML position reconstruction, a larger fiducial mass of 600 kg is added, with a leakage of lower than 10^{-6} being achieved.

Although having a larger fiducial volume will increase the expected ER and NR background rate of events within the TPC volume by 0.88 counts / 1000 year, these would be taken into account by using the Outer Detector and Skin veto systems utilised by LZ. An ER/NR discrimination of greater than 99.5% would also be utilised in helping to improve the increase in events caused by having a larger fiducial volume.

5.9 Conclusion

This study looked at using a novel method for position reconstruction for wall events using Machine Learning, which were trained on simulated wall events using both 1D and 2D input. Accurate position reconstruction is vital in ensuring that events taking place due to radon plate-out (see the previous chapter), as well as ER events caused by gamma-ray backgrounds, do not enter the fiducial volume.

The study used representative values in the Mercury/LZap case for comparative studies with ML techniques using CNNs. Although these values may differ slightly from the actual values used in the actual LZ experiment, this study highlighted the potential gains in position reconstruction that might be made by deploying such ML techniques in such experiments.

Although both the 1D and 2D neural networks performed well in reconstructing the position, the 2D neural network was able to give a better leakage ratio compared to the 1D network. This led to an increase of the wall distance from 4cm to 2cm closer to the TPC edge, and an increase in fiducial volume change of 600 kg. Increasing the fiducial mass by 600 kg would be equivalent to running the current LZ experiment at 5.6 tonnes an extra $\sim$ 110 days. This shows the impact that increasing the size of the fiducial volume wall from 4cm to 2cm would give. Taking this increase in backgrounds and fiducial volume into account, there would still be an improvement in LZ sensitivity to 40 GeV$/c^2$ WIMPs of 1.41×10^{-48} cm^2.

This study only investigated using neural networks for position reconstruction near the edges of the TPC. This was carried out by the large number of events near the walls in the training set. In order to achieve high accuracy position reconstruction for events at all radii, events with known positions would have to be simulated for all radii. These large number of events would also lead to difficulty in training and optimisation, with a large range of radii leading to difficulties in the network being able to generalise.

Chapter 6

Event Classification

And He is the One Who created the day and the night, the sun and the moon — each travelling in an orbit.

Quran [21:33]

Being able to classify events accurately is vital to the LZ experiment to help discern between backgrounds and signal events. When classifying events, the main aim is to be able to distinguish between single scatter (SS) and multiple scatter (MS) events from each other. Since WIMPs are only expected to produce single scatter events, whereas other backgrounds such as highly energetic gamma-ray events can produce multiple scatters, developing a framework that can distinguish between the two is key in being able to look at the desired physics.

This chapter will look at the current method of event classification employed by LZ, as well as presenting an ML proof of concept framework to show how analysing the time series waveforms can also give strong results for event classification.

6.1 Overview

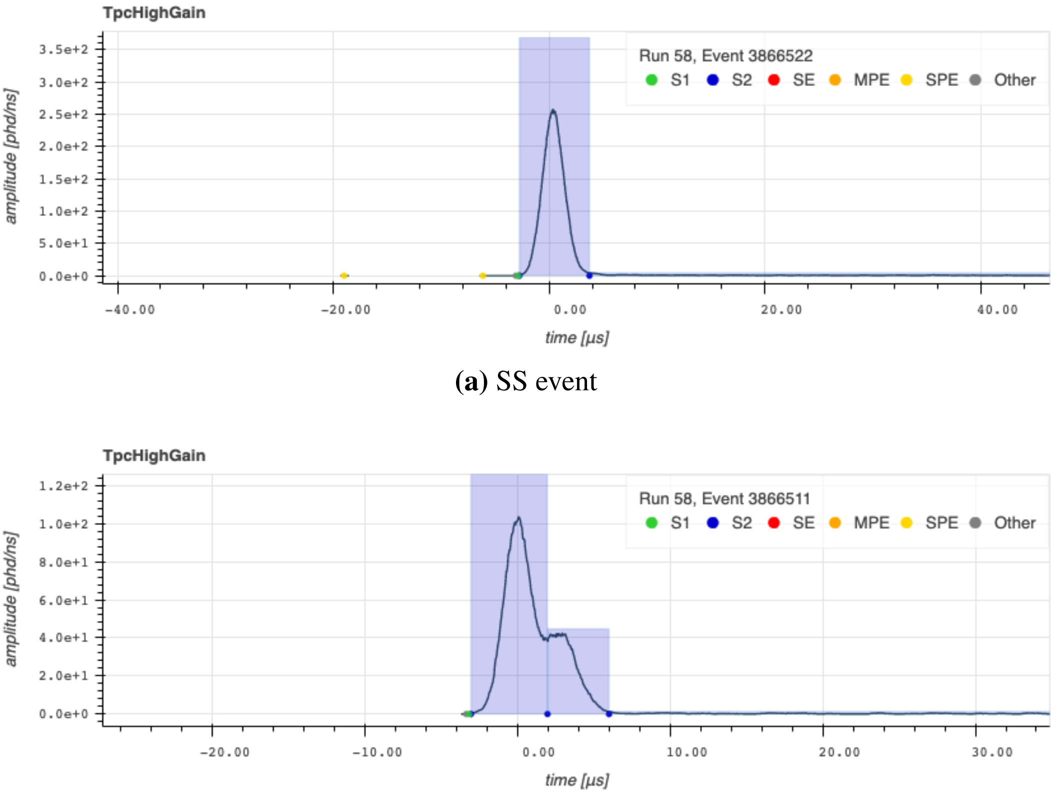
The LZ detector will record over $\sim 10^{10}$ events during its first science run, with a mixture of background and signal events mixed together. With all this data, it is important to be able to classify each event accurately. Currently, LZ looks at the position of events within the fiducial volume, and whether an event position takes place within 2 mm of each other within xyz space, and with the S2s being within 10 *ns* within each other.

An example of the difference between a single and a multiple scatter is seen in figure 6.1, where a single scatter will produce a single S2 pulse whereas a multiple scatter event would produce a merged S2 pulse superimposed onto each other. These merged S2 pulses can be incorrectly classified as being one single S2. Hence, when taking the reduced quantities of the pulses - which is done to ease the data analysis, the incorrect pulse areas are taken, thus giving an error when analysing our WIMP region of interest. To try and improve on this, a time series random forest classifier was developed to analyse the raw waveforms summed over each PMT to try and classify single scatter from multiple scatter events, and will be discussed in the next section.

6.2 Time Series Random Forest Classifier

The `sktime` implementation of the Time series forest classifier was used [162] when classifying the different events to classify between single, multiple and other scatter events. It works by randomly selecting shapelets [163] - a section of data within the time domain, and feeds them into a forest of decision trees. These decision trees are then modelled with the random forest approach [107], with temporal features calculated over the shapelets. By feeding the shapelets into an ensemble of decision trees, we can more generality to our classifier in being able to find the temporal features which distinguish different class labels within our data. These can then be used to make decision tree boundaries within our ensemble, thus allowing for classification to take place. More information on the time series forest classifier can be found in [163].

To try and improve on the classification of single and multiple scatter events, a time series random forest classifier was implemented. This is a novel type of ma-



(a) SS event

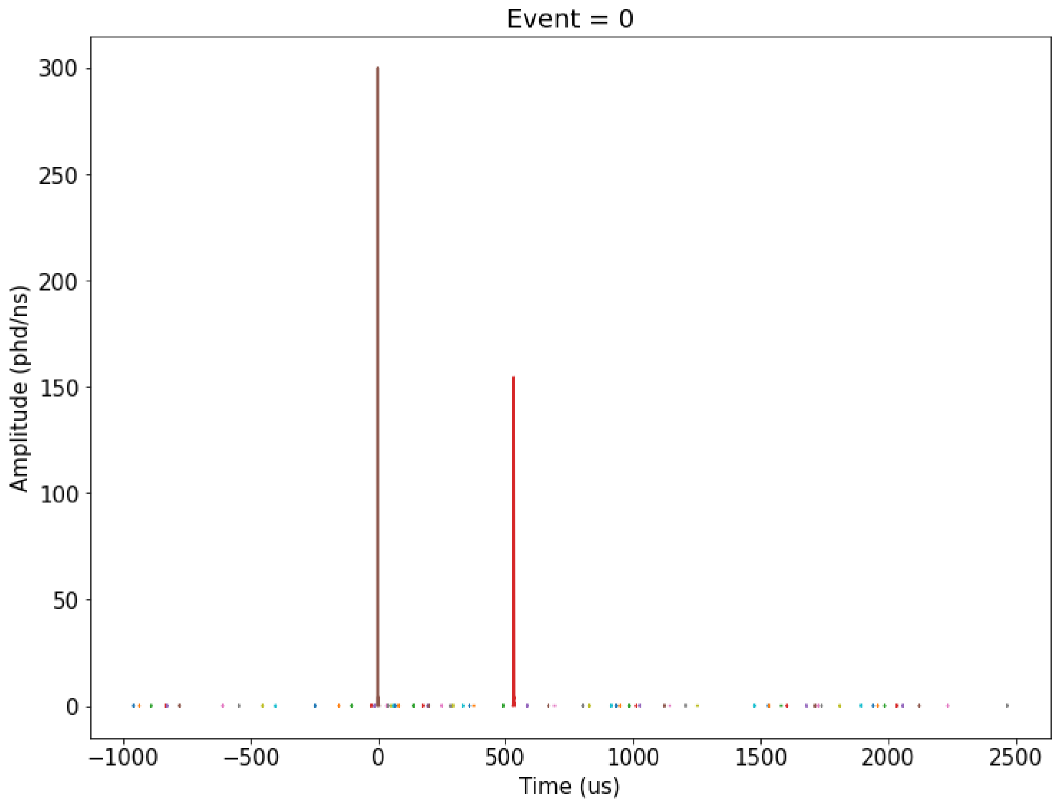
(b) MS event

Figure 6.1: Plots comparing both single and multiple scatter events.

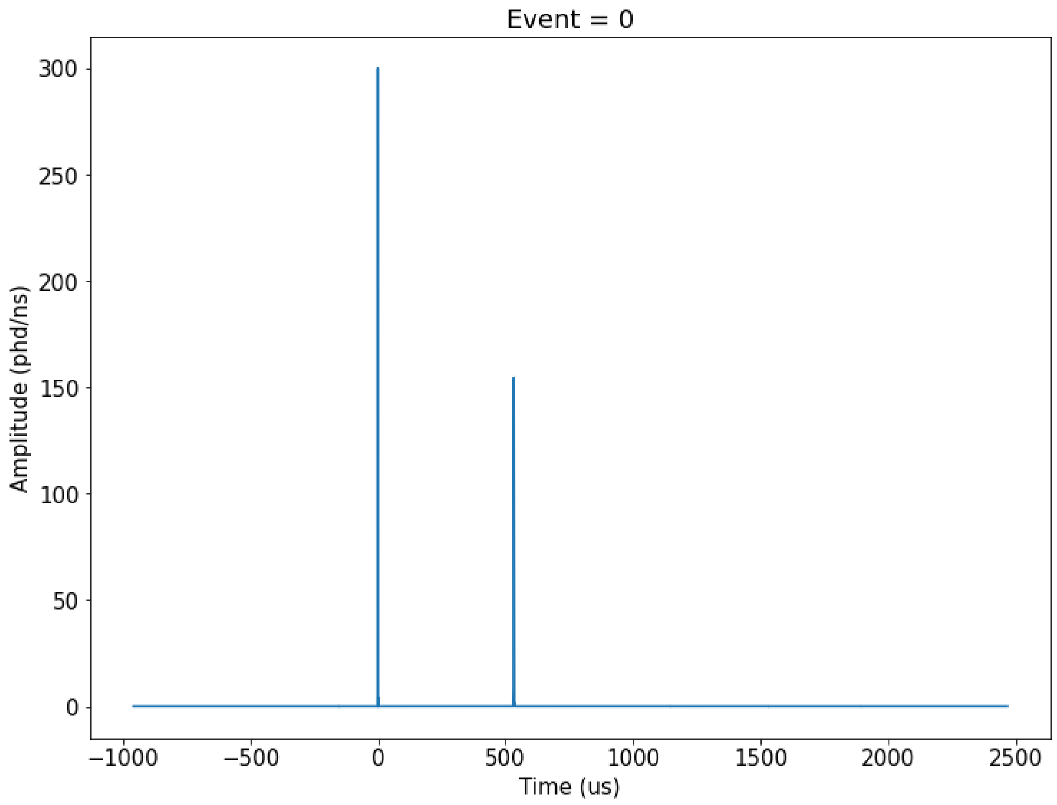
chine learning classifier which analyses the time-series waveforms produced by the PMTs, and classifies them on an event level.

The time-series waveforms were analysed instead of the reduced quantities, as the reduced quantities are created by running LZap onto the waveforms. Hence, it was necessary to try and use training data which consisted of unprocessed data.

However, by using the waveforms instead of the reduced quantities, the size of the amount of data being analysed was considerably larger, as the event waveforms would take up a larger amount of data. Since LZ uses pulse only digitisation, whereby only signals above a certain base-threshold is recorded on the event level, the pulses within an event had to be zero-padded to ensure the timing between pulses was maintained. This can be seen in figure 6.2, where the ‘deadtime’ between pulses was zero-padded.



(a) SS event



(b) MS event

Figure 6.2: Plots comparing both single and multiple scatter events. In the MS event, since the even took place within a short time frame, the S2s are overlapped. This would make it seem as if it is one S2 pulse, instead of two distinct scatters close together.

6.2.1 Training dataset

To train the time series random forest classifier, accurate labels for events were needed. This was achieved by hand scanning ~ 600 events, and labelling the events as either Single Scatter, Multiple Scatter, or Other. The test dataset contained ~ 35, events for each class label.

The hand scanning was undertaken with over 15 people taking part, with each event being checked by two people to correctly classify the events.

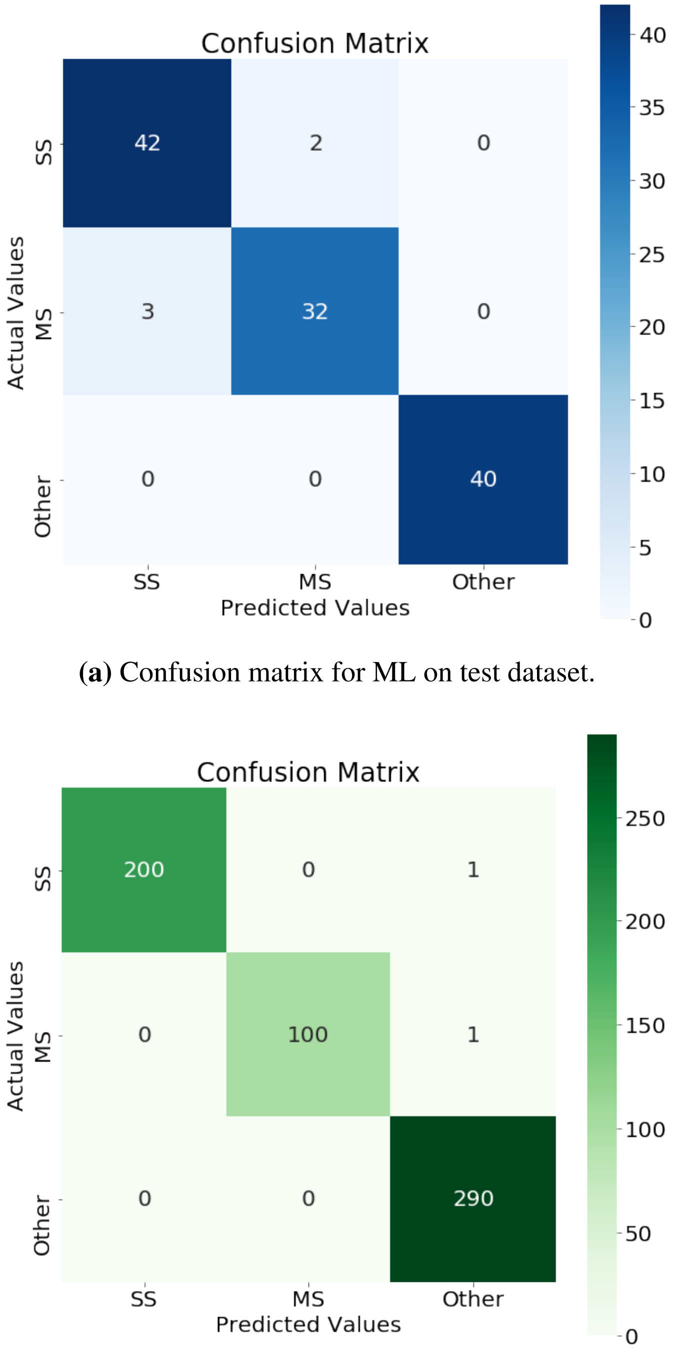
6.3 Discussion

Figure 6.3 shows the confusion matrix for the time series classifier applied to the test dataset. It shows what the classifier output was for each class label, and indicates that the classifier was able to predict the class labels correctly 96% of the time. Although this is lower compared to LZap efficiency of 98%, this accuracy would be improved with more data being available in the training set.

The LZap confusion matrix was calculated with the training dataset, as it would have provided a more coherent picture of the accuracy of LZap in determining between single scatter and multiple scatter events. This was not possible when testing the accuracy of the time series random forest classifier, as the training dataset was already ‘seen’, hence to get a meaningful value of the accuracy, the smaller test dataset was used.

6.4 Conclusion

An innovative method to distinguish between single scatter, multiple scatter and other events was developed using a time series random forest classifier. This classifier takes in the raw time-series input of the PMTs, and classifies them by looking for the temporal differences in pulse shapes, which result from the different types of scatters. This is in contrast with the implementation currently used in LZ, which is dependent on analysing the reduced quantities (RQs) of each pulse within an event,



(a) Confusion matrix for ML on test dataset.

(b) Confusion matrix for the current LZap implementation on the training dataset. The larger dataset was used for LZap as it is already pre-trained, hence the training dataset was already unseen; thus a larger dataset would give a more accurate picture of the accuracy.

Figure 6.3: Image showing the confusion matrices for the Time Series Random Forest classifier (blue) and LZap (green).

as well as looking at the time and xyz position of each event. Although the accuracy of the ML algorithm was lower compared to the current implementation used by LZ, it was shown that even with a training dataset of ~ 600 events, an accuracy of 96% was achieved. This accuracy is likely to improve significantly given a larger training dataset.

Different time-series machine learning methods [164] could also be used to improve on the accuracy given, such as kernel-based methods. These methods employ dynamic time warping (DTW) [165] to transform the time-series dataset into a higher dimensional space. This would then allow for differences in the temporal features specific to different events to be classified more easily.

6.5 Further Work

There are many avenues that future work may lead to within this area of using time-series machine learning methods for event classification. The most important feature when using this method is the ability to classify different types of events based solely on the waveforms produced by the PMTs. This could be classifying specific background events caused by ^{222}Rn or $^{83\mathrm{m}}$Kr, something only currently possible by placing linear cuts on the reduced quantities produced by each pulse within the event.

Chapter 7

Conclusion and Outlook

What you (O humanity) have been given of knowledge is but little.

Quran [17:85]

The work presented in this thesis focused on using Machine Learning techniques for the LZ experiment. Chapter 4 focused on using a Random Forest classifier to determine background BiPo events produced by the decay of ^{222}Rn. This background is important, as it produces events within our region of interest for WIMP searches due to its daughters ^{214}Pb undergoing naked-beta decay. It was found that by using a Random Forest classifier, BiPo events could be classified which were outside the region classically associated with them. This was done by inputting more features to the classifier than what is classically analysed when searching for BiPos.

Chapter 5 then presented work on implementing a convolutional neural network (CNN) for position reconstruction specifically for wall events. By using CNNs, the fiducial volume of a test experiment resembling LZ could be increased from 5.6 tonnes to 6.2 tonnes. This increase in fiducial volume size would lead to an improvement in LZ sensitivity for 40 GeV/c^2 WIMPs of 1.41×10^{-48} cm^2, compared with 1.76×10^{-48} cm^2 for 5.6 tonnes LXe at 90% CL. The increase in fiducial volume by 600 kg is equivalent to running LZ in its current discovery mode with a fiducial volume mass of 5.6 tonnes by over 110 days.

Finally, a brief study on the use of machine learning techniques on the time-series waveforms produced by the PMTs to classify different types of events was introduced, with an accuracy of 96% being found. Although the accuracy was slightly lower compared to the current method employed by LZ, this was likely caused by having a smaller training dataset of only ~ 600 events. Given that the main data domain of the LZ experiment is time-series data, this method provides an excellent way to analyse and classify different events.

The construction phase of LZ has now been complete, with the data taking phase of Science Run 1 (SR1) imminent. Given that the use of Machine Learning and Artificial Intelligence within the Physics community will increase in the years to come, the implementation of such technologies is set to revolutionise the insights produced by fundamental physics experiments in the coming years. This places the LZ experiment and similar LXe TPCs in an exceptional position to explore more areas within the WIMP parameter space, and may even allow it to finally answer the question of the nature of dark matter.

Appendix A

BiPo code

Scripts used for BiPo analysis in chapter 4.

```
import uproot
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from mpl_toolkits.mplot3d import Axes3D
import scipy.fftpack
import os
from IPython.display import clear_output

#####
#   These functions below are to do with the creating dataframes by taking in
#   the lzap RQ files .
#
#
#####

def initialiser(PATH):
    """
    Will return FILE and KEYS for an lzap file in PATH.
    PATH must be string
    """
    file = uproot.open(PATH)
    keys = file.keys()

    sub = b'Events'
    EventKeys = [s for s in keys if sub in s]

    FILE = file[EventKeys[0]]
    TPCHG = FILE['pulsesTPCHG']

    return FILE, TPCHG

def classifier_pulses(KEY, debug = False):
    """
    Will take in the KEY (i.e. TPCHG) and return classification in the form of a list
    """

    classes = KEY['pulsesTPCHG.classification'].array()
    classifications = []
    n = 0

    while n < classes.shape[0]:
        pulse = classes[n]
        decoded = [x.decode('utf-8') for x in pulse]
        classifications.append(decoded)
        n = n + 1

    pulse_classifications = []
    for alist in classifications:
        for blist in alist:
            pulse_classifications.append(blist)

    return pulse_classifications


def sequence_maker(KEY):
    """
```

```python
    Will make a sequence of what pulses have arrived.
    1 - S1
    2 - S2
    3 - SPE (Single PhotoElectron)
    4 - MPE (Multiple PhotoElectron)
    5 - SE (Single Electron)
    Hence will return: [1,3,3,3,3,2,4,4,4,4] etc
    """

    S1_prob = KEY['pulsesTPCHG.s1Probability'].array()
    S2_prob = KEY['pulsesTPCHG.s2Probability'].array() *2
    SPE_prob = KEY['pulsesTPCHG.singlePEprobability'].array() *3
    MPE_prob = KEY['pulsesTPCHG.multiplePEprobability'].array() *4
    SE_prob = KEY['pulsesTPCHG.singleElectronProbability'].array() *5
    other_s1_prob = KEY['pulsesTPCHG.otherProbability'].array() *6
    other_s2_prob = KEY['pulsesTPCHG.otherS2Probability'].array() *7

    sequence = S1_prob + S2_prob + SPE_prob + MPE_prob + SE_prob + other_s1_prob + other_s2_prob
    S1_S2 = S1_prob + S2_prob + other_s1_prob + other_s2_prob

    sequence = sequence.astype(int)
    S1_S2 = S1_S2.astype(int)


    return sequence, S1_S2

def coincidence_cut(input_df):
    """
    Will apply the coincidence (remove all with PMT < 3)

    """

    input_df = input_df[input_df.coincidence > 2]  # Coincidence cut > 2

    return input_df

def applying_cuts(input_df):
    ’’’
    Will apply the cuts to a dataframe
    ’’’

    output_df = coincidence_cut(input_df)
    return output_df




def eventID_maker(KEY):
    """
    Will make the EventID list using an RQ (preferably pulseArea)
    """
    # This cell will make the EventIDs so that it is easier to know which pulse belongs to which Event
    RQ = KEY[’pulsesTPCHG.s1Probability’].array()
    eventID = []
    n = 0
    for event in RQ:
        num = len(event)
        evt = np.ones(num) * n
        eventID.append(evt)
        n = n + 1
    eventID = np.array(eventID)
    eventID = np.concatenate(eventID).ravel()

    return eventID

def eventID_maker2(ROOT, KEY):
    """
    Will make the EventID AND EventID per raw file.
    Args:
    ROOT; FILE
    KEY; TPCHG
    """
    RQ = KEY['pulsesTPCHG.s1Probability'].array()
    rawfile = ROOT['eventHeader.rawFileName'].array()

    values, counts = np.unique(rawfile, return_counts=True)

    eventID = []
    raw_eventID = [np.arange(x) for x in counts]
    raw_eventID = np.concatenate(raw_eventID).ravel()
    raw_evtID = []


    n = 0
    idx = 0
    raw_idx = 0

    for event in RQ:
        num = len(event)
        evt = np.ones(num) * n
        eventID.append(evt)

        raw_evt = np.ones(num) * raw_eventID[n]
        raw_evtID.append(raw_evt)


        n = n + 1
    raw_evtID = np.concatenate(raw_evtID).ravel()


    eventID = np.array(eventID)
    eventID = np.concatenate(eventID).ravel()

    return eventID, raw_evtID


def rawfile_maker(ROOT, KEY):
    """
    Will take in the root file struct (i.e. ROOT = file[’Events;71’] ),
    as well as another KEY i.e. TPCHG (where TPCHG_71 = file[’Events;71’][’pulsesTPCHG.’])
    and return a list of all the filenames
    """

    rawfile = ROOT[’eventHeader.rawFileName’].array()
    RQ = KEY[’pulsesTPCHG.s1Probability’].array()

    names_raw = []

    n = 0

    while n < rawfile.shape[0]:
        filename = rawfile[n]
        for pulse in RQ[n]:
            names_raw.append(filename)

        n = n + 1


    return names_raw




def initial_df(KEY, debug = False):
    """
    Will make the initial dataframe with the main RQs
    """

    pulseArea = KEY['pulsesTPCHG.pulseArea_phd'].array()
    pulseID = KEY['pulsesTPCHG.pulseID'].array()
    #s1Probability = KEY['pulsesTPCHG.s1Probability'].array()
    TBA = KEY['pulsesTPCHG.topBottomAsymmetry'].array()
    xpos = KEY['pulsesTPCHG.s2Xposition_cm'].array()
    ypos = KEY['pulsesTPCHG.s2Yposition_cm'].array()
    promptFaction = KEY['pulsesTPCHG.promptFraction50ns'].array()
    peakAmp = KEY['pulsesTPCHG.peakAmp'].array()
    coincidence = KEY['pulsesTPCHG.coincidence'].array()
    rq_start = KEY['pulsesTPCHG.pulseStartTime_ns'].array()
    rq_end = KEY['pulsesTPCHG.pulseEndTime_ns'].array()
    pulse_type, _ = sequence_maker(KEY) # To make an array which will return the pulse types

    if debug:
        print ("Loaded in the RQs successfully")

    eventID = eventID_maker(KEY)
    classifications = classifier_pulses(KEY)
    if debug:
        print("Events and Classifications loaded")

    df = pd.DataFrame({'eventID_RQ': eventID,
                   'pulseID': pulseID.flatten(),
                   'pulseType': pulse_type.flatten(),
                   'classification': classifications,
                   'pulseArea': pulseArea.flatten(),
                   'TBA': TBA.flatten(),
                   'promptFaction': promptFaction.flatten(),
                   'peakAmp': peakAmp.flatten(),
                   'coincidence': coincidence.flatten(),
                   'start_time':rq_start.flatten(),
                   'end_time':rq_end.flatten()
                  })
```

```python


    return df

def raw_df_maker(ROOT, KEY):
    """
    Function which will make a df of [RawFIleName, EventID, PulseID]
    Args:
    ROOT: i.e. FILE = file['Events;71']
    KEYL i.e. TPCHG = file['Events;71'][TPCHG]
    """

    rawFiles  = rawfile_maker(ROOT, KEY)
    eventID, raw_eventID = eventID_maker2(ROOT, KEY)
    pulseID = KEY['pulsesTPCHG.pulseID'].array().flatten()
    classes = classifier_pulses(KEY)




    df = pd.DataFrame({"rawFileName" : rawFiles,
                       "eventID_Raw" : raw_eventID,
                      "eventID_RQ" : eventID,
                      "pulseID" : pulseID,
                     "pulseType": classes})

    return df





def appender_df(DF, KEY, title, rq_val):
    """
    This function will append a column of rq_val to the initial dataframe

    Input:
    DF: initial dataframe
    KEY: the uproot key e.g TPCHG
    title: type(str); what title you want the new column to be
    rq_val: type(str); the rq_entry in the KEY
    """

    RQ = KEY[rq_val].array().flatten()

    #col = pd.DataFrame({title: RQ})
    DF[title] = RQ


    return DF

def raw_finder(rawdf, evtID, pulse_id, show = False):
    """
    Function which will return the raw_file_fame when giving in as args
    the eventID and pulseID that you want to know
    """
    df = raw_df[(raw_df[’eventID_RQ’]==evtID) & (raw_df[’pulseID’]==pulse_id)]

    if show == True:
        return df.iloc[0][0]
    else:
        return df

def df_file_maker(PATH):
    import os
    """
    Will make a df for all the RQ files found within PATH.
    Must be str with / at the end i.e. data/background/
    """

    #finding the root files
    files = os.listdir(PATH)
    files.sort()
    file_path = [PATH + file for file in files]

    rq_files = []

    final_df = pd.DataFrame([])

    i = 0
    while i < len(file_path):

        file = uproot.open(file_path[i]) #background data
        keys = file.keys()

        #rq_files.append(f)

        sub = b'Events'
        EventKeys = [s for s in keys if sub in s]

        FILE = file[EventKeys[0]]
        TPCHG = FILE['pulsesTPCHG.']

        df = initial_df(TPCHG)
        final_df = final_df.append(df)



        rq_files.extend([files[i]] * df.shape[0])


        print(files[i], " Files remaining = " , len(file_path) - i
    )
```

```python
        i = i + 1
    clear_output(wait=True)

    final_df.insert(0, 'RQ_File_Name', rq_files, True )
    print('Completed. Loaded in ', len(file_path), ' files')

    return final_df

def df_file_appender(PATH, DF, title, rq_val):
    """
    This function will append a column of rq_val to the initial dataframe

    Input:
    PATH: path of the list of lzap RQs that you have
    DF: initial dataframe
    title: type(str); what title you want the new column to be
    rq_val: type(str); the rq_entry in the KEY

    """

    files = os.listdir(PATH)
    files.sort()
    files = [PATH + file for file in files]

    col_df = pd.DataFrame([])

    test_empty = []
    n = 0

    for f in files:
        file = uproot.open(f)
        keys = file.keys()

        sub = b'Events'
        EventKeys = [s for s in keys if sub in s]

        FILE = file[EventKeys[0]]
        TPCHG = FILE['pulsesTPCHG.']

        RQ = TPCHG[rq_val].array().flatten()
        #RQ2 = RQ[:,np.newaxis]

        test_empty.append(RQ)
        n = n + 1
        print (n)
        clear_output(wait=True)

    test_empty = np.concatenate(test_empty).ravel()

    DF[title] = test_empty

    #DF.insert(0,title, test_empty, True )
    print('Completed')

    return DF


#####
#
#   These functions are to do with the actual BiPo tagging
#
#####


def time_diff(idx1, idx2, start_array):
    '''
    Function which will take the indices of 2 pulses from the same event, and return the time difference
    (in nanoseconds) between them.

    idx1, idx2 - type: int; index of pulse 1 and pulse 2
    start_array - type: array; array with the start times
    '''

    diff = start_array[idx2] - start_array[idx1]

    return diff

def tagger(dataframe, event):
    """
    This function will take in a dataframe of all the events, and return another dataframe with columns:
    [EventID, pulseID_1, pulseID_2, start_time_1, start_time_2,'start_1', 'start_2', time_diff(ns)]

    The purpose is to find the time difference between two S1 pulses for BiPo tagging for a specific event

    dataframe: dataframe
    event: EventID that you want to look at

    """
    evt_df = dataframe[dataframe.eventID_RQ == event]
    #file_name = evt_df['RQ_File_Name'].to_list()
    evt = evt_df['eventID_RQ'].values
    types = evt_df['pulseType'].values
    tba = evt_df['TBA'].values
    #classes = evt_df['classification'].to_list()
    times = evt_df['start_time'].values
    id_0 = evt_df['pulseID'].values
    areas = evt_df['pulseArea'].values
    prompt_faction = evt_df['promptFaction'].values
    end_times = evt_df['end_time'].values


    n = 0
    test_array = np.empty((0,14), dtype=int)

    while n < types.shape[0]:

        if types[n] == 1.0:

            m = n + 1
            #print(n,m)
            while m < types.shape[0]:

            #if types[n] == 1.0:
                if types[m] == 1.0:
                    diff = time_diff(n,m,times)

                    arr = np.array([evt[n],
                                    id_0[n], id_0[m],
                                    areas[n], areas[m],
                                    times[n], times[m],
                                    tba[n], tba[m],
                                    end_times[n], end_times[m],
                                    prompt_faction[n], prompt_faction[m],
                                    diff])
                    test_array = np.append(test_array, [arr], axis=0 )
                    #print(n, m)
                m = m + 1

        n = n + 1
    arr = pd.DataFrame(test_array, columns=['EventID_RQ',
                                            'pulseID_1','pulseID_2',
                                            'area_1', 'area_2',
                                            'start_1', 'start_2',
                                            'TBA_1','TBA_2',
                                            'end_1','end_2',
                                            'promptFaction_1','promptFaction_2',
                                            'time_diff_(ns)'])

    return arr



def tagger2(dataframe, event, max_diff = 164000):
    """
    This function will take in a dataframe of all the events, and return another dataframe with columns:
    [EventID, pulseID_1, pulseID_2, start_time_1, start_time_2,'start_1', 'start_2', time_diff(ns)]

    The purpose is to find the time difference between two S1 pulses for BiPo tagging for a specific event

    dataframe: dataframe
    event: EventID that you want to look at
    max_diff: the max time_diff between two s1 events

    """
    evt_df = dataframe[dataframe.eventID_RQ == event]
    #file_name = evt_df['RQ_File_Name'].to_list()
    evt = evt_df['eventID_RQ'].values
    types = evt_df['pulseType'].values
    tba = evt_df['TBA'].values
    #classes = evt_df['classification'].to_list()
    times = evt_df['start_time'].values
    id_0 = evt_df['pulseID'].values
    areas = evt_df['pulseArea'].values
    prompt_faction = evt_df['promptFaction'].values
    end_times = evt_df['end_time'].values

    peak_amp = evt_df['peakAmp'].values
    peak_time = evt_df['peak_time'].values
    pulseArea50 = evt_df['pulseArea50'].values
    pulseArea100 = evt_df['pulseArea100'].values
    rmsWidth = evt_df['rmsWidth'].values
    FWHM = evt_df['FWHM'].values


    n = 0
    test_array = np.empty((0,26), dtype=int)

    while n < types.shape[0]:

        if types[n] == 1.0:

            m = n + 1
            #print(n,m)
            while m < types.shape[0]:

            #if types[n] == 1.0:
                if types[m] == 1.0:
                    diff = time_diff(n,m,times)

                    if diff < max_diff:

                        arr = np.array([evt[n],
                                        id_0[n], id_0[m],
                                        areas[n], areas[m],
                                        times[n], times[m],
                                        tba[n], tba[m],
                                        end_times[n], end_times[m
   ],
                                        prompt_faction[n],
   prompt_faction[m],
                                        peak_amp[n], peak_amp[m],
                                        peak_time[n], peak_time[m
   ],
                                        pulseArea50[n],
   pulseArea50[m],
                                        pulseArea100[n],
   pulseArea100[m],
                                        rmsWidth[n], rmsWidth[m],
                                        FWHM[n], FWHM[m],
                                        diff])
                        test_array = np.append(test_array, [arr],
   axis=0 )
                    #print(n, m)
                m = m + 1

        n = n + 1
    arr = pd.DataFrame(test_array, columns=['EventID_RQ',
                                            'pulseID_1','pulseID_2
   ',
                                            'area_1', 'area_2',
                                            'start_1', 'start_2',
                                            'TBA_1','TBA_2',
                                            'end_1','end_2',
                                            'promptFaction_1','
   promptFaction_2',
                                            'peakAmp_1','peakAmp_2
   ',
                                            'peak_time_1','
   peak_time_2',
                                            'pulseArea50_1','
   pulseArea50_2',
                                            'pulseArea100_1','
   pulseArea100_2',
                                            'rmsWidth_1','
   rmsWidth_2',
                                            'FWHM_1','FWHM_2',
                                            'time_diff_(ns)'])

    return arr
```

```




def s1_finder(df):
    """
    Function which will run through the df of all the events, and return another dataframe with all pairings
    of an S1 followed by another S1 within an event, and the time between then
    """

    rq_files = df[’RQ_File_Name’].unique()


    event_IDs = []

    n = 0

    while n < rq_files.shape[0]:
        events = df[df[’RQ_File_Name’] == rq_files[n]][’eventID_RQ’].unique()
        event_IDs.append(events)
        clear_output(wait=True)
        print(rq_files.shape[0] - n)
        n = n + 1



    n = 0
    m = 0

    tagged_df = pd.DataFrame([])
    while n < rq_files.shape[0]:
        rq_df = df[df[’RQ_File_Name’] == rq_files[n]]
        evt_df = pd.DataFrame([])

        event = event_IDs[n]
        while m < event.shape[0]:
            event_df = tagger2(rq_df, event[m])
            evt_df = evt_df.append(event_df)


            clear_output(wait=True)
            print(n,m, event.shape[0]-m)
            m = m + 1

        filename = [rq_files[n] for _ in range(evt_df.shape[0])]
        evt_df.insert(0,'RQ_File_Name', filename)
        tagged_df = tagged_df.append(evt_df)

        m = 0
        n = n + 1

    #tagged_df.to_csv('tagged_df.csv')
    return tagged_df
    print('Finished')
```

Script for training Decision Tree and Random Forest classifiers.

```
import uproot
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import os
from IPython.display import clear_output

#####
# Code for training ML for BiPo
#
#####

# ---------- Decision Tree Classifier
```

```


from sklearn.tree import DecisionTreeClassifier
classifier = DecisionTreeClassifier(criterion='entropy', random_state=0)
classifier.fit(X_train, y_train)

DecisionTreeClassifier(ccp_alpha=0.0, class_weight=None, criterion='entropy',
                       max_depth=None, max_features=None, max_leaf_nodes=None,
                       min_impurity_decrease=0.0, min_impurity_split=None,
                       min_samples_leaf=1, min_samples_split=2,
                       min_weight_fraction_leaf=0.0, presort='deprecated',
                       random_state=0, splitter='best')


y_pred = classifier.predict(X_test)

cm = confusion_matrix(y_test, y_pred)

accuracy_score(y_test, y_pred)

# ---------- Feature Importance
feature = []
score = []
for name, importance in zip(X_train.columns, classifier.feature_importances_):
    feature.append(name)
    score.append(importance)
    print(name, "=", importance)



```

```
# ---------- Random Forest Classifier
from sklearn.ensemble import RandomForestClassifier

clf = RandomForestClassifier(n_jobs=2, random_state=0)


clf.fit(X_train, y_train)

RandomForestClassifier(bootstrap=True, ccp_alpha=0.0, class_weight=None,
                       criterion=’gini’, max_depth=None, max_features=’auto’,
                       max_leaf_nodes=None, max_samples=None,
                       min_impurity_decrease=0.0, min_impurity_split=None,
                       min_samples_leaf=1, min_samples_split=2,
                       min_weight_fraction_leaf=0.0, n_estimators=100, n_jobs=2,
                       oob_score=False, random_state=0, verbose=0,
                       warm_start=False)

y_pred = clf.predict(X_test)


cm = confusion_matrix(y_test, y_pred)
accuracy_score(y_test, y_pred)
```

Appendix B

Linear Leakage Fraction

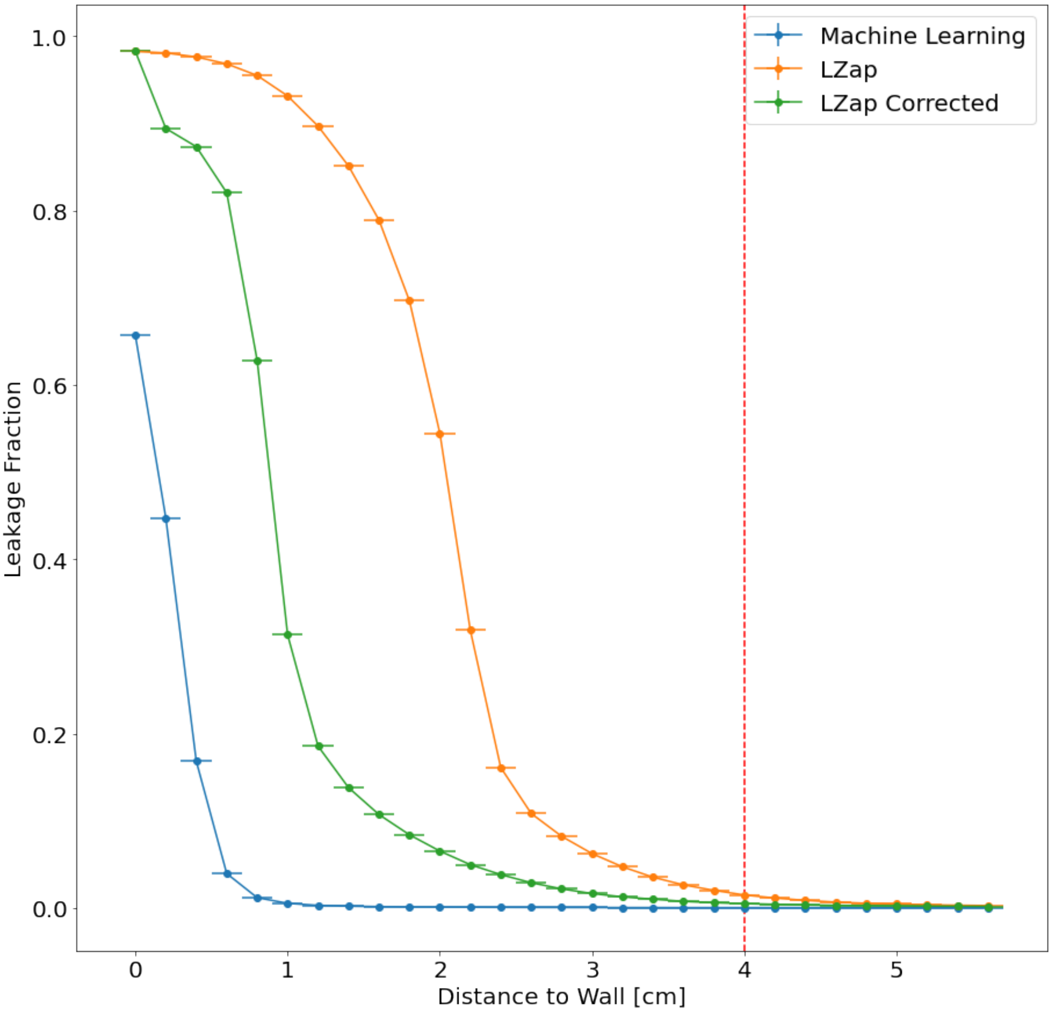


Figure B.1: Image showing the leakage of ML and LZap and LZap corrected (Mercury) on the test dataset with a linear y-axis. The LZap corrected takes into account the drift of electrons towards the centre of the TPC. It is shown separately in this image, but in further images where only LZap in mentioned, the correction is applied.

Bibliography

Bibliography

[1] Bing-Lin Young. A survey of dark matter and related topics in cosmology. *Frontiers of Physics*, 12(2):1–219, 2017.

[2] Giuseppe Gavazzi. 21 centimeter study of spiral galaxies in the coma supercluster. *The Astrophysical Journal*, 320:96–121, 1987.

[3] Gianfranco Bertone and Dan Hooper. History of dark matter. *Reviews of Modern Physics*, 90(4):045002, 2018.

[4] Fritz Zwicky. On the masses of nebulae and of clusters of nebulae. *The Astrophysical Journal*, 86:217, 1937.

[5] Fritz Zwicky. Die rotverschiebung von extragalaktischen nebeln. *Helvetica physica acta*, 6:110–127, 1933.

[6] Vera C Rubin and W Kent Ford Jr. Rotation of the andromeda nebula from a spectroscopic survey of emission regions. *The Astrophysical Journal*, 159: 379, 1970.

[7] KG Begeman and Broeils et al. Extended rotation curves of spiral galaxies: Dark haloes and modified dynamics. *Monthly Notices of the Royal Astronomical Society*, 249(3):523–537, 1991.

[8] Massimo Persic, Paolo Salucci, and Fulvio Stel. The universal rotation curve of spiral galaxies—i. the dark matter connection. *Monthly Notices of the Royal Astronomical Society*, 281(1):27–47, 1996.

[9] Bruce Hoeneisen et al. A study of dark matter with spiral galaxy rota-

tion curves. *International Journal of Astronomy and Astrophysics*, 9(02): 71, 2019.

[10] JR Brownstein and JW Moffat. Galaxy rotation curves without nonbaryonic dark matter. *The Astrophysical Journal*, 636(2):721, 2006.

[11] Claude Carignan and Chemin et al. The extended hi rotation curve and mass distribution of m31. *The Astrophysical Journal Letters*, 641(2):L109, 2006.

[12] Przemek Mróz, Andrzej Udalski, and et al Skowron. Rotation curve of the milky way from classical cepheids. *The Astrophysical Journal Letters*, 870 (1):L10, 2019.

[13] Douglas Clowe and Bradač et al. A direct empirical proof of the existence of dark matter. *The Astrophysical Journal Letters*, 648(2):L109, 2006.

[14] M Markevitch, S Randall, and Clowe et al. Dark matter and the bullet cluster. In *36th COSPAR Scientific Assembly*, volume 36, page 2655, 2006.

[15] CS Kochanek and White et al. Clusters of galaxies in the local universe. *The Astrophysical Journal*, 585(1):161, 2003.

[16] Alexey Vikhlinin and Kravtsov et al. Chandra sample of nearby relaxed galaxy clusters: Mass, gas fraction, and mass-temperature relation. *The Astrophysical Journal*, 640(2):691, 2006.

[17] George F Smoot. The cosmic microwave background spectrum. *arXiv preprint astro-ph/9705101*, 1997.

[18] Yashar Akrami and Ashdown et al. Planck 2018 results-iv. diffuse component separation. *Astronomy & Astrophysics*, 641:A4, 2020.

[19] John C Mather and Fixsen et al. Calibrator design for the cobe* far infrared absolute spectrophotometer (firas). *The Astrophysical Journal*, 512(2):511, 1999.

[20] Peter AR Ade, N Aghanim, and et al Armitage-Caplan. Planck 2013 results.

xv. cmb power spectra and likelihood. *Astronomy & Astrophysics*, 571:A15, 2014.

[21] Scott Dodelson. *Modern Cosmology*. Elsevier, 2003. ISBN 9780122191411.

[22] Katherine Freese. Status of dark matter in the universe. In *The Fourteenth Marcel Grossmann Meeting On Recent Developments in Theoretical and Experimental General Relativity, Astrophysics, and Relativistic Field Theories: Proceedings of the MG14 Meeting on General Relativity, University of Rome "La Sapienza", Italy, 12–18 July 2015*, pages 325–355. World Scientific, 2018.

[23] Veniamin Sergeevich Berezinsky, Vyacheslav Ivanovich Dokuchaev, and Yu N Eroshenko. Small-scale clumps of dark matter. *Physics-Uspekhi*, 57 (1):1, 2014.

[24] Volker Springel and Frenk et al. The large-scale structure of the universe. *nature*, 440(7088):1137–1144, 2006.

[25] Raul E. Angulo and Oliver Hahn. Large-scale dark matter simulations. *Living Reviews in Computational Astrophysics*, 8(1), Feb 2022. ISSN 2365-0524. doi: 10.1007/s41115-021-00013-z. URL `http://dx.doi.org/10.1007/s41115-021-00013-z`.

[26] Carlton Baugh and P Murdin. Correlation function and power spectra in cosmology. *Encyclopedia of Astronomy and Astrophysics,(IOP, London, UK, 2006)*, 2006.

[27] Christopher T et al Hill. Natural theories of ultralow mass pseudo nambu-goldstone bosons: Axions and quintessence. *Physical Review D*, 66(7): 075010, 2002.

[28] Roberto D Peccei and Helen R Quinn. Cp conservation in the presence of pseudoparticles. *Physical Review Letters*, 38(25):1440, 1977.

[29] Steven Weinberg. A new light boson? *Physical Review Letters*, 40(4):223, 1978.

[30] Stephen J Asztalos and et al Carosi. Squid-based microwave cavity search for dark-matter axions. *Physical review letters*, 104(4):041301, 2010.

[31] J Govenius, RE Lake, and Tan et al. Detection of zeptojoule microwave pulses using electrothermal feedback in proximity-induced josephson junctions. *Physical review letters*, 117(3):030802, 2016.

[32] Charles Alcock, RA Allsman, and Alves et al. The macho project: microlensing results from 5.7 years of large magellanic cloud observations. *The Astrophysical Journal*, 542(1):281, 2000.

[33] C Afonso, JN Albert, J Andersen, R Ansari, É Aubourg, P Bareyre, JP Beaulieu, G Blanc, X Charlot, F Couchot, et al. Limits on galactic dark matter with 5 years of eros smc data. *Astronomy & Astrophysics*, 400(3): 951–956, 2003.

[34] Ch Alcock, RA Allsman, D Alves, R Ansari, E Aubourg, TS Axelrod, P Bareyre, J-Ph Beaulieu, AC Becker, DP Bennett, et al. Eros and macho combined limits on planetary-mass dark matter in the galactic halo. *The Astrophysical Journal*, 499(1):L9, 1998.

[35] RA Allsman, DR Alves, and Axelrod et al. Macho project limits on black hole dark matter in the 1-30 solar mass range. *arXiv preprint astro-ph/0011506*, 2000.

[36] Bernard Carr, Kazunori Kohri, and Sendouda et al. Constraints on primordial black holes. *Reports on Progress in Physics*, 84(11):116902, 2021.

[37] V Pierro, LIGO Scientific Collaboration, Virgo Collaboration, et al. Astrophysical implications of the binary black hole merger gw150914. 2016.

[38] Simeon Bird and Cholis et al. Did ligo detect dark matter? *Physical review letters*, 116(20):201301, 2016.

[39] A Kashlinsky. Ligo gravitational wave detection, primordial black holes, and the near-ir cosmic infrared background anisotropies. *The Astrophysical Journal Letters*, 823(2):L25, 2016.

[40] Simon DM White and Frenk et al. Clustering in a neutrino-dominated universe. *The Astrophysical Journal*, 274:L1–L5, 1983.

[41] S.F King. Neutrino mass models. *Reports on Progress in Physics*, 67(2): 107–157, Dec 2003. ISSN 1361-6633. doi: 10.1088/0034-4885/67/2/r01. URL `http://dx.doi.org/10.1088/0034-4885/67/2/R01`.

[42] Scott Dodelson and Lawrence M Widrow. Sterile neutrinos as dark matter. *Physical Review Letters*, 72(1):17, 1994.

[43] Howard Baer and Choi et al. Dark matter production in the early universe: beyond the thermal wimp paradigm. *Physics Reports*, 555:1–60, 2015.

[44] Benjamin W Lee and Steven Weinberg. Cosmological lower bound on heavy-neutrino masses. *Physical Review Letters*, 39(4):165, 1977.

[45] Gerard Jungman and et al Kamionkowski. Supersymmetric dark matter. *Physics Reports*, 267(5-6):195–373, 1996.

[46] Peter Athron, Csaba Balázs, and et al Buckley. Combined collider constraints on neutralinos and charginos. *The European Physical Journal C*, 79(5):1–52, 2019.

[47] Alexandre Arbey and Battaglia et al. Implications of lhc searches on susy particle spectra. *The European Physical Journal C*, 72(1):1–14, 2012.

[48] Serguei Chatrchyan, Vardan Khachatryan, Albert M Sirunyan, A Tumasyan, W Adam, T Bergauer, M Dragicevic, J Erö, C Fabjan, M Friedl, et al. Search for supersymmetry at the lhc in events with jets and missing transverse energy. *Physical review letters*, 107(22):221804, 2011.

[49] Arvind Rajaraman, William Shepherd, Tim MP Tait, and Alexander M Wi-

jangco. Lhc bounds on interactions of dark matter. *Physical Review D*, 84 (9):095013, 2011.

[50] Abdessamad Abada and Salah Nasri. Constraining the parameters of a model for cold dark matter. *Trends in Modern Cosmology*, 2017.

[51] S Berlendis, E Cheu, and Delitzsch et al. Constraints on mediator-based dark matter and scalar dark energy models using root s= 13 tev pp collision data collected by the atlas detector. 2019.

[52] A Drukier and Leo Stodolsky. Principles and applications of a neutral-current detector for neutrino physics and astronomy. *Physical Review D*, 30(11): 2295, 1984.

[53] David G Cerdeno and Anne M Green. Direct detection of wimps. *arXiv preprint arXiv:1002.1912*, 2010.

[54] Mark J Reid. The distance to the center of the galaxy. *Annual review of astronomy and astrophysics*, 31(1):345–372, 1993.

[55] Justin I Read. The local dark matter density. *Journal of Physics G: Nuclear and Particle Physics*, 41(6):063101, 2014.

[56] Collaboration Gaia and et al Brown. Gaia data release 2 summary of the contents and survey properties. *Astronomy & Astrophysics*, 616(1), 2018.

[57] Andrzej K Drukier and Freese et al. Detecting cold dark-matter candidates. *Physical Review D*, 33(12):3495, 1986.

[58] Frank J Kerr and Donald Lynden-Bell. Review of galactic constants. *Monthly Notices of the Royal Astronomical Society*, 221(4):1023–1038, 1986.

[59] Chris Kelso, Christopher Savage, and et al Valluri. The impact of baryons on the direct detection of dark matter. *Journal of Cosmology and Astroparticle Physics*, 2016(08):071, 2016.

[60] PF Smith. Dark matter detection, phys. rept. 187, 203 (1990); jd lewin and pf smith, review of mathematics, numerical factors, and corrections for dark

matter experiments based on elastic nuclear recoil. *Astropart. Phys*, 6:87, 1996.

[61] JD Lewin and PF Smith. Astropart. *Phys*, 6(8):I996, 1996.

[62] D Cerdeno. *Dark Matter 101: from production to detection.* Lecture Series, Higgs Centre of Theoretical Physics, 2016.

[63] Vitaly Chepel and Henrique Araujo. Liquid noble gas detectors for low energy particle physics. *Journal of Instrumentation*, 8(04):R04001, 2013.

[64] Laura Baudis. Direct dark matter detection: the next decade. *Physics of the Dark Universe*, 1(1-2):94–108, 2012.

[65] Marc Schumann. Direct detection of wimp dark matter: concepts and status. *Journal of Physics G: Nuclear and Particle Physics*, 46(10):103003, 2019.

[66] Q Arnaud and et al Armengaud. First germanium-based constraints on submev dark matter with the edelweiss experiment. *Physical Review Letters*, 125(14):141301, 2020.

[67] R Agnese and et al Anderson. Projected sensitivity of the supercdms snolab experiment. *Physical Review D*, 95(8):082002, 2017.

[68] R Agnese and et al Aramaki. Results from the super cryogenic dark matter search experiment at soudan. *Physical review letters*, 120(6):061802, 2018.

[69] DS Akerib and Bai et al. The large underground xenon (lux) experiment. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 704:111–126, 2013.

[70] Elena Aprile and Aalbers et al. Search for wimp inelastic scattering off xenon nuclei with xenon100. *Physical Review D*, 96(2):022008, 2017.

[71] GJ Alner, HM Araújo, and et al Bewick. Limits on spin-dependent wimpnucleon cross-sections from the first zeplin-ii data. *Physics Letters B*, 653 (2-4):161–166, 2007.

[72] VN Lebedenko and Araujo et al. Results from the first science run of the zeplin-iii dark matter search experiment. *Physical Review D*, 80(5):052010, 2009.

[73] BJ Mount. Lux-zeplin (lz) technical design report. Technical report, Argonne National Lab.(ANL), Argonne, IL (United States); Pacific Northwest, 2017.

[74] F Ruppin and et al Billard. Complementarity of dark matter detectors in light of the neutrino background. *Physical Review D*, 90(8):083510, 2014.

[75] George Fitzgerald Smoot et al. Review of particle physics. *Progress of Theoretical and Experimental Physics*, page 1, 2020.

[76] E Aprile and Aalbers et al. Projected wimp sensitivity of the xenonnt dark matter experiment. *Journal of Cosmology and Astroparticle Physics*, 2020 (11):031, 2020.

[77] DS Akerib and Akerlof et al. Projected wimp sensitivity of the lux-zeplin dark matter experiment. *Physical Review D*, 101(5):052002, 2020.

[78] J Aalbers and et al Agostini. Darwin: towards the ultimate dark matter detector. *Journal of Cosmology and Astroparticle Physics*, 2016(11):017, 2016.

[79] Marc Schumann, Laura Baudis, Lukas Bütikofer, Alexander Kish, and Marco Selvi. Dark matter sensitivity of multi-ton liquid xenon detectors. *Journal of cosmology and astroparticle physics*, 2015(10):016, 2015.

[80] David R Nygren. The time projection chamber. 1978.

[81] Pietro Benetti and Calligarich et al. Detection of energy deposition down to the kev region using liquid xenon scintillation. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 327(1):203–206, 1993.

[82] DS Akerib and Alsum et al. Limits on spin-dependent wimp-nucleon cross section obtained from the complete lux exposure. *Physical review letters*, 118(25):251302, 2017.

[83] E Aprile and T Doke. Liquid xenon detectors for particle physics and astrophysics. *Reviews of Modern Physics*, 82(3):2053, 2010.

[84] Carl Eric Dahl. *The physics of background discrimination in liquid xenon, and first results from Xenon10 in the hunt for WIMP dark matter*. PhD thesis, Princeton University, 2009.

[85] Akira Hitachi and Takahashi et al. Effect of ionization density on the time dependence of luminescence from liquid argon and xenon. *Physical Review B*, 27(9):5279, 1983.

[86] Peter Sorensen. Electron train backgrounds in liquid xenon dark matter search detectors are indeed due to thermalization and trapping. *arXiv preprint arXiv:1702.04805*, 2017.

[87] Jens Lindhard and et al Nielsen. Integral equations governing radiation effects. *Mat. Fys. Medd. Dan. Vid. Selsk*, 33(10):1–42, 1963.

[88] L de Viveiros and Lindote et al. Tritium calibration of the lux dark matter experiment. 2016.

[89] DS Akerib and Alsum et al. Signal yields, energy resolution, and recombination fluctuations in liquid xenon. *Physical Review D*, 95(1):012008, 2017.

[90] Tadayoshi Doke and Hitachi et al. Absolute scintillation yields in liquid argon and xenon for various particles. *Japanese journal of applied physics*, 41(3R):1538, 2002.

[91] DS Akerib, CW Akerlof, and Alsum et al. Projected wimp sensitivity of the lux-zeplin dark matter experiment. *Physical Review D*, 101(5):052002, 2020.

[92] P Brás and Lindote et al. Identification of radiopure titanium for the lz dark matter experiment and future rare event searches. 2017.

[93] Francisco Neves and Lindote et al. Measurement of the absolute reflectance of polytetrafluoroethylene (ptfe) immersed in liquid xenon. *Journal of Instrumentation*, 12(01):P01017, 2017.

[94] SJ Haselschwardt and Shaw et al. A liquid scintillation detector for radioassay of gadolinium-loaded liquid scintillator for the lz outer detector. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 937:148–163, 2019.

[95] Eryk Druszkiewicz. The data acquisition system for lz. *Journal of Instrumentation*, 11(02):C02072, 2016.

[96] DS Akerib and Araújo et al. Tritium calibration of the lux dark matter experiment. *Physical Review D*, 93(7):072009, 2016.

[97] DS Akerib and Alsum et al. Low-energy (0.7-74 kev) nuclear recoil calibration of the lux dark matter experiment using dd neutron scattering kinematics. *arXiv preprint arXiv:1608.05381*, 2016.

[98] FE Gray and Ruybal et al. Cosmic ray muon flux at the sanford underground laboratory at homestake. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 638(1):63–66, 2011.

[99] Lindote Reichhart and Lindote et al. Measurement and simulation of the muon-induced neutron yield in lead. *Astroparticle physics*, 47:67–76, 2013.

[100] V Tomasello and Robinson et al. Radioactive background in a cryogenic dark matter experiment. *Astroparticle Physics*, 34(2):70–79, 2010.

[101] Melvin W Carter and A Alan Moghissi. Three decades of nuclear testing. *Health Physics*, 33(1):55–71, 1977.

[102] DP Child and MAC Hotchkis. Plutonium and uranium contamination in soils from former nuclear weapon test sites in australia. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, 294:642–646, 2013.

[103] Yoshua Bengio, Ian Goodfellow, and Aaron Courville. *Deep learning*, volume 1. MIT press Massachusetts, USA:, 2017. ISBN 9780262035613.

[104] Annina Simon and Mahima Singh. An overview of m learning and its ap. *International Journal of Electrical Sciences Electrical Sciences & Engineering (IJESE)*, 22, 2015.

[105] Michael I Jordan and Tom M Mitchell. Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245):255–260, 2015.

[106] Amirsina Torfi and Shirvani et al. Natural language processing advancements by deep learning: A survey. *arXiv preprint arXiv:2003.01200*, 2020.

[107] Leo Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.

[108] Douglas C Montgomery, Elizabeth A Peck, and G Geoffrey Vining. *Introduction to linear regression analysis*. John Wiley & Sons, 2021.

[109] Geoffrey E Hinton and Ruslan R Salakhutdinov. Reducing the dimensionality of data with neural networks. *science*, 313(5786):504–507, 2006.

[110] John A Richards. Feature reduction. In *Remote Sensing Digital Image Analysis*, pages 403–446. Springer, 2022.

[111] Svante Wold, Kim Esbensen, and Paul Geladi. Principal component analysis. *Chemometrics and intelligent laboratory systems*, 2(1-3):37–52, 1987.

[112] Ian T Jolliffe and Jorge Cadima. Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2065):20150202, 2016.

[113] Adi L Tarca, Vincent J Carey, Xue-wen Chen, Roberto Romero, and Sorin Drăghici. Machine learning and its applications to biology. *PLoS computational biology*, 3(6):e116, 2007.

[114] Harry Surden. Machine learning and law: an overview. *Research Handbook on Big Data Law*, 2021.

[115] Jonathan J Davis and Andrew J Clark. Data preprocessing for anomaly based

network intrusion detection: A review. *computers & security*, 30(6-7):353–375, 2011.

[116] Jenna Wiens and Erica S Shenoy. Machine learning for healthcare: on the verge of a major shift in healthcare epidemiology. *Clinical Infectious Diseases*, 66(1):149–153, 2018.

[117] Alexander Radovic, Mike Williams, David Rousseau, Michael Kagan, Daniele Bonacorsi, Alexander Himmel, Adam Aurisano, Kazuhiro Terao, and Taritree Wongjirad. Machine learning at the energy and intensity frontiers of particle physics. *Nature*, 560(7716):41–48, 2018.

[118] Georges Aad, Tatevik Abajyan, B Abbott, J Abdallah, S Abdel Khalek, Ahmed Ali Abdelalim, R Aben, B Abi, M Abolins, OS AbouZeid, et al. Observation of a new particle in the search for the standard model higgs boson with the atlas detector at the lhc. *Physics Letters B*, 716(1):1–29, 2012.

[119] Christopher J Fluke and Colin Jacobs. Surveying the reach and maturity of machine learning and artificial intelligence in astronomy. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 10(2):e1349, 2020.

[120] Bradley C Whitmore and Duncan A Forbes. An objective classification system for spiral galaxies and its relationship to the [n ii]/[s ii] ratio. In *The World of Galaxies*, pages 95–98. Springer, 1989.

[121] Shankar Agarwal, Romeel Davé, and Bruce A Bassett. Painting galaxies into dark matter haloes using machine learning. *Monthly Notices of the Royal Astronomical Society*, 478(3):3410–3422, 2018.

[122] Romeel Davé, Robert Thompson, and Philip F Hopkins. Mufasa: galaxy formation simulations with meshless hydrodynamics. *Monthly Notices of the Royal Astronomical Society*, 462(3):3265–3284, 2016.

[123] Elena Aprile, J Aalbers, F Agostini, M Alfonsi, FD Amaro, M Anthony,

B Antunes, F Arneodo, M Balata, P Barrow, et al. The xenon1t dark matter experiment. *The European Physical Journal C*, 77(12):1–23, 2017.

[124] Charanjit K Khosa, Lucy Mars, Joel Richards, and Veronica Sanz. Convolutional neural networks for direct detection of dark matter. *Journal of Physics G: Nuclear and Particle Physics*, 47(9):095201, 2020.

[125] P Brás, F Neves, A Lindote, A Cottle, R Cabrita, E Lopez Asamar, G Pereira, C Silva, V Solovov, and MI Lopes. A machine learning-based methodology for pulse classification in dual-phase xenon time projection chambers. *arXiv preprint arXiv:2201.05659*, 2022.

[126] Tin Kam Ho. Random decision forests. In *Proceedings of 3rd international conference on document analysis and recognition*, volume 1, pages 278–282. IEEE, 1995.

[127] Tom M Mitchell et al. Machine learning. 1997.

[128] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[129] Martin Anthony and Peter L Bartlett. *Neural network learning: Theoretical foundations*. cambridge university press, 2009. ISBN 9780521573535.

[130] Fernanda Psihas, Micah Groh, Christopher Tunnell, and Karl Warburton. A review on machine learning for neutrino experiments. *International Journal of Modern Physics A*, 35(33):2043005, Nov 2020. ISSN 1793-656X. doi: 10.1142/s0217751x20430058. URL `http://dx.doi.org/10.1142/S0217751X20430058`.

[131] Muhammad Uzair and Noreen Jamil. Effects of hidden layers on the efficiency of neural networks. In *2020 IEEE 23rd international multitopic conference (INMIC)*, pages 1–6. IEEE, 2020.

[132] Florian Marquardt. Machine Learning and Quantum Devices. *SciPost Phys. Lect. Notes*, page 29, 2021. doi: 10.21468/SciPostPhysLectNotes.29. URL

https://scipost.org/10.21468/SciPostPhysLectNotes.29.

[133] Hong Hui Tan and King Hann Lim. Vanishing gradient mitigation with deep learning neural network optimization. In *2019 7th International Conference on Smart Computing & Communications (ICSCC)*, pages 1–4. IEEE, 2019.

[134] Saad Albawi, Tareq Abed Mohammed, and Saad Al-Zawi. Understanding of a convolutional neural network. In *2017 international conference on engineering and technology (ICET)*, pages 1–6. Ieee, 2017.

[135] Frank Rosenblatt. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, 65(6):386, 1958.

[136] Tom Fawcett. An introduction to roc analysis. *Pattern recognition letters*, 27 (8):861–874, 2006.

[137] Muhamad et al Yani. Application of transfer learning using convolutional neural network method for early detection of terry's nail. In *Journal of Physics: Conference Series*, volume 1201, page 012052. IOP Publishing, 2019.

[138] Sai Balaji. Binary image classifier cnn using tensorflow. *Medium*, Aug 2020. URL https://medium.com/techiepedia/binary-image-classifier-cnn-using-tensorflow-a3f5d6746697.

[139] Sea Agostinelli and Allison et al. Geant4—a simulation toolkit. *Nuclear instruments and methods in physics research section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3):250–303, 2003.

[140] DS Akerib and Bai et al. Luxsim: A component-centric approach to low-background simulations. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 675:63–77, 2012.

[141] Jeremy Mock and Barry et al. Modeling pulse characteristics in xenon with nest. *Journal of Instrumentation*, 9(04):T04002, 2014.

[142] Ilka Antcheva and Ballintijn et al. Root—a c++ framework for petabyte data storage, statistical analysis and visualization. *Computer Physics Communications*, 182(6):1384–1385, 2011.

[143] Glen Cowan and Cranmer et al. Asymptotic formulae for likelihood-based tests of new physics. *The European Physical Journal C*, 71(2):1–19, 2011.

[144] Wolfgang A Rolke and Lopez et al. Limits and confidence intervals in the presence of nuisance parameters. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 551(2-3):493–503, 2005.

[145] D.S. Akerib and Akerlof et al. The lux-zeplin (lz) experiment. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 953:163047, Feb 2020. ISSN 0168-9002. doi: 10.1016/j.nima.2019.163047. URL http://dx.doi.org/10.1016/j.nima.2019.163047.

[146] JR Partington. Discovery of radon. *Nature*, 179(4566):912–912, 1957.

[147] DS Akerib and Akerlof et al. The lux-zeplin (lz) radioactivity and cleanliness control programs. *The European Physical Journal C*, 80(11):1–52, 2020.

[148] Umit Utku. *Background and sensitivity studies for the LUX-ZEPLIN dark matter experiment*. PhD thesis, UCL (University College London), 2021.

[149] Jaret Heise. The sanford underground research facility at homestake. In *Journal of Physics: Conference Series*, volume 606, page 012015. IOP Publishing, 2015.

[150] MH Magalhães and Amaral et al. Radon-222 in brazil: an outline of indoor and outdoor measurements. *Journal of environmental radioactivity*, 67(2): 131–143, 2003.

[151] Lawrence Berkely National Institute. Nuclear forensic search project. *Nuclear Forensics*, 2014. URL `https://metadata.berkeley.edu/nuclear-forensics/index.html`.

[152] M Bruemmer and Nakib et al. Studies on the reduction of radon plate-out. In *AIP Conference Proceedings*, volume 1672, page 140005. AIP Publishing LLC, 2015.

[153] Glenn T Seaborg, Walter D Loveland, and David J Morrissey. Modern nuclear chemistry, 2005.

[154] James Mott. *Search for double beta decay of 82Se with the NEMO-3 detector and development of apparatus for low-level radon measurements for the SuperNEMO experiment*. PhD thesis, University College London (University of London), 2014.

[155] Theresa Fruth. *PMT Studies and Loop Antenna Development for the LZ Dark Matter Search*. PhD thesis, University of Oxford, 2019.

[156] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, et al. Scikit-learn: Machine learning in python. *the Journal of machine Learning research*, 12:2825–2830, 2011.

[157] Yong Yu, Xiaosheng Si, Changhua Hu, and Jianxun Zhang. A review of recurrent neural networks: Lstm cells and network architectures. *Neural computation*, 31(7):1235–1270, 2019.

[158] DS Akerib and Alsum et al. Position reconstruction in lux. *Journal of Instrumentation*, 13(02):P02001, 2018.

[159] LW Kastens and Cahn et al. Calibration of a liquid xenon detector with kr 83 m. *Physical Review C*, 80(4):045809, 2009.

[160] AD McNaught and A Wilkinson. Iupac compendium of chemical terminology, 2nd edn.(the “gold book”) blackwell scientific publications, 1997.

[161] A Manalaysay and Undagoitia et al. Spatially uniform calibration of a liquid xenon detector at low energies using k 83 mr. *Review of Scientific Instruments*, 81(7):073303, 2010.

[162] Markus Löning and et al Bagnall. sktime: A unified interface for machine learning with time series. *arXiv preprint arXiv:1909.07872*, 2019.

[163] Anthony Bagnall and Lines et al. The great time series classification bake off: a review and experimental evaluation of recent algorithmic advances. *Data mining and knowledge discovery*, 31(3):606–660, 2017.

[164] Hassan Ismail Fawaz and Forestier et al. Deep learning for time series classification: a review. *Data mining and knowledge discovery*, 33(4):917–963, 2019.

[165] Meinard Müller. Dynamic time warping. *Information retrieval for music and motion*, pages 69–84, 2007.