Exploring Dark Matter Interactions: Extending the LUX-ZEPLIN Search to Non-Standard WIMP–Nucleon Interactions, Anomaly Finding Techniques, and Liquid Xenon Detector Characterization

by

Chami S. Amarasinghe

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Physics)
in The University of Michigan
2023

Doctoral Committee:

Professor Wolfgang Lorenzon, Co-Chair
Associate Professor Bjoern Penning, Co-Chair
Professor Dragan Huterer
Professor Sara Pozzi
Associate Professor Joshua Spitz

Chami S. Amarasinghe

amarascs@umich.edu

ORCID iD: 0000-0001-7789-651X

This thesis is dedicated to my parents who have given me everything; a love for nature and the provisions for my journey to study it.

Acknowledgements

Being advised by Prof. Wolfgang Lorenzon was a privilege through which I sampled the creative and collaborative challenges of experimental particle physics. I am deeply grateful for your unwavering guidance. My gratitude extends to the dissertation committee, first to Prof. Josh Spitz and Prof. Sara Pozzi for your time spent on this behalf. Many special thanks to Prof. Bjoern Penning for co-advising and supporting me in my next step on the academic path. Thank you Prof. Dragan Huterer for your guidance through the years, not to mention the convincing arguments for why cosmology is the queen of physics. My heartfelt thanks go to my first research advisors Prof. Christine Aidala and Gabriele Carcassi. Profs. Ratindranath Akhoury, Greg Tarle, and Aaron Pierce, thank you very much for your kindnesses and counsel.

I owe many thanks to Dr. Richard Raymond and Maris Arthurs for introducing me to the joys of hardware. I am especially grateful to Maris and Dongqing Huang for years of partnership and friendship, and it was wonderful to have had incredible labmates like Ruben Coronel, Samara Steinfeld, Greg Rischbieter, Erin Barillier, Ilyas Adnane, Al Kucich, Yi Liu, and Douglas Woodward. Thank you to our group members Luke Korley, Michael Williams, Harvey Birch, Haley Reid, Sam Eriksen, Geertje Heuermann, Chris Wright, and Miguel Hernandez for the collegiality and sorrowful laments against the machine.

Through the LZ collaboration I met many people who I am lucky to have learned from. I am greatly indebted to Prof. Scott Kravitz for his wonderful mentorship during our initial forays into machine learning. I would also like to deeply thank Prof. Hugh

Lippincott, Sally Shaw, Alissa Monte, and Scott Haselschwardt for your support from the filling of the LZ outer detector onward. I owe so much to T.J. Whitis, David Woodward, Derek Lucero, and Alex Geffre for showing me the ways of the underground. Dr. Peter Sorensen has always been a beacon of support; thank you very much. Finally, it was a pleasure to work closely with Greg Rischbieter, Billy Boxer, and Sam Eriksen on the high energy search. I eagerly anticipate working with all of you on new challenges.

Without my friends the PhD student life would have been a bore. May the Center for Brains and Gains, which provided lifelong friendships and a sense of intellectual belonging, continue to proliferate and inspire. I am eternally grateful to my parents for supporting me on this journey to discover something that I do not have the words in my mother tongue to describe. Your patience and understanding is unbounded. Finally, the encouragement and support of my partner Bella was paramount to this work, and to really express that gratitude, compounded with all the rest, no words in any language will do.

Table of Contents

List of Tables

List of Figures

List of Appendices

Abstract

Dark matter is a prime mystery in modern physics. Independent measurements point to a quarter of the universe's mass–energy (and 85% by mass alone) being made up of a cold, collisionless species of matter. Its ubiquity in the universe, confirmed through gravitational phenomena and cosmological evidence, belies the fact that dark matter particles have not been observed to interact directly with regular atoms. A leading candidate for dark matter is the Weakly Interacting Massive Particle (WIMP), which has a mass between a proton and a complex protein, and interacts with regular matter with a strength similar to the electroweak force (mediated by the W and Z particles). A series of experiments have unsuccessfully searched for evidence of galactic WIMPs passing through Earth, culminating in the LUX-ZEPLIN (LZ) experiment's result in July 2022 that placed the most stringent upper bounds on the WIMP–nucleon interaction strength.

This thesis extends LZ's first result by expanding the energy window of the search (by a factor of 7.5) to be sensitive to non-standard WIMP–nucleon interactions. No evidence for WIMPs is found in this new analysis region, but the strengths of 54 distinct WIMP-nucleon interactions, generated by an Effective Field Theory (EFT), are constrained by the analysis. The EFT is a theoretical tool that enumerates all of the possible interactions distinguishable at the energy scale of the WIMP–nucleon scatter; this LZ EFT result provides the tightest constraints for nearly all of these interactions.

Three other topics are presented in addition to the LZ EFT analysis. Firstly, a study to characterize and remove data that mimic non-standard WIMPs is described, enabling the extension of the energy window to where such signals are expected. Secondly, the question of anomalous experimental data caused by noise, detector and signal processing effects is addressed, leading to the development of an interpretable general-purpose anomaly finding tool using techniques in machine learning. Applications of the anomaly finder on simulated and real LZ data are also shown. Thirdly, this thesis presents the conceptualization and partial results of an experimental study with the Michigan Xenon (MiX) detector to precisely measure the electron and photon production yields of liquid xenon, the detector medium employed in LZ and other WIMP detectors.

Chapter 1

Dark Matter and Direct Detection

I can never look now at the Milky Way
without wondering from which of those
banked clouds of stars the emissaries are
coming.

Arthur C. Clarke

A wealth of evidence supports the claim that there is a puzzlingly large, but invisible component of matter that fills the universe. The observational data span from the present day back to early epochs after the big bang, from the motion of stars and galaxies to the structure of superclusters. These data furnish a strong case for a species of matter that was instrumental for the structure formation in the early universe, which led to the creation of stars, planets, and eventually us, here to ponder its nature. The case for the existence of dark matter is strong. But there is virtually no data indicating that it interacts non-gravitationally, so we have no knowledge of its nature other than what can be inferred from its distribution in the universe and its cosmological history.

As to be expected given the magnitude of this mystery, there is no shortage in the volume and diversity of experiments, ideas, and opinions, nor are there any signs of slowing even as we approach a century of research. The spectrum of hypotheses addressing the observational data is broad. The dominant paradigm holds that dark matter is comprised of particles, not yet in the Standard Model (SM) of particle physics. Even within this paradigm, candidate particles for dark matter may have masses that range tens of orders of magnitude, from ‘ultralight’ bosons (10^{-22} eV) to ‘superheavy’ fundamental particles at the Planck scale (10^{19} GeV) [1, 2]. A portion of the dark matter might even be in the form of primordial black holes, each with the mass of several suns [3]. In this dizzying array, however, there are a handful of candidates that have synergies with well-motivated theories and thus warrant closer study.

One well-motivated candidate for dark matter is called the Weakly Interacting Massive Particle (WIMP). The motivation for WIMPs comes from supersymmetric extensions to the Standard Model, which predict a neutral heavy (GeV to TeV) particle that can be assigned a plausible cosmological history. This chance meeting of particle physics and cosmology, known as the ‘WIMP miracle’, continues to spawn numerous experiments dedicated to hunting WIMPs. None so far have made a discovery. Coupled with the non-detection of supersymmetric particles at the Large Hadron Collider (LHC), the null results of dark matter experiments have caused the WIMP hypothesis to wane [4]. This thesis presents an experimental search that goes beyond the simplest WIMP models, extending the search to include dark matter that may have non-standard interactions with atomic nuclei. The analysis is conducted with the first science data of the LUX-ZEPLIN (LZ) experiment, the most sensitive WIMP detector at time of writing [5]. LZ will be described in Chapter 2. Here we review the evidence for dark matter and present the WIMP.

1.1 A Wealth of Evidence

The range of scales present in the evidence for dark matter can be showcased by examining the historical account. The early evidence was primarily dynamic; the existence of invisible matter was inferred from the motion of large astronomical objects under gravity. We shall trace the gathering of this evidence in the 20th century, starting from galaxy clusters in the 1930s to individual galaxies in the 1970s. In the latter half of the century, this line of attack was complemented by the rapidly growing field of cosmology, which not only provided independent confirmations for the existence of dark matter but also made the most precise measurements of its abundance.

1.1.1 Dynamic Arguments

Galaxy Clusters

The starting points for dark matter research in the 20th century are galaxy redshift data, which record the stretching of light emitted by distant galaxies due to the expansion of space. These data encode the velocities of individual galaxies in a galaxy cluster, which can be used to calculate the dynamic mass of the entire cluster. Luminous mass, on the other hand, is obtained by counting how much light is emitted from the individual galaxies. If galaxies were exclusively made of luminous matter, then the dynamic and luminous masses would be the same. The first well-advertised evidence for the existence of dark matter was a marked disagreement between the dynamic mass and the luminous mass reported by Zwicky in 1933 [6]. This discrepancy was foreshadowed much earlier, for instance by Bessel, writing about binary star systems almost a century before Zwicky: "But light is no real property of mass. The existence of numberless visible stars can prove nothing against the evidence of numberless invisible ones" [7].

Zwicky estimated the line-of-sight dispersion velocities of galaxies in the Coma cluster, a large galaxy cluster around 100 Mpc away, and found the result larger than expected for a cluster of that size [8]. If the cluster mass was exclusively luminous, the dispersion velocities indicated that the individual galaxies were flying apart too fast to be gravitationally bound together. The presence of a large non-luminous component of matter was required to support the large dispersion velocities . Zwicky employed the virial theorem in his estimate, extracting the kinematic properties of objects that move under potential forces. For a gravitationally bound system such as a galaxy cluster

$$M\langle v_{||}^2\rangle \approx \frac{GM}{R}, \tag{1.1}$$

where M is the mass of the collection of galaxies, R is its spatial extent, $\langle v_{||}\rangle$ is the line-of-sight velocity, and G is Newton's constant. The relationship in Eq. 1.1 relies on a time averaging of the momenta and positions of the particles in a system. The time averaging converges only for stable orbits, hence restricting the applications of the theorem to properly virialized systems, an issue that cast doubt on similar studies of galaxy clusters [9]. Further, Zwicky had only estimates of the average mass of a galaxy and the physical size of the cluster, in addition to an overestimated value of Hubble's constant, all of which he used to calculate the dispersion velocities. Even further, the redshift data available at the time described only eight galaxies in the cluster.

Despite its limitations, Zwicky's result was intriguing enough for astronomers to use the same approach on other clusters. For instance, the dynamic mass of the Virgo cluster was estimated to be around a hundred times larger than the combined mass of its individual galaxies [10]. For smaller ensembles of galaxies, the dynamic evidence for dark matter becomes fraught with systematic uncertainties, but arguments in favor were put forth for galaxy groupings down to binary systems [9]. As the problem of the missing mass did not seem like it was going away, astronomers turned their attention to eliminating candidates such as gas, the intracluster medium (ICM), and compact objects like dwarf stars that could make up the dark matter. It was apparent that more information was needed to understand the anomalies in galactic dispersion velocities. The next line of evidence would come from a further reduction in scale; studies of individual galaxies themselves.

Galaxies

Galaxy rotation data provided strength to the claim that there is invisible mass accompanying astronomical objects, this time within individual galaxies. The orbital speeds of stars and dust are expected to fall off with distance from the galactic center. This is predicted by the balance of the centripetal force and gravity, leading to

$$v_{\mathrm{c}} = \sqrt{\frac{GM(r)}{r}}, \tag{1.2}$$

where v_{c} is the circular speed of the star (or dust), G is Newton's constant, $M(r)$ is the mass contained within the orbital radius r. In the outer regions of spiral galaxies where stars are sparse, if luminosity is a good tracer of mass, M is expected to be roughly constant with respect to r such that $v_c \propto 1/\sqrt{r}$ strictly holds.

Initial optical measurements of the rotation curves in spiral galaxies, particularly our neighboring Andromeda galaxy (M31), covered enough of the outer regions to indicate a tension with the expected fall off in stellar speeds [11]. Eventually with the advent of radio astronomy, measurements of the 21 cm H-α line at larger radii were made possible [12]. Until the 1970s, astronomers doubted the significance of these observations due to discrepancies between the optical and radio data, concerns of dust obscuring the spectral emissions, and other systematic uncertainties. However, most doubts were alleviated when new measurements of Andromeda were made by Rubin and Ford using their image tube spectrograph [13]. Figure 1.1 shows a rotation curve made using both optical and radio frequency measurements averaged over a set of spiral galaxies [14]. These data imply that the radial extent of spiral galaxies far exceeds the edges indicated by photometric measurements, and that a significant amount of dark matter is required for the rotation curves to maintain their flatness.

N-body simulations of galaxy formation

In the 1960's, the invention of the transistor enabled researchers studying galaxy formation to simulate the gravitational interactions between physically useful numbers of particles. Early computer simulations of N = 10^5 particles starting in circular orbits showed pronounced instabilities after orbiting only a couple of periods. An example simulation is shown in Figure 1.2, copied from Ref [16]. These simulations, although they used a tiny fraction of stars in an actual

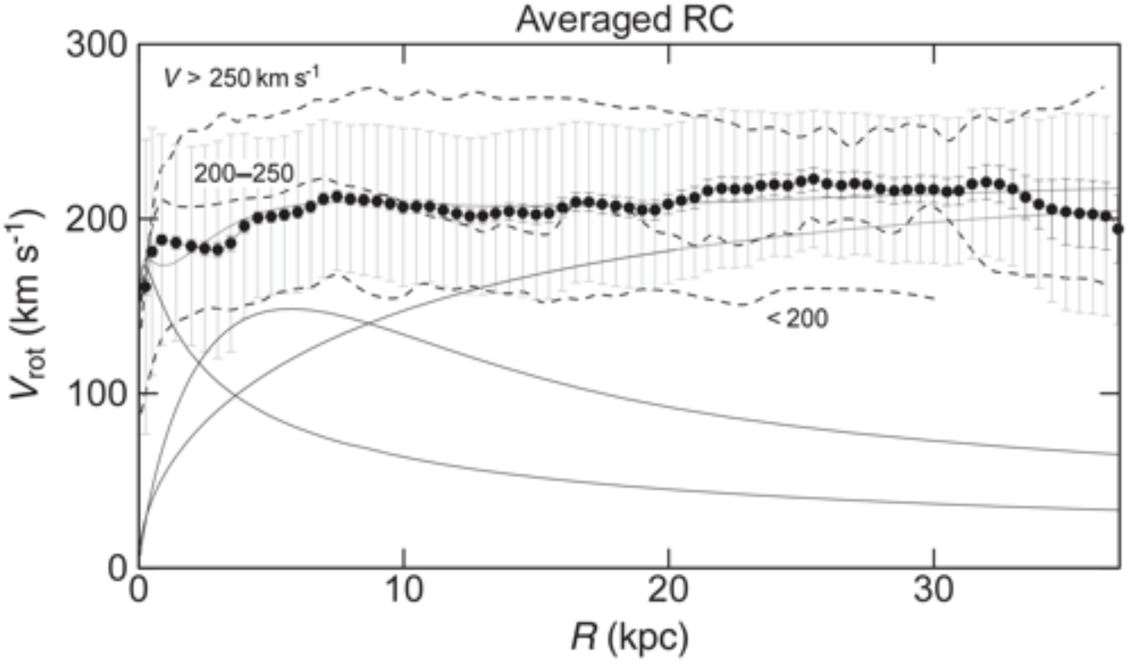


Figure 1.1: Average rotation velocities (black points) from all the galaxies listed in Ref [15], which combines optical and radio data. The three black unbroken lines indicate best fits to the data using two baryonic components (two decaying lines), and a dark matter halo component (monotonically increasing line). Figure copied from Ref [14].

spiral galaxy and made use of many approximation schemes, produced unstable galaxies that disintegrated in a short time. This was at odds with real spiral galaxies, which are clearly rotationally supported by the centripetal acceleration of gravity. The stability of spiral galactic disks, along with the formation of bars seen in nature, could be obtained by tweaking the simulation, but a more satisfactory solution was obtained by Ostriker and Peebles when they ran simulations demonstrating that a spherical massive halo extending beyond stars could stabilize the rotation of the disk [17].

Gravitational weak lensing

Light from distant galaxies on its way to Earth can be lensed by the curved spacetime around massive objects in a phenomenon known as gravitational lensing [18]. Lensing observations can be used to infer the presence of dark matter in foreground galaxy clusters. A celebrated example of this is seen in the Bullet Cluster, shown in Figure 1.3. The picture actually shows two galaxy clusters (the Bullet Cluster proper is the smaller of the two) that have undergone a merger. Most of the visible mass in the clusters is contained in the intracluster medium (ICM), which was strongly impeded in the collision, producing x-rays. The galaxies themselves are too diffuse to have collided but decelerated gravitationally, passing through mostly unaffected. Gravitational lensing of the system revealed most of the mass to also have passed through unimpeded by the collision, coincident with the galaxy distribution. The Bullet Cluster thus provides a strong piece of evidence for the existence of non-radiating collisionless mass identified as dark matter [19].

1.1.2 Cosmological evidence

The argument for dark matter is further strengthened when the cosmological evidence is added to the astrophysical observations discussed above. Dark matter has to be inescapably included in the standard model of cosmology, Λ-Cold Dark Matter (ΛCDM), which is the current best model that accounts for the cosmic microwave background (CMB), large scale structure, iso-

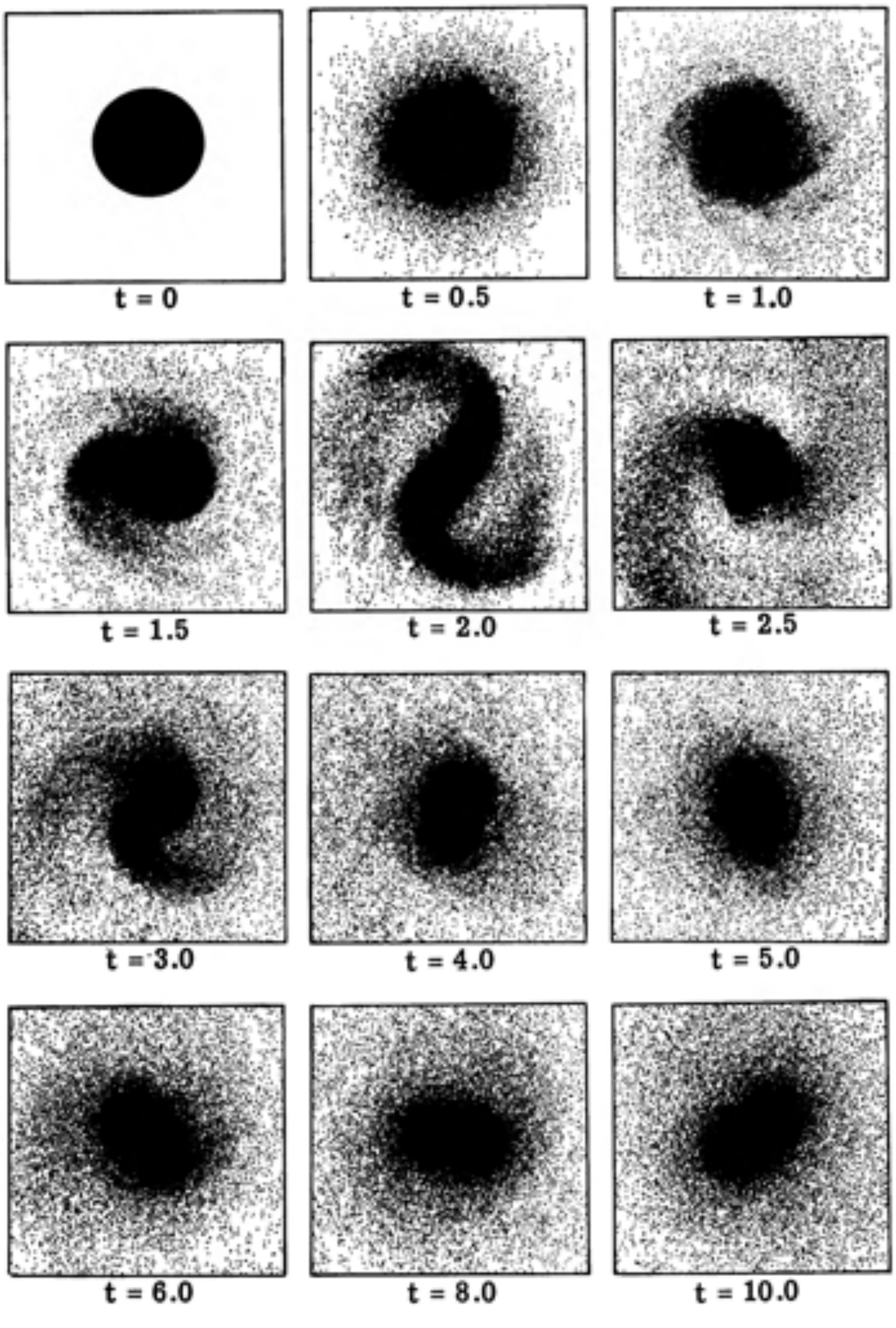


Figure 1.2: Simulation of a uniformly rotating disk of 10^5 stars under gravity. The evolution shows the formation of a barred spiral structure and its subsequent dissolution. The time unit t is the period of the orbit. Figure taken from Ref [16].

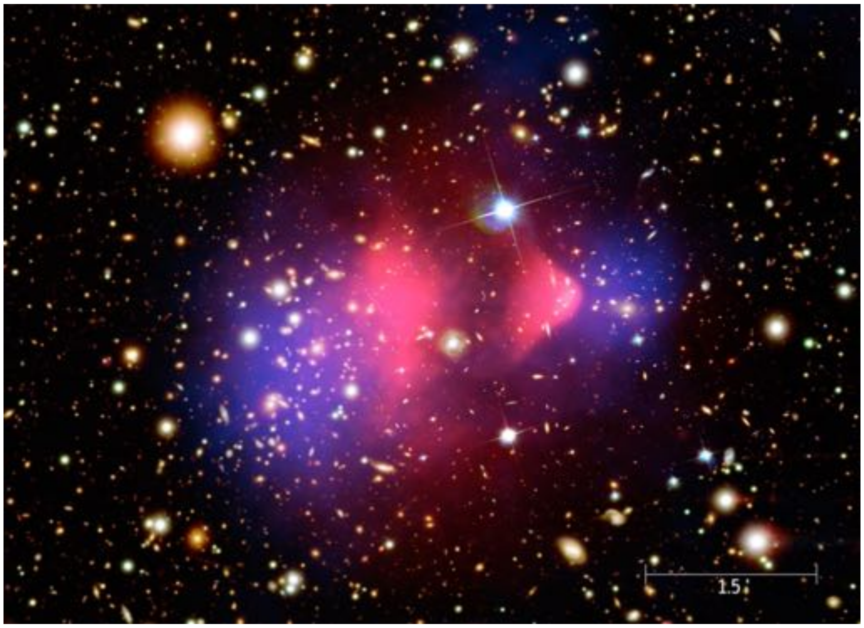


Figure 1.3: Image of the Bullet Cluster, depicting the collision of two galaxy clusters. The pink region is a Chandra X-ray Observatory image of the intracluster medium (ICM), which makes up the bulk of baryonic matter in the system, and the blue region is the actual distribution of mass calculated from gravitational lensing. Figure taken from the Chandra website [20].

topic abundances, and the accelerated expansion of the universe caused by Λ, the cosmological constant associated with dark energy [21]. The CMB in particular provides a precisely determined energy density for dark matter in the early universe. In the best fit of ΛCDM to the Planck observatory's CMB data, the matter fraction of the mass-energy density was determined to be $\Omega_m = 0.315 \pm 0.007$, with the dark matter comprising $84.4 \pm 0.2\%$ of all the matter and only the remaining sixth being baryons [22]. The remaining energy density is in the form of dark energy, with $\Omega_\Lambda = 0.6847 \pm 0.0073$ and small contributions from neutrinos and radiation.

The Cosmic Microwave Background Radiation

The CMB was produced in the primordial plasma of the Big Bang and reaches us from the edge of the observable universe. It is thus the oldest radiation we observe, and was released 400,000 years after the Big Bang when the universe was cool enough for photons to decouple from the electrons and baryons. The temperature of the CMB that we observe today corresponds to the

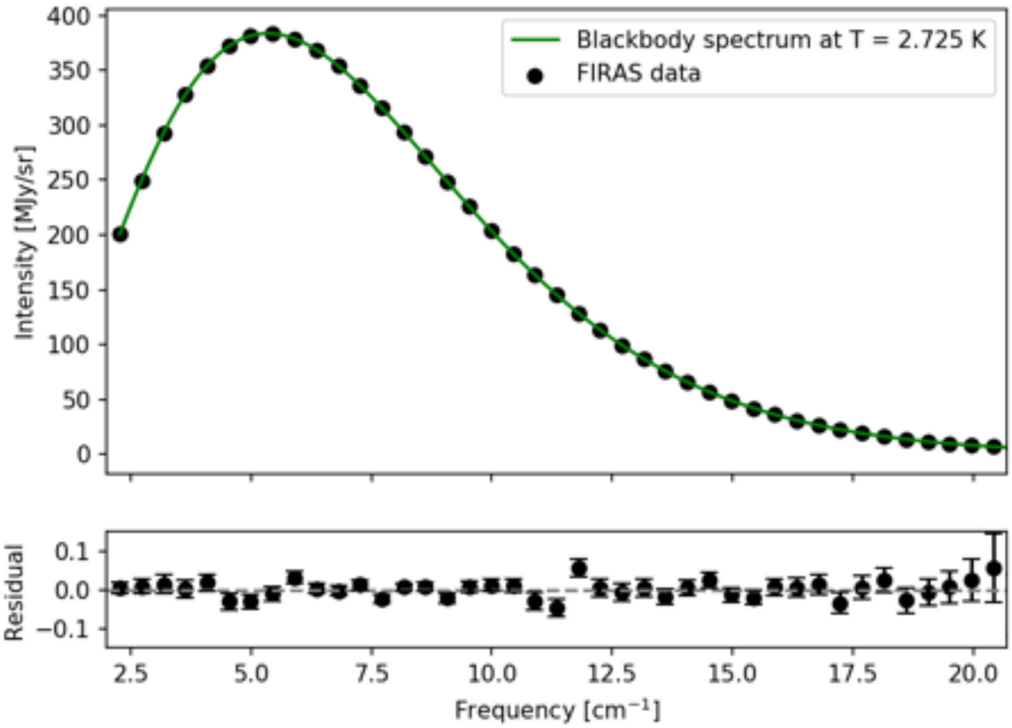


Figure 1.4: Top: Emission spectrum of the CMB determined by FIRAS on the COBE satellite (black points), shown with a theoretical black body spectrum at 2.725 K (green line). Bottom: The residual between the CMB spectrum and a perfect blackbody at 2.725 K. FIRAS had a large horn with a hole that let in the CMB radiation. The hole could be plugged with a microwave absorber, forming an isothermal cavity, which could be heated to obtain calibration data. FIRAS actually measured the residuals between the calibrated black body on the instrument and the CMB. Figure made by the author with data from Ref. [24].

primordial light that has been redshifted due to the expansion of space.

The thermal spectrum of the modern day CMB has been measured to be a near-perfect black body emission with a temperature of 2.726 ± 0.001 K and a peak spectral radiance at 160.23 GHz [23]. Figure 1.4 shows the data collected by the Far Infrared Absolute Spectrophotometer (FIRAS) flown on the Cosmic Background Explorer (COBE) satellite [24].

A host of ground based experiments, balloons, and satellites has provided detailed maps of the CMB at various angular resolutions. The highest resolution map of the CMB to date, made by the Planck collaboration is shown in Figure 1.5. The CMB is remarkably isotropic across the

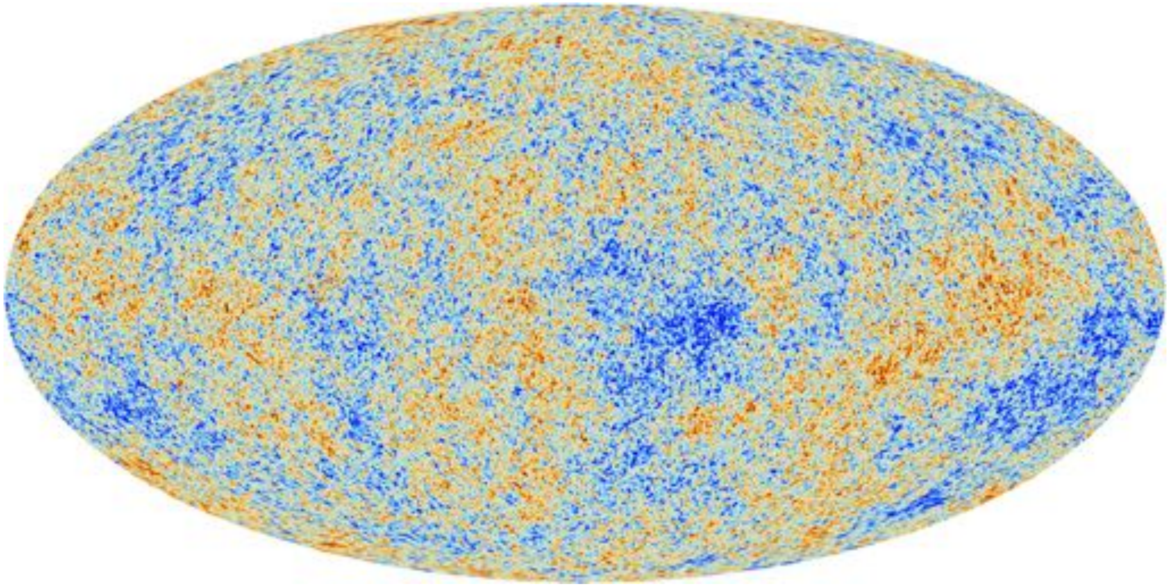


Figure 1.5: The temperature anisotropies of the CMB as observed by the Planck satellite. Figure taken from the European Space Agency website [25].

sky with deviations at the 10^{-5} level, revealing the slight density oscillations that existed at time of the photon–baryon decoupling. The temperature anisotropies on the map can be decomposed via a Fourier transform into a power spectrum, which encodes correlations of the temperature fluctuations as a function of their angular separations, and is shown in Figure 1.6. The angular scale of the fluctuations and the height of their peaks are sensitive probes of cosmological parameters such as the total energy density, dark matter fraction, and baryonic fraction for which values were given earlier.

Structure Formation

The structure today in the form of galaxies and clusters has evolved from the density fluctuations imprinted in the CMB. The origin of the density fluctuations is unclear, but cosmic inflation is known to govern the growth of the fluctuations in the early universe followed by hierarchical structure formation. Dark matter is an important component leading to the growth of structure in the late universe, and can be simulated in large N-body codes. The results of these simulations, which yield observables such as the matter power spectrum, can be compared with experiments

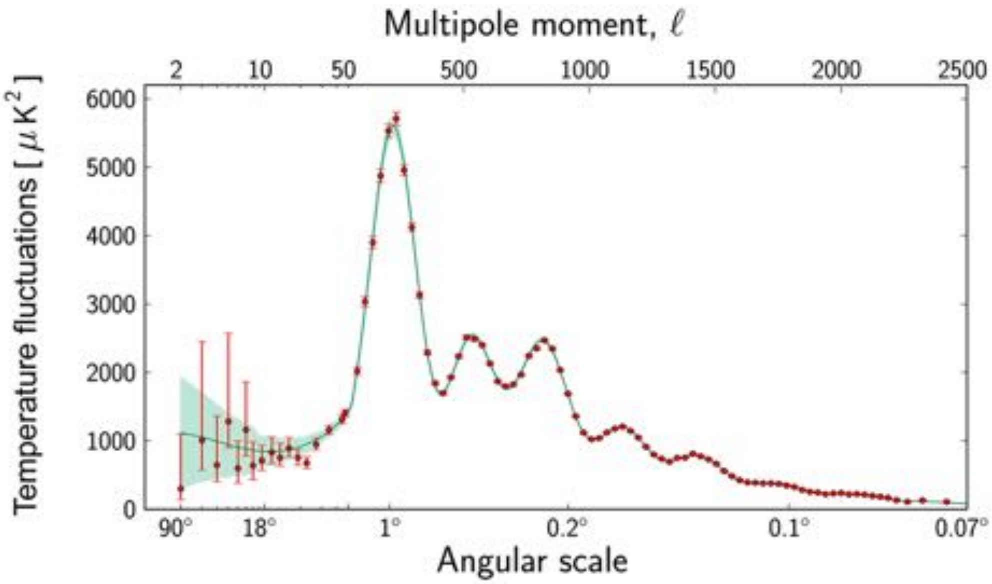


Figure 1.6: The power spectrum measured by the Planck observatory (red), showing the temperature fluctuations in the CMB at difference angular scales. The best fit of the standard model of cosmology to the data is shown in green. Figure taken from the European Space Agency website [26].

to constrain cosmological parameters such as the matter density.

Statistical measures of galaxies, clusters, and weak lensing are sensitive to the details of non-linear growth. Some state of the art cosmological simulations include effects such as baryonic pressures and the energetic processes of stars and black holes, which produce ejecta that are relevant at small distance scales. Figure 1.7 shows the result of the ILLUSTRIS N-body simulation, which makes detailed maps of not only the dark matter distribution, but that of the baryons as well [27]. Such simulations produce numbers, masses, and shapes of galaxies that are consistent with observations, indicating the necessity of dark matter for structure formation.

While the sketches above cover the range of diverse evidences for dark matter, it is not an exhaustive list. Please refer to Refs. [6, 9] for more complete accounts. We now turn to the WIMP, a leading candidate for the dark matter particle.

1.2 Weakly Interacting Massive Particles

Weakly Interacting Massive Particles (WIMPs) have been favored as candidates for dark matter due to a convergence of cosmology and particle physics, known as the WIMP miracle. Specifically, a neutral particle with mass around the electroweak scale (100 GeV) could have partially annihilated in the early universe to give a modern relic density consistent with dark matter. This idea was first promulgated by five independent papers published in a two month span in 1977, revealing the beginnings of an intense synergy between the fields of particle physics and cosmology [6]. Further work, which identified WIMPs as the constituents of the astronomers' missing mass, and as the neutral stable particle predicted by supersymmetric theories, cemented the status of the WIMP as the prime dark matter candidate in the decades that followed. A description of the cosmological history of the WIMP, following Ref. [21], is given here, followed by a discussion of WIMP scattering rates.

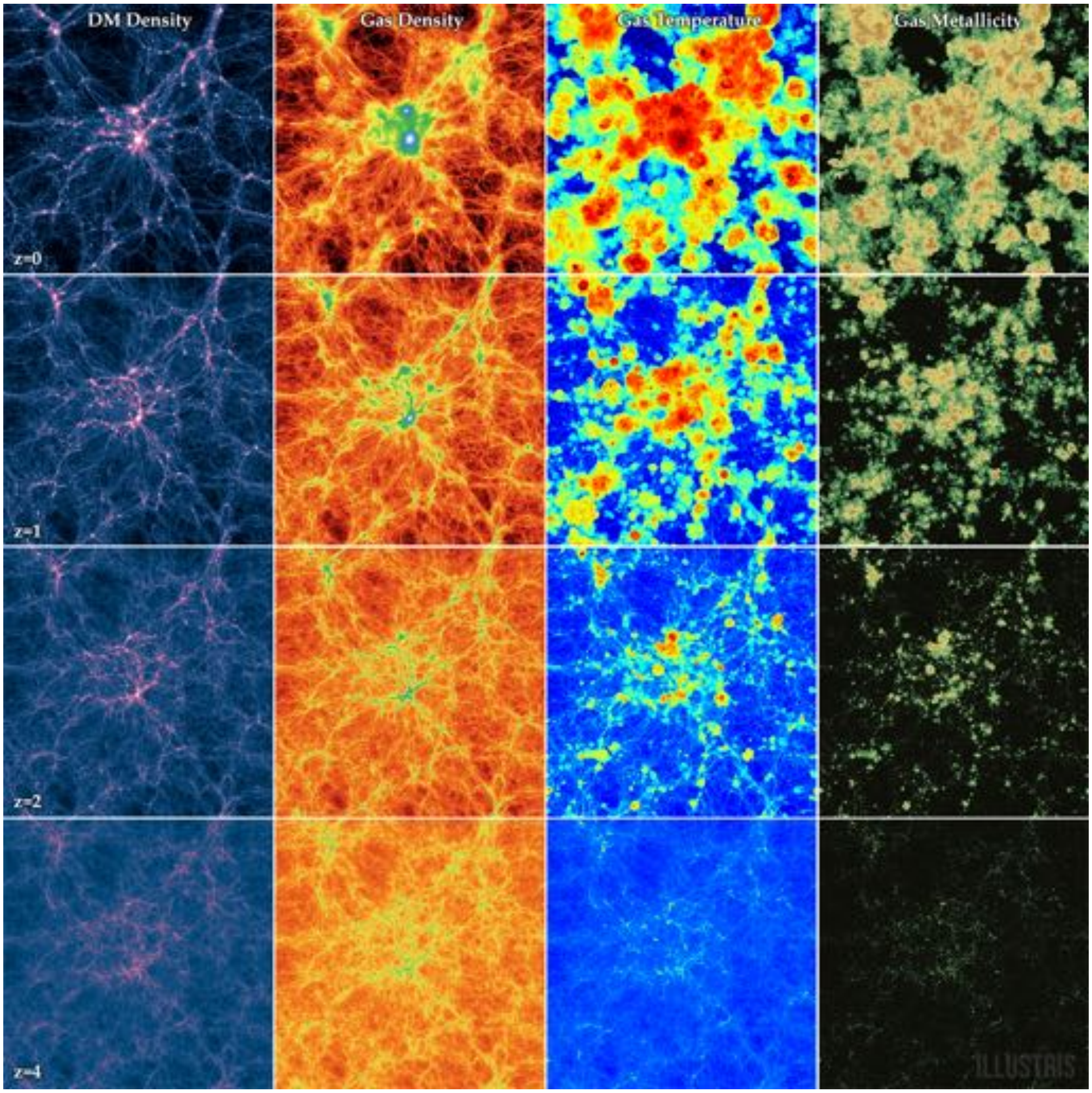


Figure 1.7: The ILLUSTRIS N-body simulation, which tracks the formation of galaxies in time with a focus on retaining non-gravitational effects. The first column shows the density of dark matter, the second column the correlated density of gas, the third column the temperature of the gas, and the fourth column the metallicity of the gas. From the top, the four rows correspond to redshift z of 0 (present day), 1, 2, and 4 (corresponding to 1.6 Gyr since the Big Bang), respectively. Figure taken from the Illustris Collaboration [27].

1.2.1 Cosmology

The abundance of WIMPs in the early universe can be modeled in time by the homogeneous Boltzmann equation in an expanding spacetime, under the assumptions of kinetic equilibrium and the thermodynamics of a classical gas. The Boltzmann equation governing the WIMP annihilation/production interaction $\chi + \chi \longleftrightarrow \psi + \psi$, where ψ stands for SM particles and χ is the dark matter, is given by

$$a^{-3}\frac{d(n_\chi a^3)}{dt} = \left(n_\chi^{(0)}\right)^2 \langle\sigma v\rangle \left[\frac{n_\psi^2}{\left(n_\psi^{(0)}\right)^2} - \frac{n_\chi^2}{\left(n_\chi^{(0)}\right)^2}\right], \tag{1.3}$$

where a is the scale factor, $\langle\sigma v\rangle$ is the thermally averaged cross section, n_χ and n_ψ are the WIMP and SM number densities, and where the superscripted number densities are species independent, i.e. with no differences in the chemical potential. The number densities are defined for a classical (neither Bose-Einstein nor Fermi-Dirac) species s as

$$n_s = g_s e^{\mu_s/T} \int \frac{d^3p}{(2\pi)^3} e^{-E_s(p)/T}, \tag{1.4}$$

and the species independent number density as

$$n_s^{(0)} = g_s \int \frac{d^3p}{(2\pi)^3} e^{-E_s(p)/T}, \tag{1.5}$$

where g_s is the degeneracy factor of the species, μ_s is the chemical potential, E_s is the phase-space energy, T is the temperature of the bath, and where the momentum integral is taken over the phase-space distribution $f(E) \propto e^{-(E-\mu)/T}$, with the chemical potential set to zero for the species independent case. Note that the system of units employed here and throughout the thesis sets the speed of light c, the reduced Planck's constant $\hbar$, and Boltzmann's constant k to unity.

In the early universe WIMPs are in equilibrium with the rest of the particles, but then freeze out when the Hubble expansion rate exceeds the WIMP annihilation rate. Since all the SM particles (except the neutrinos) are tightly coupled to each other, they are assumed to be in chemical equilibrium $n_\psi = n_\psi^{(0)}$ and Eq. 1.3 simplifies to

$$a^{-3}\frac{d(n_\chi a^3)}{dt} = \langle \sigma v \rangle \left[\left(n_\chi^{(0)} \right)^2 - n_\chi^2 \right]. \tag{1.6}$$

Changing variables from n_χ to $Y = n_\chi / T^3$, and the introduction of a time variable $x = m_\chi / T$ where m_χ is the WIMP mass gives

$$\frac{dY}{dx} = -\frac{\lambda}{x^2}\left[Y^2 - Y_{\mathrm{EQ}}^2 \right], \tag{1.7}$$

where $\lambda \propto \langle \sigma v \rangle m_\chi^3 / H$ is a ratio of the WIMP–SM coupling to the expansion rate. Numerical solutions for Eq. 1.7 are shown in Figure 1.8 for two different values for λ. In the limit of either slow cosmic expansion or large cross section ($\lambda \to \infty$) the WIMPs do not freeze out and will have Boltzmann suppressed number densities $n_\chi \propto e^{-m_\chi/T}$ until the modern day, essentially resulting in no dark matter. However, associating the observations of dark matter with a relic from freeze out, Eq. 1.7 may be used to calculate the strength of WIMP–SM interaction. An estimate using the modern abundance of $\Omega_\chi h^2 \approx 0.1$ for a 100 GeV WIMP yields a thermally averaged cross section of $\langle \sigma v \rangle \approx 10^{-26}$ cm^3/s, which is typical for electroweak interactions, hence the ‘miracle’ [21].

The WIMP hypothesis faces challenges from the lack of experimental evidence, most recently from the LZ experiment which probed cross sections as low as $\sigma = 9.2 \times 10^{-48}$ cm^2 for a 36 GeV WIMP [5] and also indirectly from the absence of evidence for supersymmetric particles at the Large Hadron Collider (LHC) [28]. Further, the strength of the WIMP hypothesis has always been weakened by the fact that the required cross section depends on m_χ, for which no strong priors exist [4]. Despite the connection between weak-scale physics and dark matter being put into question, the traditional WIMP remains a viable candidate for reasons to do with cosmology and particle phenomenology. For example, detailed model-independent calculations

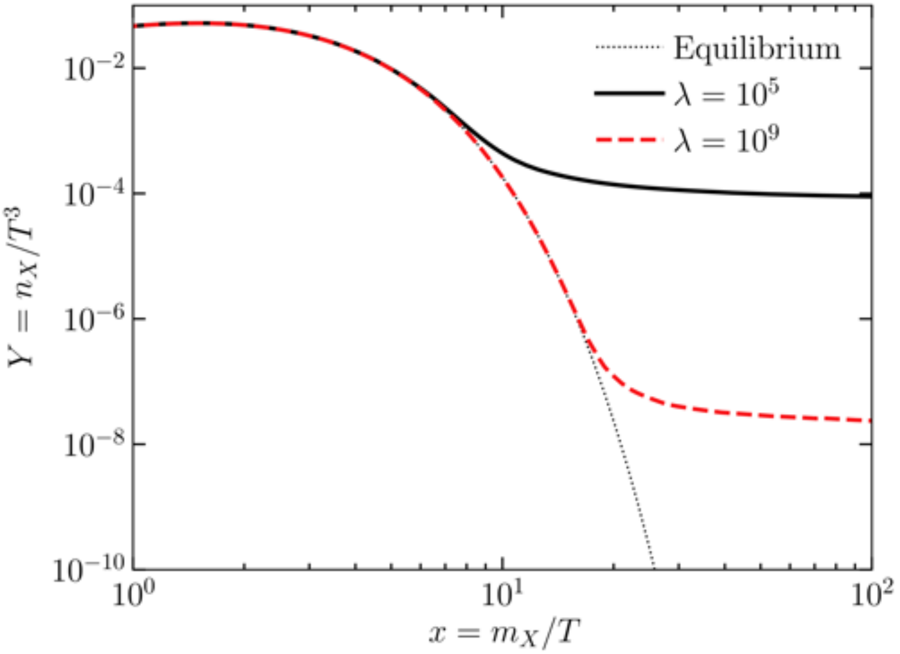


Figure 1.8: Abundance of a massive neutral particle as a function of inverse temperature. The equilibrium abundance is shown by the dotted line, and the red dashed and black lines show different values of λ, the ratio of the annihilation cross section to the Hubble expansion rate. Figure from Ref. [21].

of the thermally averaged cross section found that including a complete set of WIMP annihilation channels weakens the claim that WIMPs are nearing total exclusion [29]. On the particle physics side, the lack of WIMP observations may be due to interference effects, or momentum suppressions that diminish the WIMP–nucleon cross sections. The main subject of this thesis is a search for general WIMP–nucleon interactions that take into account degrees of freedom that may suppress the cross section.

1.2.2 Scattering rates in terrestrial detectors

The annihilation of WIMPs to SM particles in the freeze out scenario indicates a non-zero coupling between WIMPs and ordinary matter. This assumption has motivated three types of experimental searches. Firstly, the annihilation processes $\chi + \chi \to \text{SM}$ that froze the WIMPs in the early universe could still occur today in strong gravity wells where dark matter coalesces, like the centers of stars and galaxies. The detection of anomalous cosmic rays produced in such annihilation reactions may provide evidence for dark matter. Secondly, the reverse reaction $\psi + \psi \to \text{DM}$ may be achieved in colliders like the LHC, where missing momenta due to the WIMP escaping detection will provide evidence of production. Thirdly, direct detection experiments such as LZ look for the keV-scale recoils of galactic WIMPs scattering with nuclei in detectors. The scattering rate of WIMPs in terrestrial detectors is presented here.

The rate of WIMP interactions qualitatively depends on the number density of the WIMP halo, the relative WIMP speed, and the WIMP–SM interaction cross section. The differential WIMP event rate exhibits an exponential fall off in energy that is due to the WIMP velocity distribution. A complete accounting of the rate requires the inclusion of the WIMP velocity profile, nuclear form factors, and detector energy thresholds, and is given by

$$dR = \frac{\rho_0}{m_\chi m_N} v f(v) \frac{d\sigma}{d|q|^2} d|q|^2 dv, \tag{1.8}$$

where ρ_0 is the WIMP mass density, typically taken to be 0.3 GeV/cm^3 with close to a 100%

uncertainty [30], m_χ and m_N are the masses of the WIMP and nucleon respectively, v is the WIMP velocity, $f(v)$ is the velocity distribution of galactic WIMPs, σ is the WIMP-nucleon cross section, and q is the momentum transferred in the interaction [31].

Cross sections and form factors

WIMP–SM interactions would arise from the fundamental WIMP–quark and WIMP–gluon interactions, which result in scalar and spin-dependent cross sections, both tested in direct detection experiments like LZ. For neutralinos, which are stable neutral fermions predicted by the Minimal Supersymmetric Standard Model (MSSM) that serve as a benchmark WIMP, the scalar cross section is the larger and is calculated from squark and Higgs exchanges, while the spin-dependent cross section arises from the axial-vector interaction [31]. The general expression for the differential WIMP-nucleon cross section is

$$\frac{d\sigma}{d|\vec{q}|^2} = G_F^2 \frac{C}{v^2} F^2(|\vec{q}|^2), \tag{1.9}$$

where G_F is the Fermi constant, C is a complex number containing the particle physics content of the interaction, and F is the nuclear form factor that carries momentum dependence, also a dependence on the interaction type.

The complex number for the spin-independent interaction is

$$C_{\mathrm{SI}} = \frac{1}{\pi G_F^2} \left[Z f_p + (A-Z) f_n\right]^2, \tag{1.10}$$

where Z and A are the atomic and mass numbers, and f_p and f_n are the effective WIMP–proton and WIMP–neutron coupling strengths. Since isospin is a good symmetry of nuclear ground states, the assumption $f_p \approx f_n$ is reasonable, leading to an A^2 enhancement of the cross section that favors heavy nuclei like xenon. The form factor for the scalar case is the Fourier transform of the nucleon densities in the nucleus, and behaves like $F^2(E) \propto e^{-E/2E_0}$, where E is the energy transferred to nucleus, and the nuclear coherence energy E_0 is approximately $80/A^{5/3}$ MeV for xenon [32]. This exponential form factor corresponds to a purely Gaussian nucleon density. A more accurate form factor is

$$F(q) = \left[\frac{3j_1(qR_1)}{qR_1}\right]^2 e^{-(qs)^2}, \tag{1.11}$$

where $R_1 \approx 6$ fm for xenon, $s \approx 1$ fm, and where j_1 is the spherical Bessel function [33].

The c-number for the spin-dependent interaction is

$$C_{\text{SD}} = \frac{8}{\pi}\frac{J+1}{J}\left[a_p\langle S_p\rangle + a_n\langle S_n\rangle\right]^2, \tag{1.12}$$

where J is the total angular momentum of the nucleus, a_p and a_n are the WIMP–proton and WIMP–neutron spin couplings, and the bracketed quantities are the expectation values for the spin content of protons and neutrons in the nucleus. The form factor in this case is more complicated, requiring a description of how nucleon spins are distributed in the nucleus. To first order in xenon, the bulk of the spin coupling is with the unpaired neutron in ^{129}Xe and ^{131}Xe, but proton couplings are also possible due to mixing between nucleon spin states [33]. There are significant theoretical uncertainties in the spin dependent form factors arising from independent calculations [34–36]. To wrap up the discussion of interaction cross sections, note that the spin-dependent interaction does not benefit from A^2 enhancement like the scalar case.

WIMP halo

The WIMP velocity distribution is a key component of the rate and deposited energy spectrum. Galactic WIMPs are non-relativistic, with typical speeds of 0.1% the speed of light in the rest frame of Earth. For xenon nuclei elastically scattering with WIMPs of a similar mass (100 GeV), the nuclear recoil energy is then of order

$$\frac{1}{2}m_{\text{Xe}}v_{\text{WIMP}}^2 = \frac{1}{2} \times 100\text{GeV} \times (0.001)^2 = 50\,\text{keV}, \tag{1.13}$$

which is entirely determined by the kinematics. Since WIMPs behave like a dilute classical gas, assuming sufficient thermalization in the gravitational field of the Milky Way, their velocity in the galactic frame $\vec{v}_g$ can be expressed by the Maxwellian distribution, also known as the Standard Halo Model (SHM)

$$f(\vec{v}_g)d^3v_g = \frac{1}{v_0^3\pi^{3/2}}e^{-v_g^2/v_0^2}d^3v_g, \tag{1.14}$$

where v_0 is the velocity of the local standard of rest at the sun, and sets the size of the dispersion. The recommended value for v_0 is 238.0 ± 1.5 km/s [37]. The galactic WIMP velocity is related to the lab frame velocity $\vec{v}_l$ by

$$\vec{v}_g = \vec{v}_l + \vec{v}_0 + \vec{v}_\odot + \vec{v}_\oplus(t), \tag{1.15}$$

where $\vec{v}_\odot$ is the sun's peculiar velocity, and $\vec{v}_\oplus$ is Earth's velocity with respect to the sun. Each component of $\vec{v}_\odot$ is on the order of 10 km/s, while the average speed of the Earth in the sun's frame is roughly 30 km/s and the time dependence may be neglected for analyses that do not look for an annual modulation in the WIMP rate. Ignoring the form factor, which for light WIMPs is a constant of order unity, and the corrections to the circular velocity $\vec{v}_0$, integrating the Maxwellian distribution as in Eq. 1.8 from 0 to ∞ produces the exponential fall of the energy spectrum mentioned earlier.

There are two corrections made to the picture above. The distribution in Eq. 1.14 is only valid below the escape velocity of the Milky Way v_{esc}, implying a cut off in the velocity integral of Eq. 1.8. The recommended value for the escape velocity is 544 km/s [37]. There is also a threshold WIMP velocity v_{min} below which a scatter producing nuclear recoils with energy E is impossible. Since the energy deposited by the WIMP is $E = |\vec{q}|^2/(2m_N) = (1-\cos\theta)m_r^2v^2/m_N$ where $m_r = m_\chi m_N/(m_\chi + m_N)$ is the reduced mass, the minimum velocity required to deposit energy E is $v_{\text{min}} = \sqrt{Em_N/(2m_r^2)}$. Thus the velocity integral is bounded by v_{min} and v_{esc}.

The dark matter halo is almost certainly anisotropic due to the formation history of the Milky Way, and Earth might be in the vicinity of substructures that would affect the overall WIMP rates in direct detection experiments [38, 39]. These considerations have led to updated halo models

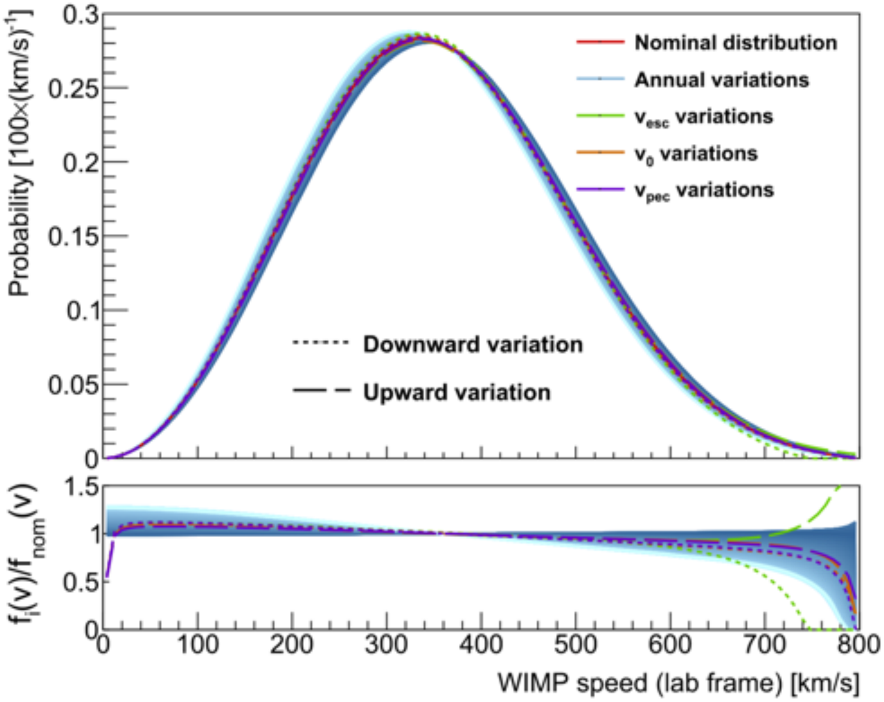


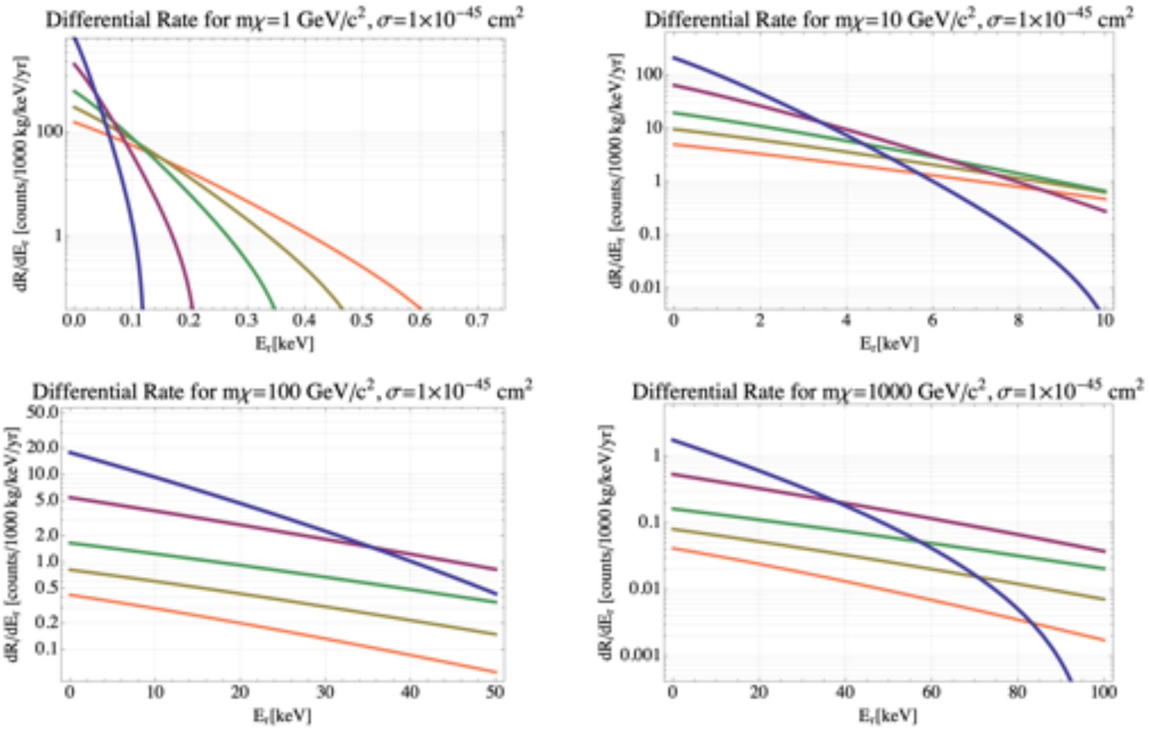
Figure 1.9: The Maxwellian WIMP velocity distribution shown for varying parameter values, from Ref. [37]. The red curve shows the distribution with recommended average values of the astrophysical parameters, and the various shades of blue indicate the changes in the shape due to annual modulation. Green, brown, and purple lines show the changes due to variations in the escape, circular, and peculiar velocities, respectively. The bottom plot shows the ratio of the variations to the recommended model.

and direct detection analyses in which the possibility of dark matter streams are accounted for [40]. While experiments are still in discovery mode, however, the Maxwellian velocity distribution is recommended for a common point of comparison. Standard values for the velocities are provided in Ref. [37]. Figure 1.9 shows the analytical distribution of WIMP velocities, along with the effects of varying certain parameters.

To wrap up the discussion of the scattering rate, Eq. 1.8 can be written as

$$\frac{dR}{dE} = \frac{\rho_0}{m_\chi m_N} \int_{v_{\text{min}}}^{v_{\text{esc}}} v f(v) \frac{d\sigma}{dE} dv, \tag{1.16}$$

where for generality, the cross section is allowed to be velocity dependent to account for interactions that couple the WIMP velocity to spin and momentum as elaborated in Chapter 3. Figure 1.10 shows the differential event rate for WIMPs scattering in various target materials for four WIMP masses [41]. The spectra for light WIMPs are suppressed at large recoil energies, highlighting the importance of a low energy threshold. In the following chapter, we present a discussion of how the design of direct detection instruments like the LZ detector makes them effective tools for WIMP discovery.



Differential Rate for $m\chi=1$ GeV/c^2, $\sigma=1\times10^{-45}$ cm^2

Differential Rate for $m\chi=10$ GeV/c^2, $\sigma=1\times10^{-45}$ cm^2

Differential Rate for $m\chi=100$ GeV/c^2, $\sigma=1\times10^{-45}$ cm^2

Differential Rate for $m\chi=1000$ GeV/c^2, $\sigma=1\times10^{-45}$ cm^2

Figure 1.10: Differential event rates for a 1 GeV (upper left), 10 GeV (upper right), 100 GeV (lower left), and 1,000 GeV (lower right) WIMP interacting with a cross section of 10^{-45} cm^2 with xenon (blue), germanium (purple), argon (green), silicon (brown) and neon (orange) targets. Figure from Ref. [41].

Chapter 2

The LUX-ZEPLIN Dark Matter Experiment

> *But it is one thing to read about dragons and another to meet them.*
>
> Ursula K. Le Guin, A Wizard of Earthsea

Shedding light on the nature of dark matter is one of the primary goals of the particle physics community in the next decade. In their decennial report on particle dark matter, the conveners of the Snowmass Community Planning Exercise recommended a strategy that delves deep, concentrating on high priority targets such as WIMPs, while searching wide, in order to explore as much parameter space available to dark matter as possible [42]. Direct detection experiments are well suited to discover dark matter due to their ability to cover a wide range of masses, and in some theoretical models, prove to be the only near-term experiments capable of discovery [43]. Figure 2.1 shows the landscape of theories that are refutable by direct detection. Further, experiments that use xenon such as LZ are sensitive to a wide range of WIMP interaction types, as evidenced in the main result of this thesis in Chapter 3.

The most promising and well-studied range of dark matter masses is between 1 GeV and

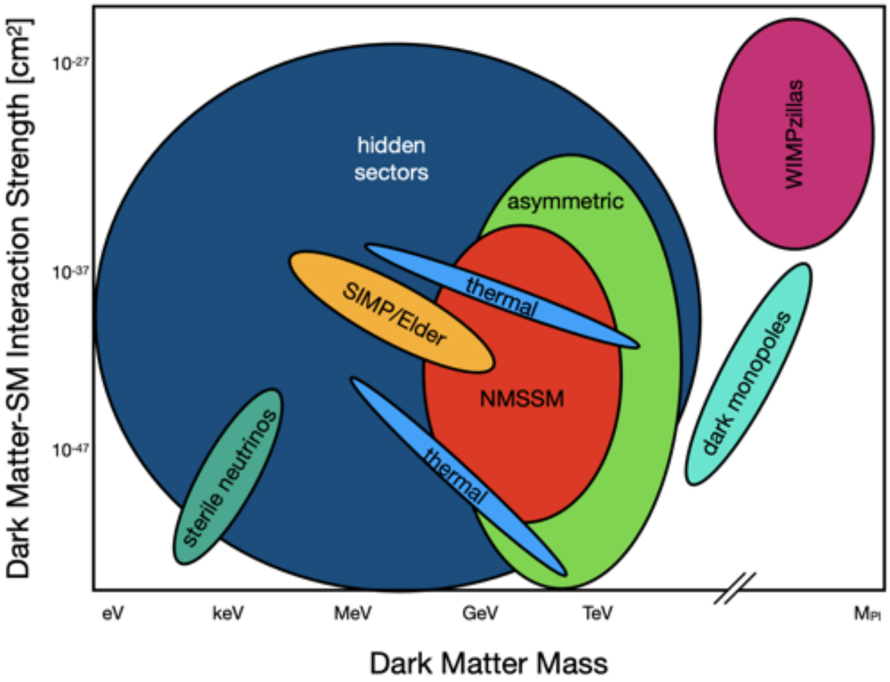


Figure 2.1: A cartoon of dark matter candidates shown in the space of interaction strength versus mass, from Ref. [42]. Among the models shown are various thermal candidates, supersymmetric candidates, asymmetric candidates [44], sterile neutrinos [45], dark monopoles [46], and hidden sector dark matter [47].

100 TeV, which includes WIMPs [42]. The basic goal of a WIMP detector is to record the rare interactions between a galactic WIMP and a target nucleus that causes the nucleus to recoil with small energies (1–100 keV). Ideally, these nuclear recoils (NRs) should be separated from background events, which are data that do not correspond to the dark matter signal. Typical backgrounds stem from electronic recoils (ERs) caused by β and γ particles from the environment interacting with the atomic electrons of the detector material, in addition to neutron interactions producing NR signals that mimic WIMPs. An excess number of NR events that fits the theoretical WIMP energy spectrum may be taken as evidence for WIMP dark matter and will finally begin to demystify their particulate nature.

2.1 Towards dark matter direct detection in Pahá Sápa

The home of the LZ experiment is the Black Hills (Lakota: Pahá Sápa) of South Dakota, a location sacred to Native Americans, and later the location of the Homestake mine, once the deepest and most productive gold mine in North America. Today, the Sanford Underground Research Facility (SURF) houses several underground experiments, including LZ. SURF is also undergoing preparations for the future Deep Underground Neutrino Experiment (DUNE). This location has a scientific pedigree, having housed the Large Underground Xenon (LUX) dark matter experiment that concluded underground operations in 2016, and going even further back to the 1960s, Ray Davis's experiment that recorded the solar neutrino deficit [48, 49]. This section describes the LZ detector along with its specializations for WIMP discovery, and some of the author's contributions to its commissioning.

2.1.1 LZ design drivers

The LZ experiment's goal requires a highly specialized instrument designed to address the challenges associated with recording a handful of dark matter events per year in a radioactive en-

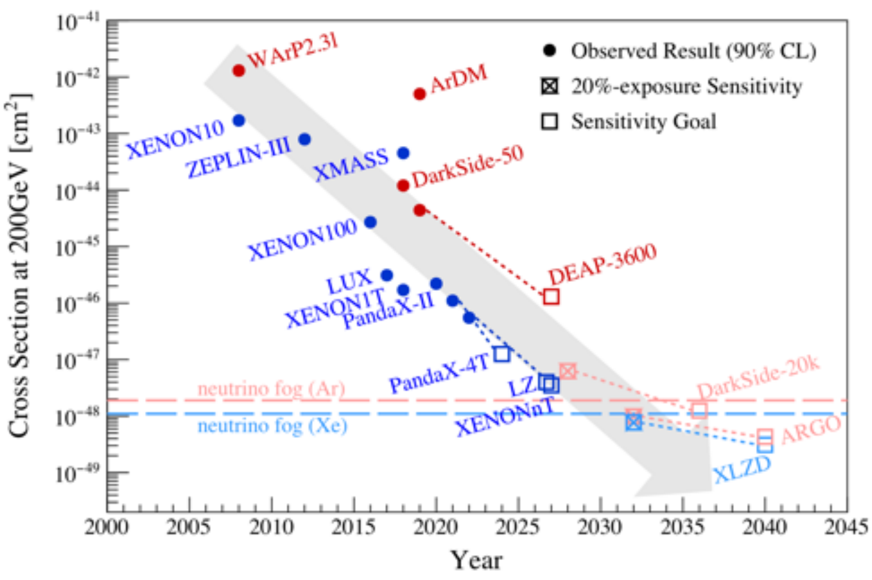


Figure 2.2: Spin independent WIMP–nucleon couplings for a 200 GeV WIMP, for LXe (blue) and liquid argon (red) experiments. The filled circles indicate measured results, while the squares indicate expectations for the final exposure (open) and 20% of the final exposure (crossed). Figure taken from Ref. [42].

vironment. The adopted solution to these challenges is a dual-phase liquid xenon (LXe) time projection chamber (TPC), operated with a system of outer detectors to veto multiply scattering backgrounds. The TPC is the heart of the LZ experiment, where the scintillations and ionizations in the LXe produced by particles scattering are recorded, enabling energy and 3D position reconstruction, along with the discrimination of ER and NR interactions.

LXe TPCs have proved to be a premier technology in the search for WIMPs, excluding the smallest WIMP–SM interaction cross sections, as shown in Figure 2.2. The following points describe the design drivers of LZ, while simultaneously motivating the use of LXe as a detector medium. A more complete description of the TPC and veto detectors is given afterwards.

Underground siting to reduce cosmic ray flux

Cosmic rays are a serious background source to experiments searching for rare events of astrophysical origin. Cosmic rays interact with atoms upon entering the Earth's atmosphere and

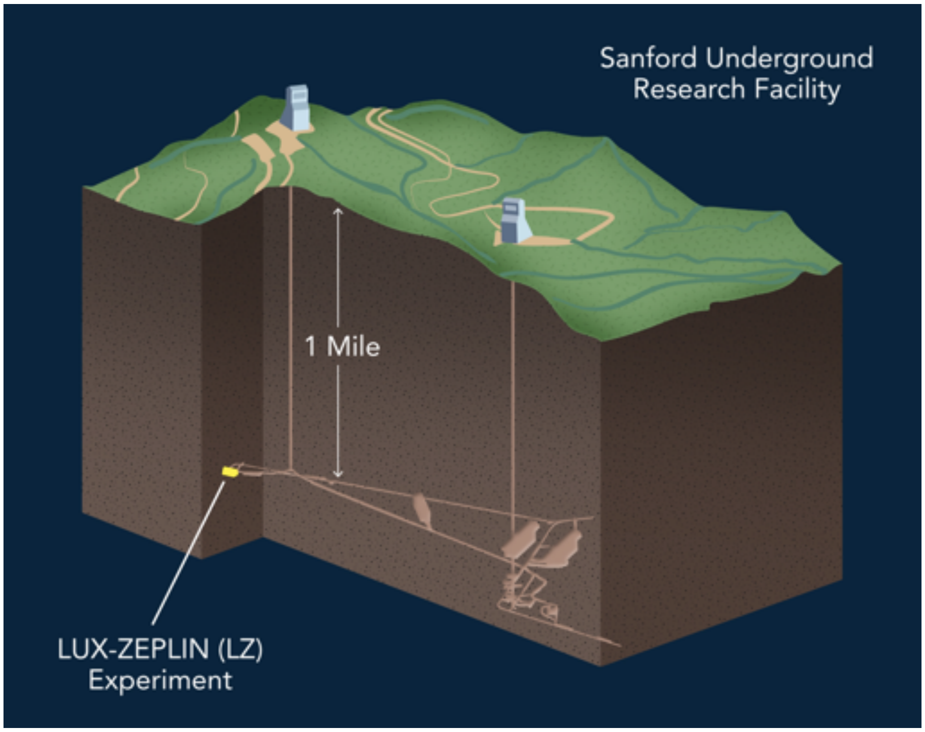


Figure 2.3: Network of shafts and tunnels operated by SURF, with the location of the LZ experiment indicated in yellow. The caverns on the right are currently being excavated for the Deep Underground Neutrino Experiment (DUNE) [51]. Picture from the SURF website [52].

create cascades of radiation that include protons, neutrons, α particles, pions, muons, and neutrinos. Most of these particles are stopped by the atmosphere except neutrinos, which pass through the Earth relatively unimpeded, constituting an irreducible background, and muons, which can penetrate into rock. Muon interactions in the LZ detector produce large scintillation signals that are readily identifiable, but muon-induced hadronic showers in the Davis cavern produce neutrons whose interactions with LXe are indistinguishable from a WIMP's.

Like the solar neutrino experiment and LUX, the LZ experiment is located at the 4850 ft (1.48 km) level of the Sanford Underground Research Facility (SURF), formerly the Homestake Mine, where the rock overburden of 4,300 m of water equivalent reduces the muon flux by a factor of 3×10^6 relative to the surface [50]. Figure 2.3 shows a schematic of the LZ experiment's location in the mine.

Large target mass to detect rare interactions

A target that is as large and dense as possible is necessary to maximize chances of a WIMP scattering in it. The LZ TPC is filled with 7 tonnes of LXe, a liquid with a density of 2,940 kg/m^3 at its boiling point, roughly three times the density of water [53]. The latest detectors of two other LXe programs, XENONnT and PandaX-4T, also exhibit target masses on the few-tonne scale, and represent a step up in sensitivity from previous experiments like XENON1T and LUX [54, 55]. Accordingly, future liquid xenon dark matter detectors are planned to be even larger. For example, the XLZD consortium, the union of three leading collaborations (LZ, XENON, and DARWIN), is planning for a TPC that could contain upwards of 40 tonnes of LXe [56].

Self-shielding and purification to reduce backgrounds

The outer regions of the LXe shield the innermost volume from radioactive particles emanated from the walls of the detector. Thus it is beneficial to scale up monolithic detectors rather than use several modular TPCs to make up the target mass. Figure 2.4 shows the exponential decrease in the rates of ER backgrounds published by LXe TPC experiments of increasing size over the years. To limit this external radiation, materials with extremely low intrinsic radioactivity were used to build LZ, following an extensive radioassay campaign [57].

For radioactive impurities that are dissolved in the LXe, purification of the LXe before the TPC is filled and during operation is crucial to ensure low backgrounds and adequate electron mobility within the LXe bulk. The xenon in LZ underwent a chromatographic process at the SLAC National Accelerator Laboratory to remove the β-emitter ^{85}Kr before being condensed in LZ. Non-noble impurities in the LXe are continuously removed during operations in LZ by flowing the xenon through a heated zirconium getter, using a circulation system that turns over the total LXe mass in 2.5 days.

Finally, natural xenon has no intrisically radioactive isotopes, except for the extremely long lived ^{124}Xe (2×10^{22} years) and ^{136}Xe (2×10^{21} years), making it an intrinsically radioclean

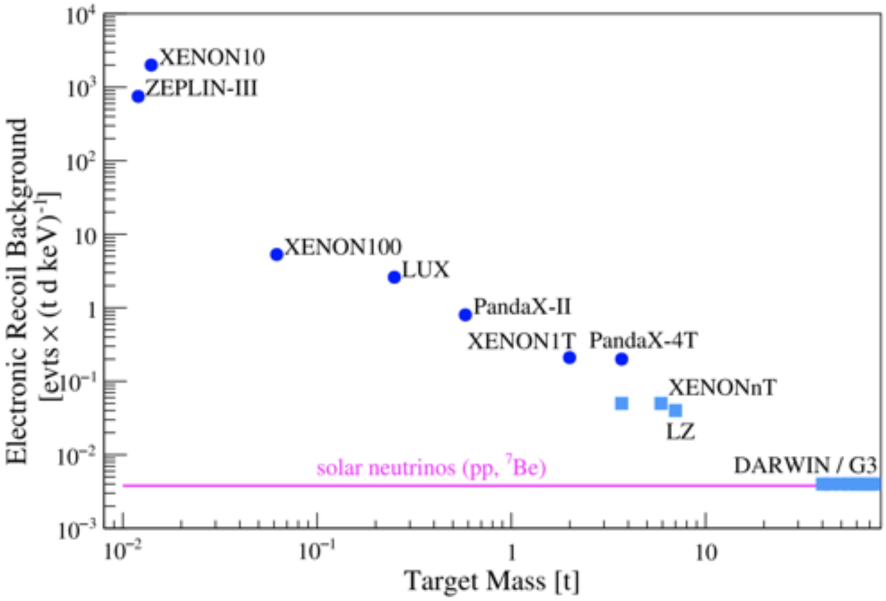


Figure 2.4: ER background rates in LXe TPCs as a function of target mass before discrimination cuts. Circles indicate best achieved limits, and squares indicate expectations, with the long shaded region showing the expectation for the future XLZD detector. Figure taken from Ref. [56].

target.

Low energy thresholds to record small energy deposits

A detector with a low energy threshold of around 1 keV is necessary to accept the low energy events produced by the elastic scattering of a WIMP off a nuclei, and especially for the detection of light (sub-GeV) WIMPs. LXe as a detector material in TPCs has a demonstrated sensitivity to NR events with energies below 1 keV using calibrated neutron sources [58, 59]. Efforts are continuing to push the measurements of LXe response to below 0.3 keV, such as the author's simulation work on neutron capture elaborated in Appendix B and published in Ref. [60].

Discrimination to isolate nuclear recoils

The signals detected by LXe TPCs allow NR events to be effectively separated from ER events, providing a nearly background-free search region. This discrimination is afforded by differences in channels through which energy is dissipated in the LXe when a xenon atom is struck; proportionally less ionization occurs in an NR compared to an ER for the same amount of scintillation. The physics of signal production in LXe is the subject of Section 2.3.

Isotopic diversity for unknown WIMP interactions

In addition to a high sensitivity for spin-independent (SI) WIMP–nucleon interactions due to the A^2 enhancement, the presence of ^{129}Xe and ^{131}Xe, isotopes with unpaired neutrons, provide good spin-dependent (SD) sensitivity as well. For extending the search to interactions beyond the standard SI and SD WIMP-nucleon couplings, as is the topic of this thesis (elaborated in Chapter 3), natural xenon provides a good nuclear laboratory that is sensitive to six distinct nuclear responses that could arise from general WIMP-nucleon interactions.

2.2 Detector description

A rendering of the LZ detector as it sits in the Davis cavern is shown in Figure 2.5. In the center is the TPC, a cylindrical chamber containing around 7 tonnes of LXe with reflective polytetrafluoroethylene (PTFE) walls and two arrays of light detecting photomultiplier tubes (PMTs) capping each end. Outside the TPC, but within a titanium cryostat is a layer of LXe (approximately 2 tonnes) called the 'xenon skin' that is optically separated from the TPC and is instrumented with its own set of PMTs. A second cryostat surrounds the inner cryostat to create a thermally insulating vacuum jacket.

A set of acrylic tanks containing gadolinium-loaded liquid scintillator (Gd-LS) hermetically surrounds the cryostats, forming one component of the outer detector (OD). The entire assembly

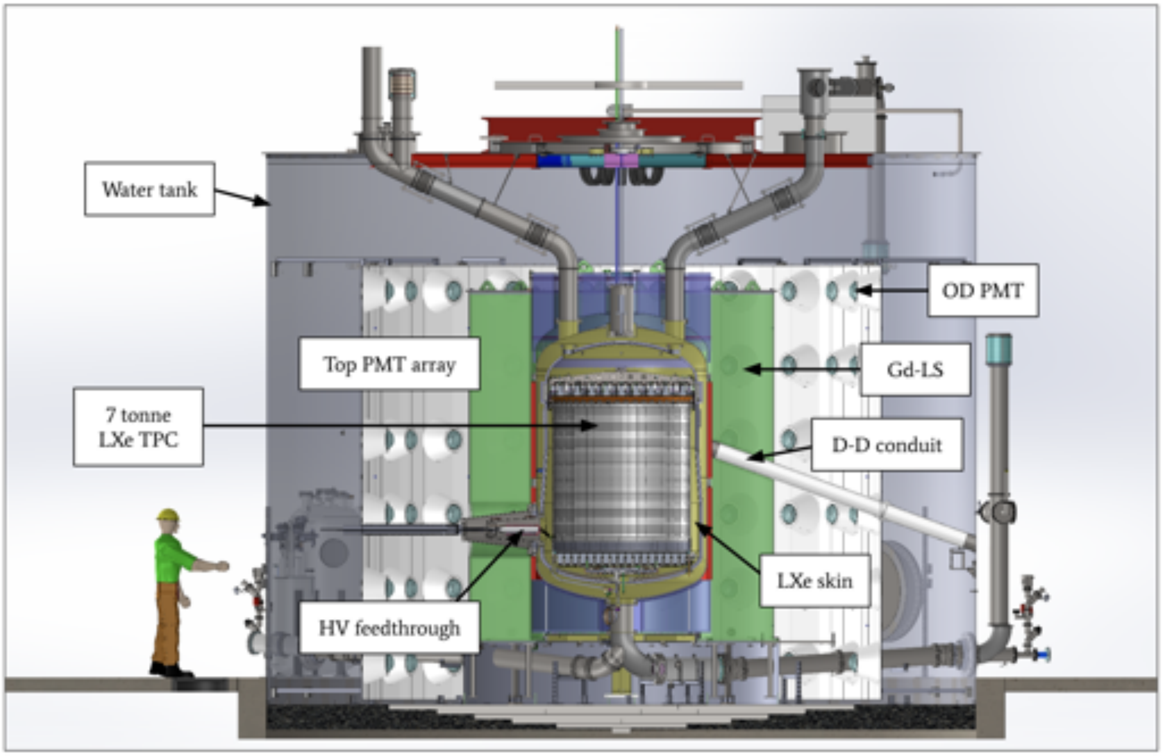


Figure 2.5: Schematic layout of the LZ detector showing the TPC, the xenon ‘skin’ around the TPC, the TPC and OD PMTs, the tanks containing gadolinium-loaded liquid scintillator (Gd-LS) in green and purple, and conduits supplying high voltage and the neutron calibration source.

is in a water tank, so there is a layer of ultra-pure water between the acrylic tanks and a final layer of PMTs, which make up the rest of the OD. Scintillation light produced in the Gd-LS in addition to Cherenkov radiation in the water are recorded by the OD PMTs.

Various conduits and feedthroughs carry electrical connections to the PMTs and high voltage (HV) for the grid electrodes, circulate the xenon through the purification system (located outside the water tank), maintain the insulating vacuum in between the cryostats, and allow external calibration sources such as the deuterium-deuterium (D-D) neutron source to penetrate to the TPC.

What follows below are brief descriptions of selected components in LZ, chosen for relevance. For full accounts of the design and performance of LZ instruments, refer to the Technical Design Report [50], the Conceptual Design Report [61], the LZ NIM paper [62], the grids design paper [63], the radioassay papers [57, 64–66], the cavern background measurement [67], the optical calibration system paper [68], and the PMT papers [69, 70].

2.2.1 Time projection chamber

Upon interaction with a particle, LXe dissipates the deposited energy in the form of scintillation light, ionization electrons, and heat as atomic motion. The design of the TPC enables the collection of the scintillation light and electrons. The scintillation light is recorded immediately by the TPC PMTs, and is known as the S1 signal. The electrons are drifted upwards using electric fields from the interaction site to the surface of the LXe and extracted into the xenon gas (GXe), where their passage through the gas creates secondary scintillation light in a process known as electroluminescence [1]. The electroluminescence process typically produces around 1,000 photons/cm per emitted electron [72]. This light is also picked up by the PMTs and is called the S2 signal. An example event in the TPC containing an S1 and S2 pulse is shown in Figure 2.6.

The configuration of the LZ TPC, including the electrode, liquid level, and PMT positions, is shown in Figure 2.7, and closer looks at the regions near the electrodes are shown in Figure 2.8. The 3-inch diameter Hamamatsu R11410-22 PMTs in the TPC, numbering around 500 split between the upper and lower arrays, were developed to have low backgrounds specifically to be used in LZ [69]. Further, the electrode grids were woven from stainless steel (304) wires of thickness 75-100 μm using a custom loom developed at SLAC [63]. In order from the bottom of the TPC these grids are the bottom PMT shield, cathode, gate, and anode, collectively responsible for three distinct electric field regions. Between the bottom PMT shield grid and the cathode is the reverse field region (RFR), where the electric field points upwards. Particle interactions in the RFR are not used for analysis due to the reversed field, but backgrounds for the main analysis in this thesis originate from it, and are the subject of Chapter 4. Between the cathode and the gate, still completely submerged in LXe, is the drift region, 1.45 m tall with a downward pointing

---

[1] This method is distinct from typical liquid argon TPCs, where the electrons are collected directly onto the grid wires having passed through two induction planes [71].

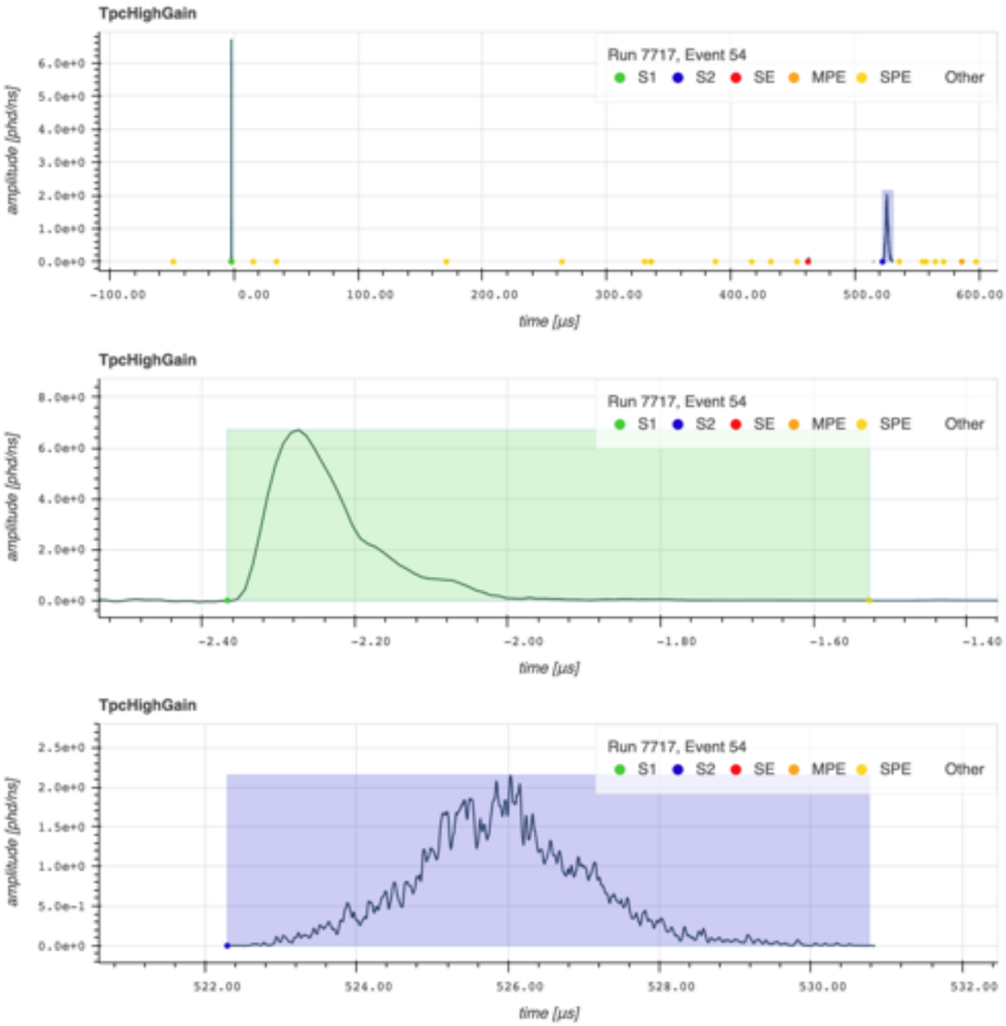


Figure 2.6: Top: Waveform of a single scatter event taken during the first science run, containing an S1 pulse and an S2 pulse, obtained by summing the waveforms of top and bottom PMTs. The time separation of the pulses, known as the drift time, is around 520 μs, corresponding to an event depth of halfway down the TPC. Center: Zoomed in view of the S1 pulse, which consists of 892 photons that the PMTs detected (phd). The pulse has a width of around 100 ns, largely set by the timing of the PMTs. Bottom: Zoomed in view of the S2 pulse, which has a far larger area (4.5×10^3 phd) and width (few microseconds). The relatively larger size of the S2 pulse indicates the large signal multiplication that occurs during electroluminescence.

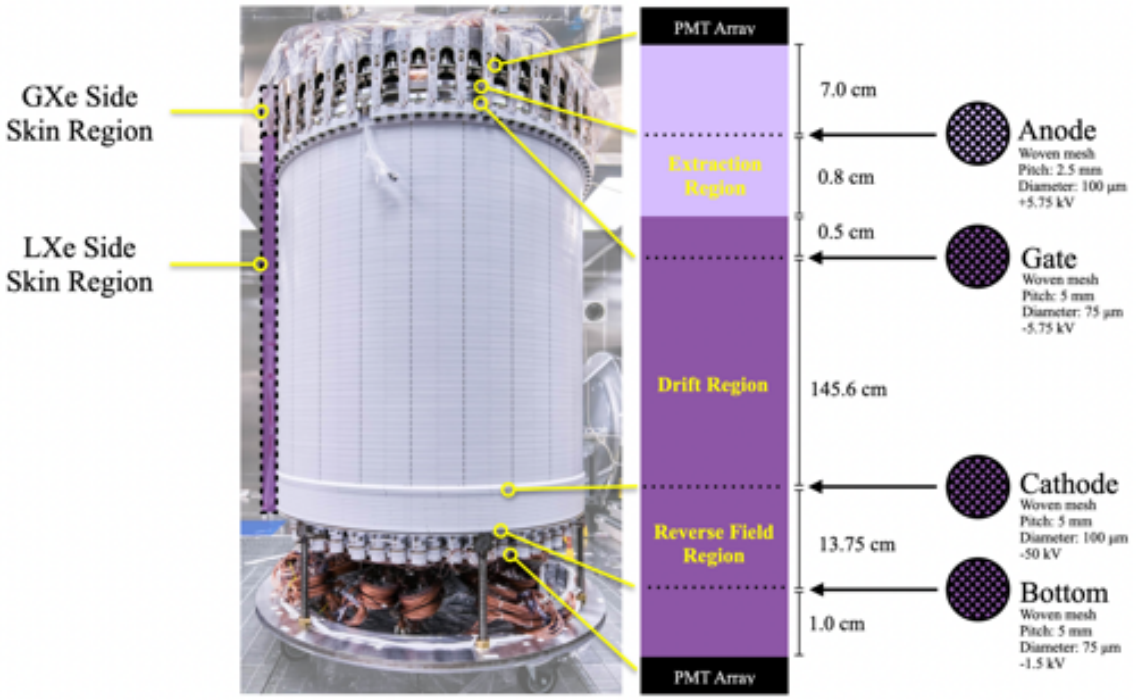


Figure 2.7: Photograph of the LZ TPC being assembled on the surface lab at SURF. The regions occupied by LXe, GXe, and the grid electrodes are indicated, along with their respective dimensions. Details about the mesh grids including the nominal operational voltages they were designed for are also shown. Figure taken from Ref. [63].

electric field. Above the gate grid is the LXe surface, a layer of gaseous xenon (GXe), and the anode grid. The GXe below the anode constitutes the extraction region, with a stronger electric field than the drift region. The drift and extraction regions are central to the operation of the TPC because they facilitate the transfer of charge from the interaction site to the extraction region where the S2 signal is produced. During the first science run (SR1) of LZ, the drift field was 193 V/cm and the extraction field was 7.3 kV/cm [5].

PMTs and photon measurement units

‘Photons detected’, or ‘phd’, is the unit commonly used throughout this thesis to measure the amount of light the PMTs saw. This unit is derived from the voltage time series that is digitized from the output of a PMT. The voltage time series is produced when a series of photons

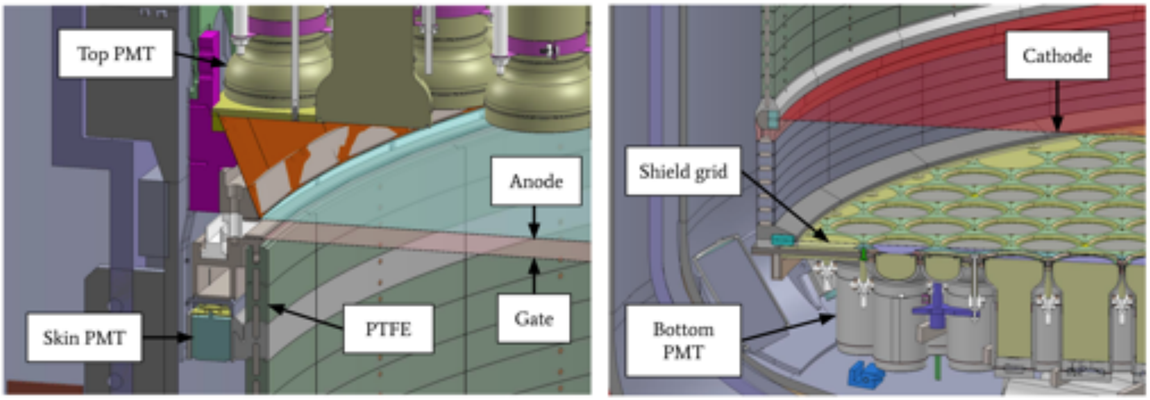


Figure 2.8: Left: Rendering of the top edge of the TPC, showing the top PMT array, a skin PMT, the anode and gate grids (the LXe surface, not shown, lies between these grids), and the PTFE wall. Right: The bottom edge of the TPC showing the bottom PMT array, the shield grid protecting the array from the cathode HV, the RFR below the cathode, and the cathode grid itself.

strikes the PMT cathode, photoionizing it, and the photoelectrons are accelerated towards the base of the PMT. On their way to the base, the photoelectrons pass through a series of dynodes held at roughly equal voltage spacings, and are thus multiplied by a constant factor at each dynode. The gain of a PMT G, which measures how many electrons are produced for each starting photoelectron, is thus given by

$$G = k\delta^{\gamma}, \tag{2.1}$$

where k is a constant, δ is the multiplication factor at each dynode, and γ is the number of dynodes [73]. The amplitude of the digital pulse can be converted to the number of starting photoelectrons using the gain, resulting in the photoelectrons (PE) unit.

To trace back the signal further to the number of photons incident on the PMT cathode, the efficiency of producing photoelectrons, called the quantum efficiency of a PMT, is required. The quantum efficiencies for the LZ TPC PMTs are around 30% for xenon vacuum ultraviolet (VUV) light in a cryogenic setting, but are subject to change with the wavelength of the incident light and temperature [70]. Around 20% of the time, an incident VUV photon produces two photoelectrons, instead of one [70]. This is known as the double-photoelectron effect, and is taken into account in the unit conversion from PE to phd. Tracing the signal even further back to the site of the interaction where the S1 photons were produced, or to the extraction region where the S2 photons were produced via electroluminescence requires a characterization of the signal collection efficiencies of the detector.

Light and charge detection efficiencies

The photon detection efficiency (PDE) is the number of photons incident on the PMTs for each S1 photon emitted at an interaction, and has been measured in SR1 of LZ to be $g_1 = 0.114 \pm 0.002$ phd/photon [5]. The PDE varies with the location of the interaction, as it is dependent on the solid angle the PMTs subtend, in addition to the reflectance of the PTFE walls, grids, and LXe-GXe surface. The quantity g_1, summarizing the PDE of the TPC, is defined to be the PDE evaluated at the radial center of the TPC ($x = y = 0$), and halfway down the length. The S1 areas used in analysis are corrected for position dependence, which causes the PDE to decrease for interactions near the wall (more PTFE reflections and less solid angle), and increase towards the bottom (less reflection off the LXe surface and more solid angle).

The charge detection efficiency (CDE) is greater than unity due to the electroluminescence process producing thousands of photons per extracted electron. Since this only occurs in the extraction region, the position dependence of the CDE is restricted only to the radial position of the interaction. The radial dependence is encoded in the photon detection efficiency in the GXe, $g_1^{\text{gas}}(x,y)$, which is one of four factors that determines the CDE. The other factors include the electron extraction efficiency, ε_{ext}, which is a function of the extraction field, the light yielded by the GXe in response to the extracted electrons, Y_e, and the exponential loss $\exp^{-t/\tau}$ of electrons as they drift upwards and attach to impurities, where t is the time spent drifting, and τ is the electron lifetime. The quantity g_2 summarizing the CDE of the TPC is defined for $t = 0$ (top of TPC; no electron loss to impurities) and $x = y = 0$, given by

$$g_2 = Y_e \varepsilon_{\text{ext}} g_1^{\text{gas}}(0,0). \tag{2.2}$$

The SR1 measurement yielded $g_2 = 47.1 \pm 1.1$ phd/electron [5].

Energy reconstruction

TPCs rely on the scintillation and ionization signals to reconstruct energy. Reconstructing the energy of an event is achieved by noting that the number of photons and electrons produced is proportional to the energy deposited, such that

$$E = \frac{W}{L}(N_{\text{ex}} + N_i), \tag{2.3}$$

where E is the energy deposited, W is the average energy required to ionize or excite an electron in LXe, L is the 'Lindhard' factor that accounts for energy lost to heat, N_{ex} is the number of excitations, and N_i is the number of ionizations. The associated LXe microphysics is discussed in section 2.3.

For interactions with the atomic electrons of xenon, L is constant with respect to energy, and thus the fraction of heat may be absorbed into W. Eq. 2.3 may be rewritten in terms of the signal sizes and gain parameters, to give

$$E = W\left(\frac{\text{S1c}}{g_1} + \frac{\text{S2c}}{g_2}\right), \tag{2.4}$$

which may be used to reconstruct the energy of electronic recoil (ER) events. For nuclear recoils (NRs), the Lindhard factor L depends on energy, leading to a parameterized scheme for the NR energy reconstruction [74].

A useful form of Eq. 2.4 is

$$\frac{S2}{E} = -\frac{g_2}{g_1}\frac{S1}{E} + \frac{g_2}{W}, \tag{2.5}$$

which casts the energy-normalized signal sizes as linearly anti-correlated variables. This anti-correlation may be used to extract g_1, g_2, or W if data exists for a set of monoenergetic ER sources, and is the subject of Appendix A

Position reconstruction

Since the S1 pulse is recorded immediately after the interaction occurred, and the S2 pulse is recorded only after the electrons have drifted to the LXe surface, the time difference between the S1 and S2 pulses is a measure of the z (vertical) position of the interaction. Electrons drift at a constant speed in noble liquids under constant electric fields; the electrostatic acceleration is balanced by the drag of coherent scattering with atoms [75]. In LXe, for fields of 100 V/cm and higher, the drift speed is on the order of 10^5 cm/s [75]. Because of the constant drift speed, the drift time is often used in this thesis as a proxy for vertical distance.

The (x, y) reconstruction of an event is possible by calculating the centroid of the S2 light hit pattern imprinted on the top PMT array. A good position reconstruction is essential as a well-determined position will allow for a more accurate correction to the pulse sizes, resulting in a better energy resolution. The layout of the top PMTs has been optimized to improve the position reconstruction of events near the TPC walls, and is shown in Figure 2.9. In LZ, the MERCURY algorithm, which employs a maximum likelihood method, is used for position reconstruction and yielded a 1σ resolution of 0.7 mm near the cathode electrode along the cylindrical axis of the TPC, which increases by a factor of two near the walls [5, 76].

2.2.2 Veto systems

A plausible claim of dark matter discovery will be supported by a comprehensive knowledge of background events, supplied by the energy and position reconstruction capabilities of the TPC, ER/NR discrimination, and the veto system. Equipping the regions neighboring the TPC with additional radiation detectors allows neutrons, γ rays, and other multiply scattering particles to be tagged and excluded from the analysis. The ability to reject these backgrounds allows regions of LXe near the TPC walls to be used for the WIMP search, resulting in a larger fiducial volume. Further, these detectors can be used to assess the local radiation environment produced by both

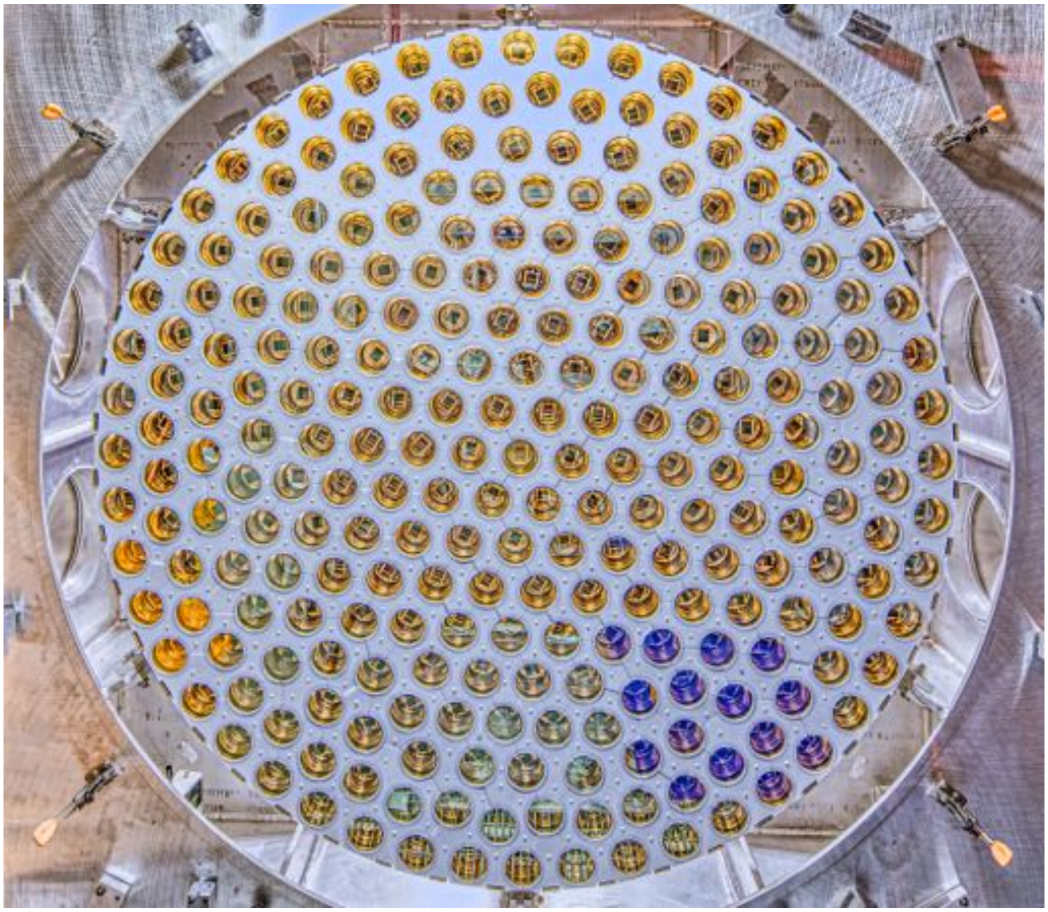


Figure 2.9: Hybrid pattern of the top PMT array with a hexagonal packing in the center transitioning to concentric rings on the outside. Photograph by Matthew Kapust.

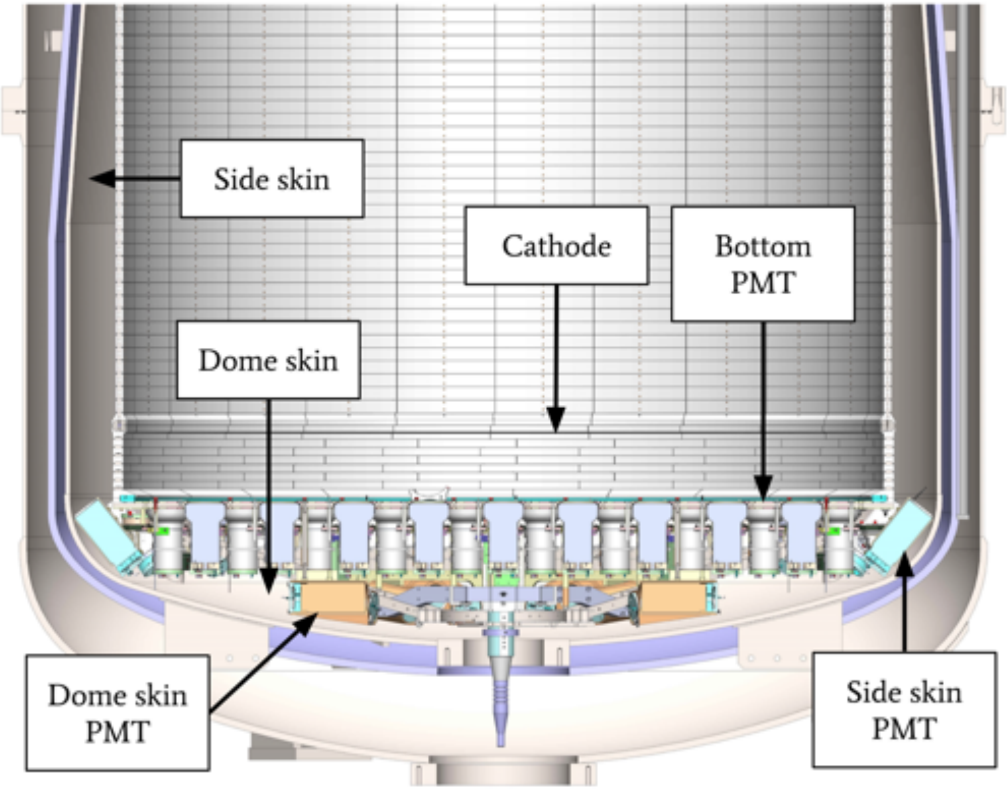


Figure 2.10: Rendering of the bottom half of the LZ cryostats, showing the xenon skin and the bottom skin PMTs.

internal (within the TPC) and external sources. The LZ veto system is composed of the xenon skin and the outer detector (OD), which are described here.

Xenon skin veto

The xenon skin is the layer of LXe outside the TPC and within the wall of the inner cryostat vessel, with a mass of more than 2 tonnes. The cylindrical shell of LXe outside the TPC wall, called the side skin, is instrumented with 93 1-inch PMTs on the top half and 20 2-inch PMTs on the bottom half [50]. The LXe below the bottom PMT array, called the dome skin, is instrumented with 18 more 2-inch PMTs. Figure 2.10 depicts these regions.

The goal of the xenon skin is to identify scatters of γ rays from decays in detector materials, or from radiative neutron capture. It is designed to identify events that deposit more than 100 keV in over 95% of the skin's volume [50]. The close proximity of the xenon skin to the TPC facilitates the rejection of γ rays with MeV energies with a high efficiency since the associated mean free path in LXe is a few centimeters. These prevalent γ rays originate from the decays of rare earth element contaminants in detector materials and from radiative neutron capture. Neutrons with MeV energies, such as those produced by the spontaneous fission of the trace element ^{238}U and from (α, n) reactions, travel tens of centimeters in LXe before interacting. The surrounding OD is therefore better suited to tag neutrons that have interacted in the TPC; but the skin helps by registering the γ rays from radiative neutron capture emitted from the OD, providing a coincidence tag. A secondary goal is to reject events that are indicative of a light leak between the TPC and skin; an unaddressed light leak could have the harmful effect of distorting the S1 and S2 pulse areas of an ER event to mimic a WIMP (NR) event. In SR1, events with pulses in the skin within 0.5 μs of an S1 pulse in the TPC were removed from the WIMP search data [5].

The skin was also used to remove events from the radioactive decay of ^{127}Xe, a cosmogenically activated isotope which decays to the ground state of ^{127}I, producing atomic de-excitations accompanied by γ rays. The TPC signals produced by the low energy atomic emissions pose a background to the WIMP search, especially for decays near the wall where the γ rays are likely to escape the TPC without interacting. The skin readily identified these decays in SR1, reducing the ^{127}Xe background by a factor of 5 [5].

Outer detector

The OD is comprised of a near-hermetic layer of liquid scintillator (LS), with mass 17.5 tonnes, surrounding the cryostats, held in a set of clear acrylic tanks and observed by an array of 120 8-inch OD PMTs. A schematic of the OD acrylic tanks are shown in Figure 2.11, and a photograph of the empty assembled tanks with the surrounding OD PMTs is shown in Figure 2.12.

The LS is based on the organic compound linear alkyl benzene, and has been chemically loaded with the element gadolinium (Gd) at a concentration of 0.1% by mass [50]. The isotopes

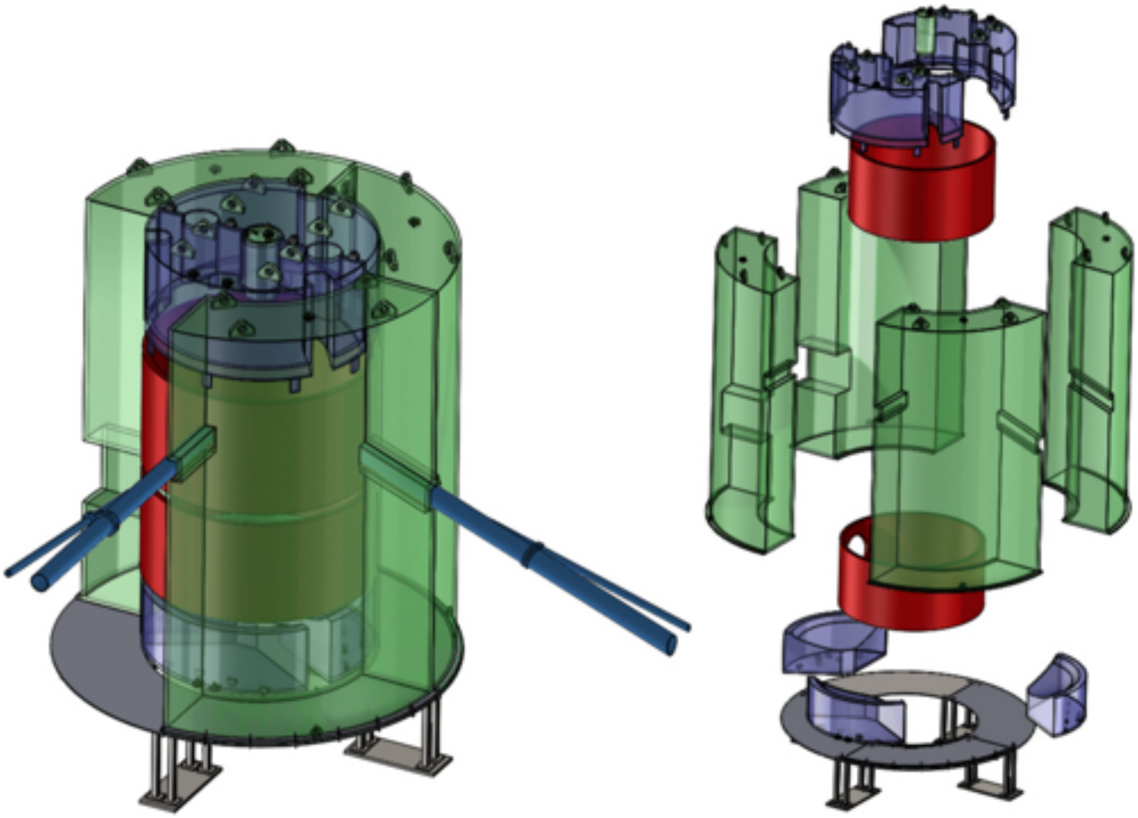


Figure 2.11: Left: Assembled layout of the ten LZ OD tanks (green and purple) and water displacers (red). Right: Exploded view of the tanks, showing the two top acrylic tanks, the four side acrylic tanks that hold nearly 90% of the LS, and the three bottom acrylic tanks. The small cylindrical vessel between the top tanks is replaced with a yttrium beryllium (YBe) photoneutron source during calibrations. Figures taken from Ref. [50].

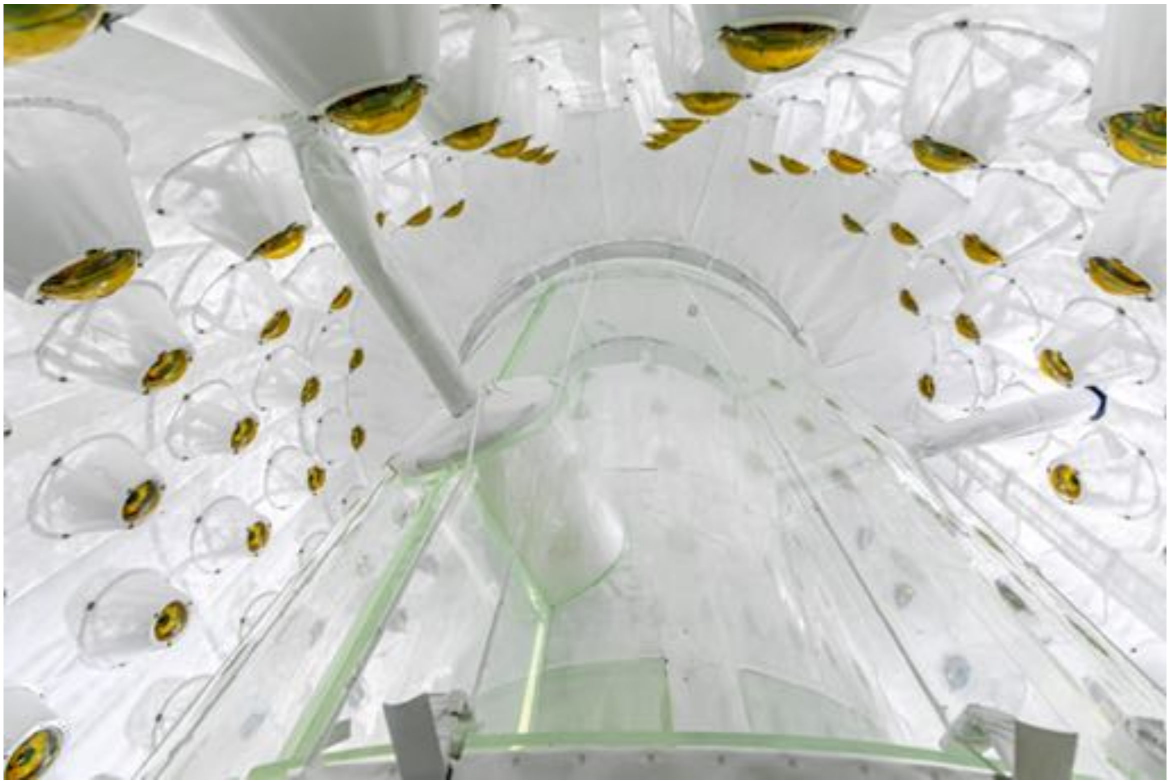


Figure 2.12: Photograph of the side acrylic tanks prior to filling. Surrounding the tanks are 120 8-inch PMTs that observe scintillation light from the LS. The space between the tanks and the PMTs will be occupied by water after filling.

^{157}Gd and ^{155}Gd have some of the largest neutron capture cross sections among stable nuclides (2.6×10^5 barns for ^{157}Gd [77]), and thus serve as efficient and fast-acting neutron absorbers. The γ rays from the neutron captures on Gd (more than 90% of the time) and hydrogen cause the LS to scintillate, indicating the presence of a neutron. Gd-LS was previously used as the primary target in the three large reactor neutrino experiments at Chooz, Palo Verde, and Daya Bay [78–80].

In LZ, the position-averaged neutron tagging efficiency of the OD was measured using americium-lithium (AmLi) neutron calibration sources to be $88.5 \pm 0.7\%$ [5]. Events where an S1 pulse in the TPC was followed by a large pulse in the OD within a 1,200 μs window were removed for the SR1 WIMP search [5]. Figure 2.13 shows the trajectory of multiply scattering neutrons during SR1 across the TPC, skin, and OD detectors, illustrating the importance of the veto systems.

Filling the outer detector

The author spent several stints at SURF working on the assembly and commissioning of a few components of LZ. For instance, in the spring of 2020, the inline radon reduction system (iRRS) was assembled in Ann Arbor, shipped to South Dakota, and installed in LZ by a Michigan team that included the author. This operation has been described in the PhD thesis of the author's colleague Maris Arthurs [82]. Onsite work conducted the following year is sketched in the vignettes below.

When the summer of 2021 began, the OD tanks were assembled and ready to be filled along with the water tank. Simultaneous filling of the water and Gd-LS was required because otherwise, the liquid would exert large hydrostatic pressure on the sides of the acrylic tanks. This co-filling was monitored using level sensors read out via LZ's slow control interface, which had to be calibrated before filling.

Each acrylic tank had an inlet for Gd-LS introduction and a vent outlet that terminated in a

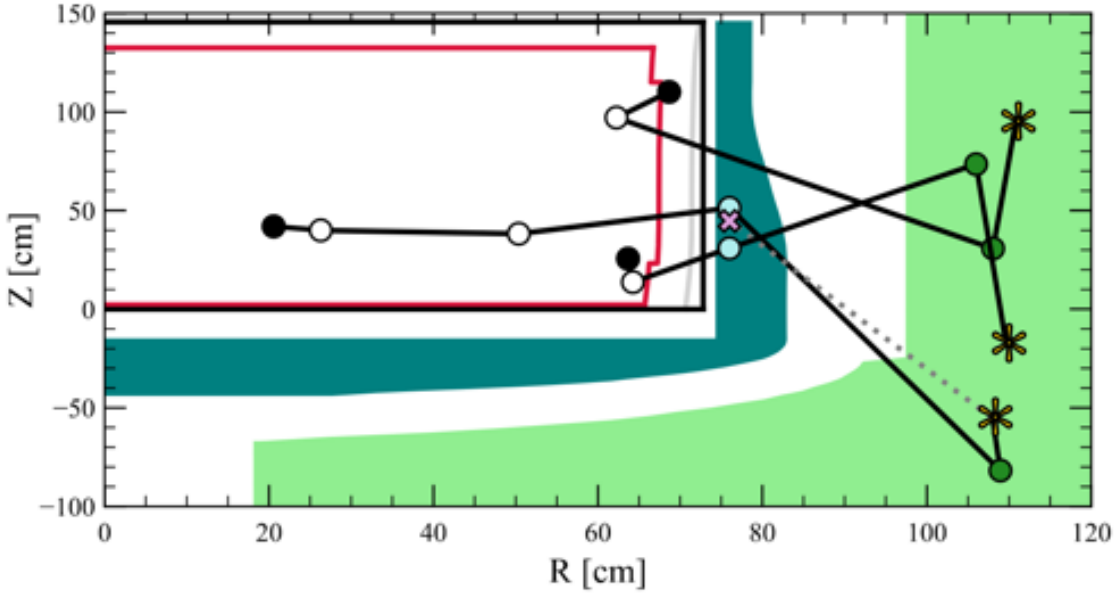


Figure 2.13: Positions of neutron events showing the scattering path of across the TPC and veto detectors with a coincidence window of 400μs. The red outline indicates the fiducial volume of the TPC, the black line is its physical extent, the teal is the xenon skin volume, and the outermost green region is the LS in the OD tanks. Black points indicate interactions with the shortest drift time in the TPC, while the white circles represent the remaining interactions. Interactions in the skin and OD are colored blue and green, respectively. Neutron captures in the OD are marked with a star, and the resulting γ ray deposit in the skin are marked with a pink cross. Plot taken from Ref. [81].

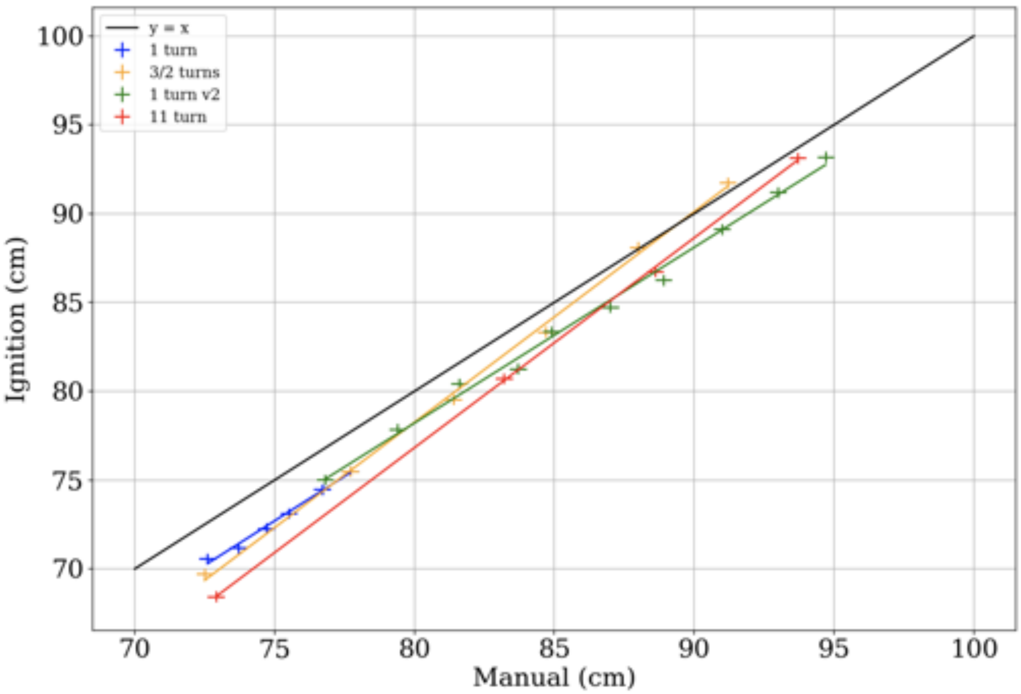


Figure 2.14: Level sensor readout from Ignition, the LZ slow control software, versus the manual measurement on the test vessel, shown for various fill rates (quantified as turns of a valve).

bubbler for pressure equalization. The inlet for each tank was fed from a fill line (one per tank grouping), connected in parallel to 6 m long laser rangefinder tubes that measured the liquid level; due to the common inlet line, the liquid level was expected to be equal in the tank and sensor tube. A clear 10 gallon test vessel was placed in the water tank to test the calibration of the level sensors and the Gd-LS distribution system. Figure 2.14 shows the comparison of the liquid level rise during a test fill measured by the level sensor and manually on the test vessel using a tape measure. The error on the level sensor, coupled with the coarseness of control in the Gd-LS distribution system, was deemed acceptable to fill within the pressure ratings of the acrylic tanks (a level difference of less than 20 cm between the Gd-LS and water was required). Transient behavior was observed on the level sensor readings, shown in Figure 2.15, which could have been caused by a variety of reasons, including a delayed pressure equalization and foam formation inside the level sensor.

It was realized before the filling began that the small cross-sectional area of the 1/2-inch

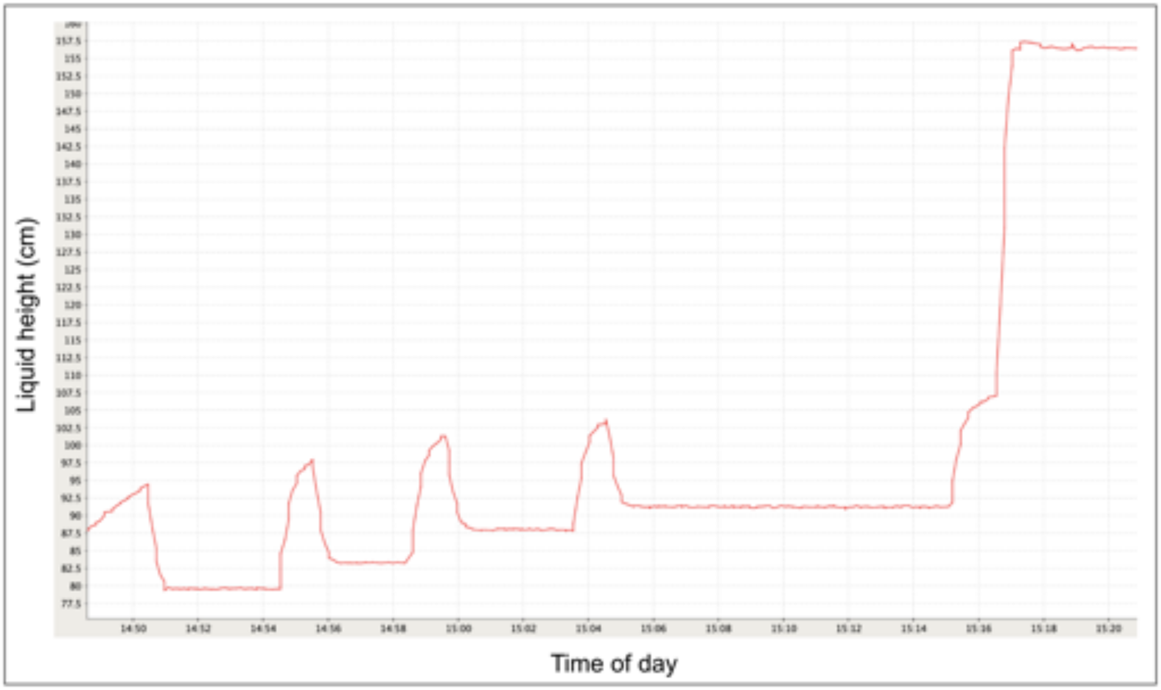


Figure 2.15: Level sensor reading as a function of time over a 30 min interval during which five bouts of filling occurred, reaching the top of the test vessel (first four spikes), and into the thin vent line of the test vessel (large final jump). Hypotheses for the spikes include foaming of the Gd-LS in the level sensor tube, and an accelerated filling of only the sensor tube caused by differences in inlet hose lengths.

diameter PTFE vent lines leading out from the top of the acrylic tanks posed a dangerous risk of damaging the tanks. In the intended final state of the OD, the Gd-LS level would rest in reservoirs above the water tank; the vent lines would be completely full. When getting to this stage there would be a vulnerable period where the vent lines are partially full, where for example, a small temperature change or inlet flow instability could cause the Gd-LS level to rise much faster than the surrounding water level, creating in effect a giant thermometer. The large hydrostatic pressure exerted by the small amount of Gd-LS in the vent tubes would have cracked the tanks. The potential danger to the acrylic vessels was solved by the rapid installation of 4-inch diameter tubes that branched off the fill lines. The purpose of these tubes was to act as a volume buffer during the vent-filling period.

With these preparations the OD and water tank fill commenced. First, the Gd-LS was brought underground in 120 drums, where it was extracted barrel-by-barrel into a staging reservoir above the water tank. Nitrogen gas was bubbled through the Gd-LS to remove any dissolved oxygen, which is a quenching agent that reduces the light yield of the scintillator [83]. Small amounts of Gd-LS were distributed into the acrylic tanks to further test the filling system. An ultraviolet (UV) photograph taken during this initial fill (LS fluoresces in UV light) is shown in Figure 2.16.

Despite issues such as leaks of Gd-LS from the staging reservoir, preferential filling of the acrylic tanks, and poor pressure equilibration, the OD was successfully co-filled with the water tank over a month-long operation in 24-hour shifts, in LZ's first instance of round-the-clock underground shifts. Figure 2.17 shows the progress of filling the three acrylic tank groupings in June 2021.

2.3 Microphysics of liquid xenon

This section describes the salient physical processes in LXe that produce the VUV photons and electrons that manifest as S1 and S2 signals, respectively, in the TPC. The nature of the initial

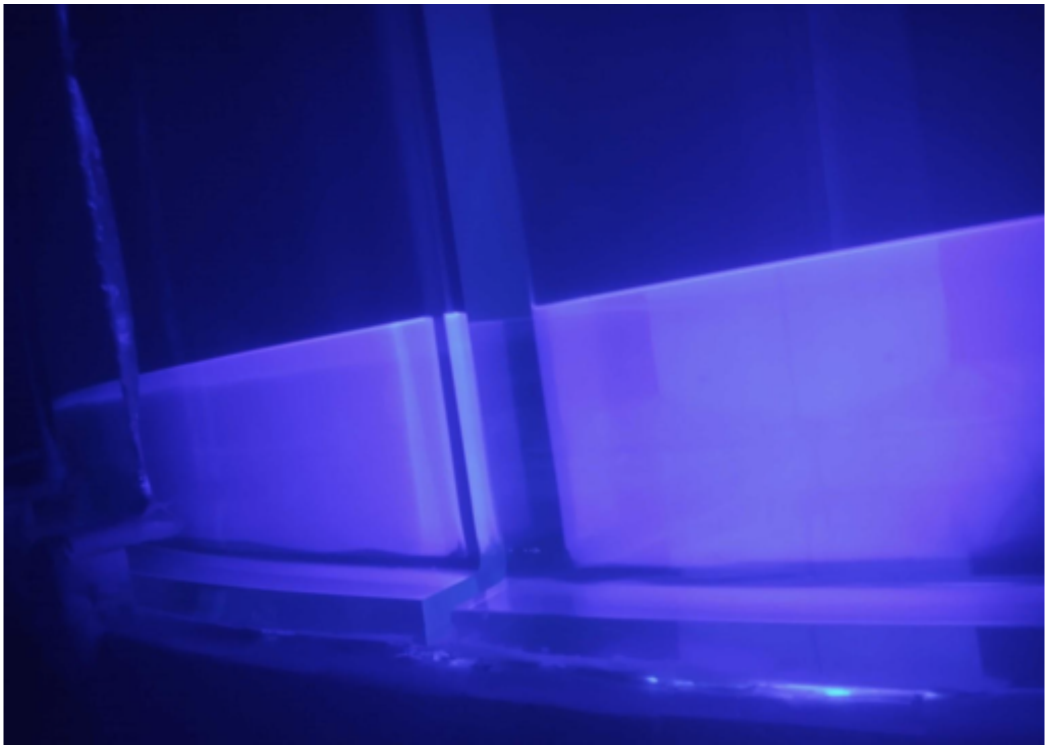


Figure 2.16: UV photograph of the bottom acrylic vessels filled with a few centimeters of Gd-LS.

energy deposit determines the mix of energy dissipation channels that contribute to the photon and electron yields, resulting in the ER/NR discrimination properties of LXe. Several energy scales involved in signal production are discussed, motivating studies to measure LXe work functions and yields, detailed in Chapters A and B.

2.3.1 Quanta production in xenon

LXe, like most noble liquids, possesses a structure similar to solids due to weak (0.01 eV) van der Waals forces arising from spontaneously polarized xenon atoms [84]. For example, there is direct evidence of electronic band stuctures in the liquid forms of xenon, argon, and krypton [85]. The interplay of this structure with the energy dissipation in the liquid gives rise to a rich microphysics phenomenology that is continously being revealed by experiments.

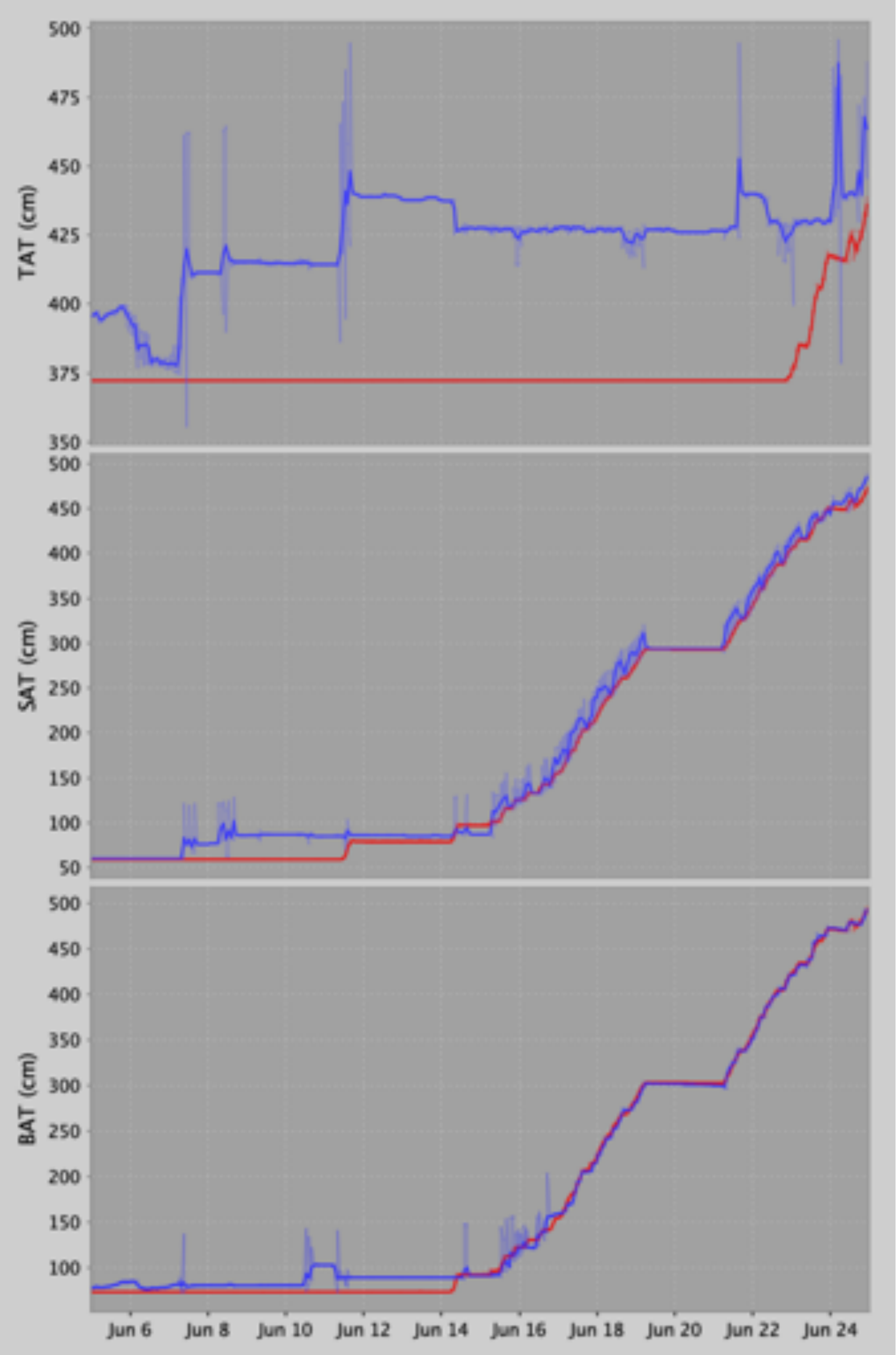


Figure 2.17: Measured liquid level (blue) of each tank grouping as a function of time for the top tanks (top), side tanks (middle), and bottom tanks (bottom). The red lines show the target values, calculated from measurements of the water level such that the level difference is less than 20 cm. The top tanks are off-target because of persistent issues that affected their pressure equalization.

Energy dissipation channels

The three channels through which energy is dissipated in LXe broadly are light, charge, and heat. Only the light and charge channels are observable in a TPC, and are the only channels facilitated by atomic and molecular transitions; heat is dissipated via motion under the threshold of atomic transitions. The atomic transitions of interest are electron excitations from the ground state to higher energy levels, and ionization where electrons are completely liberated. After a particle interaction in LXe depositing energy E_0, the energy sharing that immediately follows (with no secondary particle interactions) is formalized with the Platzmann equation

$$E_0 = N_i^0 \bar{E}_i + N_{\text{ex}}^0 \bar{E}_{\text{ex}} + N_i^0 \varepsilon, \tag{2.6}$$

where N_i^0 is the number of ionized atoms with average ionization energy $\bar{E}_i$, N_{ex}^0 is the number of excited atoms with average excitation energy $\bar{E}_{\text{ex}}$, and where ε is the average kinetic energy of the electrons from ionization [84]. Note that Eq. 2.6 does not take into account the kinetic energy transferred in the initial collision, nor the collisions, excitations, and ionizations of secondary particles.

The charge, mass, and energy of the impinging particle also play roles in signal production. For instance, a γ ray could Compton scatter, depositing an initial amount of energy, and then scatter again in the immediate vicinity of the first interaction. This results in a different electron to VUV photon ratio than if the γ ray had been completely photoabsorbed initially. Similarly, a neutron may elastically scatter with several nuclei in the same vicinity, resulting in a proportionally larger kinetic energy transfer compared to a γ ray. Therefore, the microphysics is best suited to be tracked with software such as the Stopping and Range of Ions in Matter (SRIM) package, and the parameterizable Noble Element Simulation Technique (NEST) [86, 87].

For the purposes of characterizing the signal production properties of target media, the W-

value, the average energy required to ionize or excite an electron in LXe, is defined as

$$W = \frac{E_0}{N_i + N_{\text{ex}}}, \tag{2.7}$$

where N_i is the total number of ionization electrons (accounting for secondary particles), and where N_{ex} is the total number of excited atoms. The currently accepted value for LXe is $W = 13.7 \pm 0.2$ eV [88], but there are recent discrepant measurements that motivated us to revisit the assumptions of the present W-value definition, discussed in Appendix A.

Recombination

After ionization, some free electrons combine with ions and produce excitations in a process known as recombination. In fact, most of scintillation light observed in an event originates from recombined excited atoms [88]. Recombination reduces the number of liberated electrons, such that

$$N_i^r = N_i(1 - r), \tag{2.8}$$

where N_i^r is the number of electrons left after recombination, and where r is the recombination fraction. The number of excited atoms after recombination, N_{ex}^r, correspondingly increases such that

$$N_{\text{ex}}^r = N_{\text{ex}} + rN_i. \tag{2.9}$$

Random fluctuations in r therefore cause the detected numbers of VUV photons and electrons to be anti-correlated, a fact that is taken advantage of in the energy reconstruction. In the framework where the W-value is common to both ionization and excitation, such as in Eq. 2.7, recombination simply shuffles the numbers of quanta such that

$$N_i^r + N_{\text{ex}}^r \to N_i(1 - r) + N_{\text{ex}} + rN_i = N_i + N_{\text{ex}}. \tag{2.10}$$

This simplification is a boon to dark matter and neutrino TPCs that use electric fields to extract ionization signals. Introducing an electric field lowers the recombination fraction, affecting the numbers of electrons and photons read out (although total quanta remains the same). The common W-value framework allows the energy calibration process to remain somewhat independent of the phenomena associated with recombination.

Scintillation

The process by which VUV light is emitted is described in the following three-body interaction between xenon atoms

$$e^- + \mathrm{Xe} \rightarrow \mathrm{Xe}^* + e^-, \tag{2.11}$$

$$\mathrm{Xe}^* + \mathrm{Xe} \rightarrow \mathrm{Xe}_2^{*,\nu}, \tag{2.12}$$

$$\mathrm{Xe}_2^{*,\nu} + \mathrm{Xe} \rightarrow \mathrm{Xe}_2^* + \mathrm{Xe}, \tag{2.13}$$

$$\mathrm{Xe}_2^* \rightarrow 2\mathrm{Xe} + \gamma_{\mathrm{VUV}}, \tag{2.14}$$

where the initial excitation Xe^* is due to an electron e^-, $\mathrm{Xe}_2^{*,\nu}$ is a dimer with vibrational energy, Xe_2^* is the vibrationally and non-radiatively de-excited dimer, and γ_{VUV} is the 175 nm (7.1 eV) VUV photon emitted [84].

The Xe^* exciton radiatively de-excites to the ground state, but the resulting photon is resonantly absorbed and re-emitted by neighboring atoms. It is only when Xe^* accesses molecular degrees of freedom via Xe_2^* that the de-excitation photon is of low enough energy to pass through LXe unimpeded. The dissociation of the Xe_2^* dimer occurs when it de-excites from either the singlet $^1\Sigma_u^+$ state or triplet $^3\Sigma_u^+$ state, each with lifetimes on the order of 1 ns and 10 ns, respectively [84].

In addition to the direct exciton production in Eq. 2.12, an ion and electron recombining will also produce an exciton by means of shedding energy to heat via vibrational and molecular degrees of freedom [89]. The three modes of scintillation: singlet, triplet, and recombination transitions, despite producing the same VUV photon, each have characteristic timings that have been measured in xenon [90].

Charge amplification

Electrons surviving recombination are drifted up in the electric field of the TPC, and extracted into the GXe. The drift in the LXe can be modeled as a process where electrons scatter off xenon atoms with an effective potential that represents the single-atom potential corrected for long range polarization forces of neighboring atoms [91]. The effective potential can be determined from the liquid structure, measured with neutron diffraction experiments, and the model is solved using the Boltzmann equation with a collisional term [84]. An electron cloud drifting in an electric field is subject to self-repulsion, spreading it in transverse and longitudinal directions. For the typical sizes and fields employed in a LXe TPC, the spread is on the order of a millimeter, providing an advantage for precise position reconstruction [92, 93].

For typical drift fields, the energies of drifting electrons are usually under 1 eV, which are below the excitation threshold of LXe (around 10 eV). It may be energetically favorable, however, for impurities in the liquid to capture electrons, forming negative ions with reduced drift speeds that result in a smaller S2 yield. The high liquid purity of LXe TPCs is a therefore major purification achievement; in LZ, electron lifetimes as large as 8 ms were measured, around 8 times larger than the longest possible drift time [56].

Having flowed past the gate electrode into an LXe region with a stronger electric field, the electrons are extracted into the GXe, where the field is even stronger by a factor of the LXe dielectric constant (1.85) [94]. The extraction efficiency is dependent on the extraction field, and has been observed to saturate for liquid fields above 7.5 kV/cm [95]. The extraction process is not completely understood and might be complicated by a local potential minimum below the liquid surface induced by competing forces on the electrons, one due to the electric field and other due to repulsion from the LXe-GXe dielectric boundary [96]. For instance, delayed electron emission up to milliseconds after the S2 pulse has been observed in many LXe TPCs, consistent with a trapping potential [97–99].

Following extraction, the electrons scatter and excite the atoms in the GXe, which scintillate via the three-body excimer channel described in subsection 2.3.1. The resulting 175 nm light is thus identical to the primary (S1) scintillation light, and is mostly detected by the top PMT array. The width of the S2 pulse, typically on the order of a microsecond, is set by the size of the gas gap (above the liquid and below the anode grid) and the magnitude of the extraction field in the GXe. In addition to the excimer-based electroluminescence, there are observations of non-VUV light emission below excitation energies resulting from neutral bremsstrahlung and partial atomic transitions [100, 101]. However, the yields of these emissions appear to be too low to impact current dark matter detectors.

ER NR discrimination

A crucial feature of LXe TPCs is their capacity to distinguish between NR and ER events. To first order, the two types of interactions differ in the fraction of energy that is lost to heat. In the Lindhard-Scharff-Schiott (LSS) formulation of energy partitioning in condensed media, the fraction lost to kinematic atomic motion, L, in a single nuclear scatter is parameterized as a function of the incoming particle energy; for a neutral projectile with 10 keV kinetic energy, L is approximately 0.2 [102–105]. The fraction described above is unity for ER events, since scattering with light atomic electrons does not generate much kinetic energy, as opposed to scattering with heavier nuclei. Further, L does not include the kinetic energy of liberated electrons with sub-excitation energies. Therefore, the disparity between the heat channels of NR and ER interactions results in two different energy scales used for the two types of events. The energy reconstructed in an NR interaction is related to the amount of energy deposited by an electron of equal initial energy by

$$E_{\mathrm{ee}} = L E_{\mathrm{nr}}, \tag{2.15}$$

where E_{ee} is the ‘electron-equivalent’ energy of the observed NR energy E_{nr}. In practice, a different relationship that accounts for discrepancies in the LSS model is used; see for instance Ref [106].

However, ER–NR discrimination irrespective of energy is possible in LXe, indeed necessary to identify WIMPs depositing various amounts of energy in LXe. The ratio of the primary scintillation (S1) photons to ionization electrons (that result in the S2 pulse) is fundamentally different for ER and NR events; the S1/S2 ratio is always smaller for NR events than for ER events. The two-fold explanation for this phenomenon includes differences in the initial exciton to ion ratio immediately following the ER or NR interaction, and different recombination effects from the distinct ionization densities between ER and NR events [106]. The discrimination is showcased in the LZ TPC response to ER and NR calibration sources during SR1, shown in Figure 2.18. The leakage fraction, defined as the fraction of ER events falling below the NR median, is typically 10^{-3} to 10^{-2}, depending on the energy deposit [106].

With the description of the LZ detector and principles of detection, reconstruction, and discrimination given in this chapter, an account of the dark matter search can be presented. Chapter 3 describes a search for WIMPs extending up to the largest NR energies ever considered in an analysis.

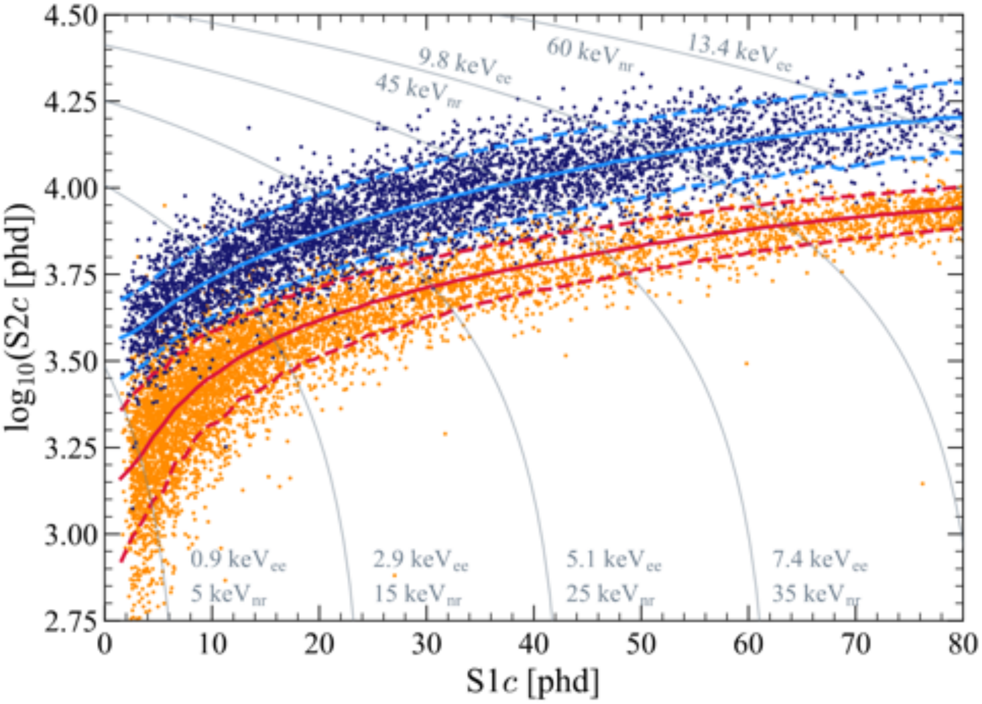


Figure 2.18: Calibration data in the S1 and $\log_{10}$S2 space (corrected for position), with tritium (blue) constituting a β source and neutrons (orange) an NR source. The medians of the simulated ER and NR bands are indicated in solid lines, and the dotted lines indicate 10% and 90% quantiles. Gray lines show contours of constant energy. Figure taken from Ref [5].

Chapter 3

Effective Field Theory Search in LZ

It was astonishing what a line you could draw if you had two points of physical constants and aesthetic preferences.

Marissa Lingen, How we know they have faces

3.1 Introduction

The Effective Field Theory (EFT) analysis of the LZ collaboration's first science run (SR1) is presented in this chapter. These results extend the WIMP search (WS) result of SR1, which probed spin-independent (SI) and spin-dependent (SD) interactions between WIMPs and nuclei [5]. The goal of the EFT is to account for all the relevant degrees of freedom, in addition to spin, at the energy scale of the WIMP–nucleon scattering process. This approach creates a comprehensive set of interactions which produce nuclear recoils at higher energies than the SI and SD interactions, motivating a larger energy window. In the EFT analysis the energy window is a factor of 7.5 larger than the WS, reaching up to 270 keV_{nr}. World-leading upper limits for the strengths of 56 WIMP-nucleon interactions comprising the EFT are presented in this chapter, and are the most constraining for a general WIMP in the GeV to TeV mass range. Critical to the expansion of the energy window is a tool to remove a high energy NR-mimicking background, which is the subject of Chapter 4.

3.1.1 The momentum-independent result

In July 2022 the LZ collaboration released its first WS results containing world leading upper limits for SI and SD interactions, shown in Figures 3.1 and 3.2, respectively. The SI and SD search strategies are focused on achieving a low energy threshold, since the respective nuclear recoil spectra peak at zero energy and fall off exponentially due to the Maxwellian tail of the WIMP velocity distribution. Further, the SI and SD scattering matrix elements are independent of the momentum transferred and hence do not lead to suppressed cross sections in the low recoil energy limit.

The SI and SD interactions are a subset of the EFT; they are both naturally generated as the operators O_1 and O_4, described in subsection 3.1.2. The EFT search keeps the low energy threshold of the WS while extending the high energy cutoff. In terms of the position-corrected scintillation signal S1c, the EFT analysis window of 3 phd $<$ S1c $<$ 600 phd contains the WS energy window of 3 phd $<$ S1c $<$ 80 phd, providing consistency checks of the SI and SD interactions between the WS and EFT analyses. The various other operators generated by the EFT exhibit couplings not only among WIMP and nucleon spins, but also among momentum transfer and the characteristic velocity of nucleons in the rest frame of the nucleus as well. The resulting cross sections of some EFT operators vanish in the limit of zero recoil energy, motivating a look at higher energies.

3.1.2 The non-relativistic dark matter effective field theory

The idea of the EFT is to simplify a large space of interactions, stemming from numerous DM models and the associated couplings to the Standard Model, to a much smaller set of nuclear responses that are indistinguishable from each other. One way to view the larger space of testable

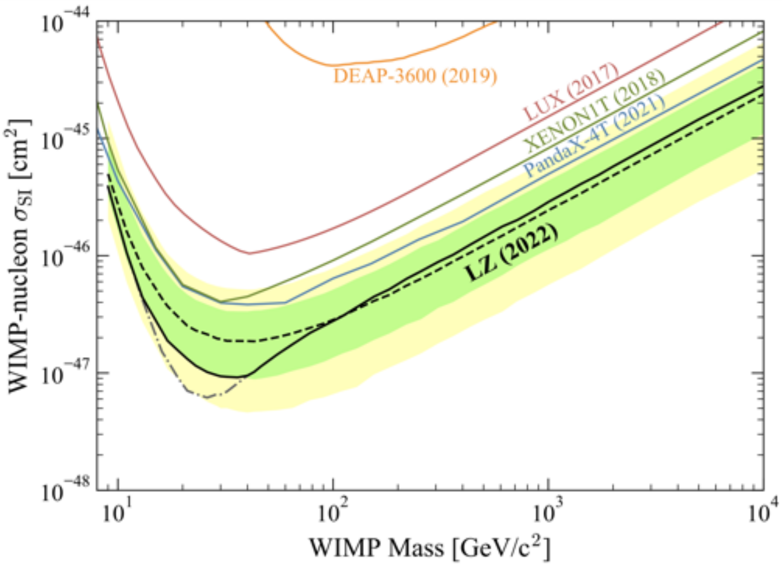


Figure 3.1: The SI WIMP cross section 90% confidence limits (black) as a function of WIMP mass, with 1σ (green) and 2σ (yellow) bands. The dotted line shows the median expected sensitivity projection. Also shown are the PandaX-4T, XENON1T, LUX, and DEAP-3600 limits. Figure from Ref. [5].

models is to parameterize how their cross sections scale with momentum transfer in the low energy limit [107]. This approach naturally cleaves the search data into a low-energy, momentum-independent region, as in the WS, and a higher energy momentum-dependent region as in the EFT search. However, the non-relativistic EFT (NREFT) goes further and uses all the relevent degrees of freedom to build from the bottom up all the possible interactions direct detection experiments should be able to detect [108]. The framework therefore captures momentum, spin, and velocity dependences of the cross sections, up to leading order. The choice of the degrees of freedom are set by the context of the WIMP-nucleon scattering process. For instance, the velocity of the WIMP wind is around 10^{-3}c, so the EFT can be non-relativistic. Secondly, the maximum momentum transferable to a nucleon from a WIMP scatter is around 200 MeV, implying that the nuclear structure of the xenon target is relevant.

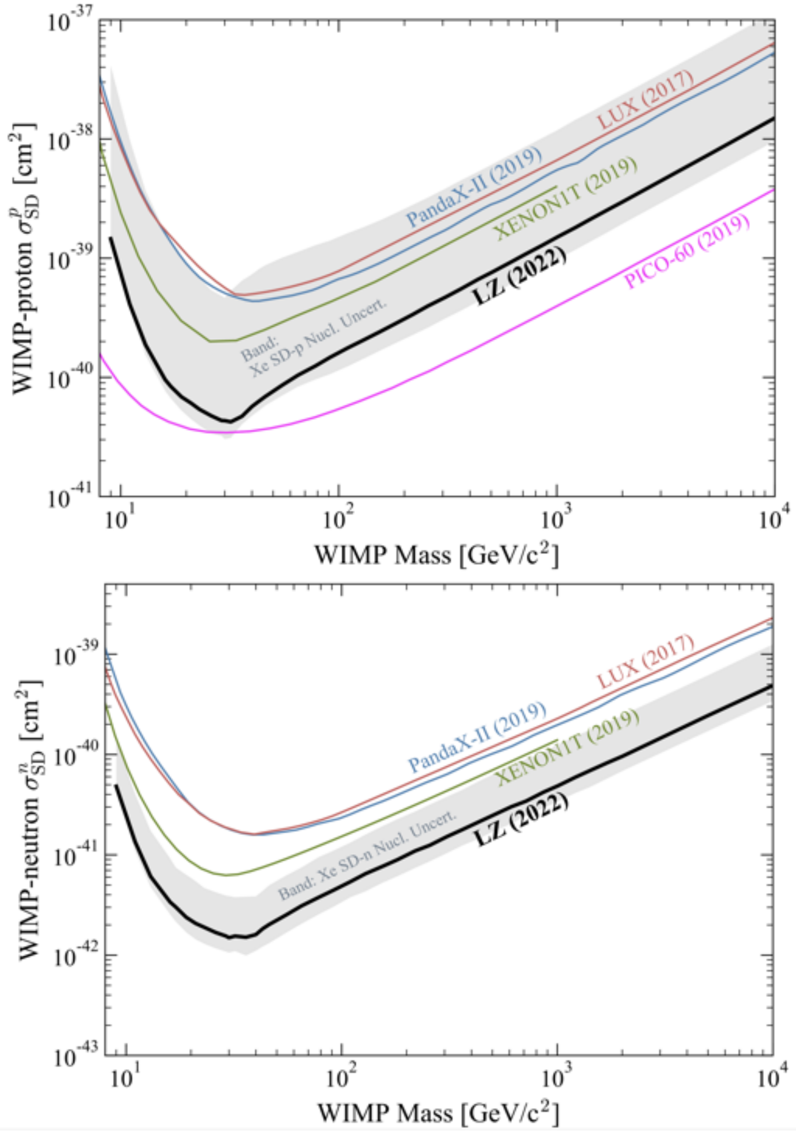


Figure 3.2: The SD WIMP-proton (top) and WIMP-neutron (bottom) cross section 90% confidence limits (black) as a function of WIMP mass, with uncertainty bands (gray) resulting from xenon nuclear correction factors. Figure from Ref. [5].

The general Lagrangian density for the elastic scattering process $\chi + N \to \chi + N$ is

$$\mathcal{L}_{\text{int}} = \sum_{I=\{\text{s, v}\}} \sum_{i} c_i^I O_i \chi\chi NN, \tag{3.1}$$

where sums are performed over isospin conserving and violating interactions, over all the possible operators O_i, and where χ is the dark matter field and N is the nucleon field. In the isospin basis, interactions are classified according to whether the WIMP couples equally to protons and neutrons (isoscalar) or not (isospin-violating, or isovector). There is an alternate basis, the (p,n) basis (adopted in the WS result), that projects the coupling strength along WIMP-proton and WIMP-neutron axes. Both bases are equivalent and are related with the relations

$$c_{\text{s}} = \frac{1}{2}(c_p + c_n), \tag{3.2}$$

$$c_{\text{v}} = \frac{1}{2}(c_p - c_n), \tag{3.3}$$

where c_p and c_n are the proton and neutron coupling strengths, respectively. The isospin basis is used in this work since isospin is an approximately conserved symmetry in xenon nuclei [109]. Further, the isospin basis facilitates comparisons with the results of experiments using nuclei with varying amounts of isospin violation [110].

The operators O_i are built using four non-relativistic, Hermitian three-vectors. These quantities are

$$i\vec{q},\ \vec{v}^{\perp} = \vec{v} + \frac{\vec{q}}{2\mu_N},\ \vec{S}_{\chi},\ \vec{S}_N, \tag{3.4}$$

where $\vec{q}$ is the momentum transfer, $\vec{v}$ is the relative velocity of the WIMP, μ is the reduced mass of the WIMP-nucleon system, and where S_{χ} and S_N are the spins of the WIMP and the nucleon, respectively. The quantity $\vec{v}^{\perp}$ is constructed using the Hermitian quantity $i\vec{q}$ to be invariant under the exchange of incoming and outgoing particles. Using these variables, the condition for energy conservation is $\vec{q} \cdot \vec{v}^{\perp} = 0$.

Up to the second power of q, the following non-relativistic operators make up the EFT. They are

$$
\begin{aligned}
O_1 &= 1_\chi 1_N \\
O_2 &= \left(v^\perp\right)^2 \\
O_3 &= i\vec{S}_N \cdot \left(\frac{\vec{q}}{m_N} \times \vec{v}^\perp\right) \\
O_4 &= \vec{S}_\chi \cdot \vec{S}_N \\
O_5 &= i\vec{S}_\chi \cdot \left(\frac{\vec{q}}{m_N} \times \vec{v}^\perp\right) \\
O_6 &= \left(\vec{S}_\chi \cdot \frac{\vec{q}}{m_N}\right)\left(\vec{S}_N \cdot \frac{\vec{q}}{m_N}\right) \\
O_7 &= \vec{S}_N \cdot \vec{v}^\perp \\
O_8 &= \vec{S}_\chi \cdot \vec{v}^\perp \\
O_9 &= i\vec{S}_\chi \cdot \left(\vec{S}_N \times \frac{\vec{q}}{m_N}\right) \\
O_{10} &= i\vec{S}_N \cdot \frac{\vec{q}}{m_N} \\
O_{11} &= i\vec{S}_\chi \cdot \frac{\vec{q}}{m_N} \\
O_{12} &= \vec{S}_\chi \cdot \left(\vec{S}_N \times \vec{v}^\perp\right) \\
O_{13} &= i\left(\vec{S}_\chi \cdot \vec{v}^\perp\right)\left(\vec{S}_N \cdot \frac{\vec{q}}{m_N}\right) \\
O_{14} &= i\left(\vec{S}_\chi \cdot \frac{\vec{q}}{m_N}\right)\left(\vec{S}_N \cdot \vec{v}^\perp\right) \\
O_{15} &= -\left(\vec{S}_\chi \cdot \frac{\vec{q}}{m_N}\right)\left(\left(\vec{S}_N \times \vec{v}^\perp\right) \cdot \frac{\vec{q}}{m_N}\right).
\end{aligned}
$$

Each operator represents a possible interaction in the non-relativistic limit of any elastic WIMP–nucleon scattering processes. The exception to this is O_2, which contains two powers of $\vec{v}^\perp$, and cannot result as the leading order non-relativistic reduction of a covariant operator. It is therefore omitted.

3.1.3 Nuclear embedding

The WIMP–nucleon operators have to be embedded in a nucleus and integrated over the nucleons in order to obtain the interaction cross section and rate. In the standard SI and SD calculation of the WIMP–nucleus scattering rate, the transition probabilities with the nuclear charge and spin current densities are summed over bound nucleons. The nuclear responses elicited are the vector charge (M) operator for the SI case, and combinations of the axial longitudinal (Σ'') and axial transverse (Σ') electric operators for the SD case [109].

It is assumed that the WIMP–nucleus matrix element is obtained by summing the WIMP–nucleon matrix element over all the nucleons. The EFT takes into account the velocity parameter $\vec{v}^{\perp}$, which is a function of the relative incoming velocity $\vec{v}$ and the internal nucleon velocities (given by $\vec{q}/\mu_N$) [109]. The latter quantity vanishes in the point nucleus approximation, but is responsible for a parametric enhancement of the cross sections when accounted for. While nucleon velocities are odd under parity, which is a respected symmetry in nuclear ground states, they may pair with spatial operators such as $e^{-i\vec{q}\cdot\vec{x}}$ to produce parity-even interactions. Three novel nuclear responses, Δ, Φ', and Φ'', arise due to the compositeness of the nucleus that do not arise in general from SI and SD interactions. Therefore the general amplitude of a WIMP-nucleon interaction contains all six nuclear responses, which reduces to the SI/SD case with only M, Σ', and Σ'' in the limit $q^2 \to 0$ and $\vec{v}^{\perp 2} \to 0$.

3.1.4 Non-relativistic matching

The NREFT can be obtained from relativistic WIMP-nucleon interactions. In general, there are two Lorentz-covariant scalar bilinears, given by $\mathbf{1}$ and γ^5, and four vector bilinears, given by P^{μ}, $P^{\mu}\gamma^5$, $i\sigma^{\mu\nu}q_{\nu}$, and $\gamma^{\mu}\gamma^5$. Therefore, there is a total of $2^2 + 4^4 = 20$ possible products giving a scalar Lagrangian density. Each of these interactions could be reduced to the non-relativistic limit, and shown to be linear combinations of the operators O_i presented above. Instead of constructing the

EFT this way, the four non-relativistic degrees of freedom (momentum, velocity, nucleon and WIMP spins) allow a simpler enumeration of the possible operators. However, as an example, the non-relativistic reduction of the axial operator γ^5 is presented below.

The operator γ^5 can be expressed in terms of the incoming and outgoing momenta using the Gordon identity (from Ref. [111])

$$\bar{u}(p')\gamma^\mu u(p) = \bar{u}(p')\left[\frac{p'^\mu + p^\mu}{2m} + \frac{i\sigma^{\mu\nu}q_\nu}{2m}\right]u(p), \tag{3.5}$$

which can be modified to obtain the modified Gordon identity for the axial-vector interaction

$$\bar{u}(p')\gamma^\mu\gamma^5 u(p) = \bar{u}(p')\left[\frac{\gamma^5 q^\mu}{2m} + \frac{i\gamma^5\sigma^{\mu\nu}(p'+p)_\nu}{2m}\right]u(p). \tag{3.6}$$

The identity required for the axial interaction can then be obtained using the Dirac equation

$$(\gamma^\mu p_\mu - m)u(p) = 0, \tag{3.7}$$

which gives

$$\bar{u}(p')\gamma^5 u(p) = \frac{p_\mu}{m}\bar{u}(p')\gamma^5\gamma^\mu u(p). \tag{3.8}$$

The axial-Gordon identity is therefore

$$\bar{u}(p')\gamma^5 u(p) = \bar{u}(p')\left[\frac{\gamma^5 q\cdot p}{2m^2} + \frac{i\gamma^5\sigma^{\mu\nu}(p'+p)_\nu p_\mu}{2m^2}\right]u(p). \tag{3.9}$$

The first term on the right hand side vanishes because

$$q \cdot p = (p' - p) \cdot p \tag{3.10}$$

$$= p' \cdot p - p^2 \tag{3.11}$$

$$= (m^2 - \vec{p}' \cdot \vec{p}) - m^2 \tag{3.12}$$

$$= -\vec{p}' \cdot \vec{p} \ll m^2. \tag{3.13}$$

The remaining term can be written in terms of its temporal and spatial indices to obtain

$$\gamma^5 \sigma^{\mu\nu} p_\mu (p' + p)_\nu = \gamma^5 \left[\sigma^{0\nu} m (p' + p)_\nu + \sigma^{i\nu} p_i (p' + p)_\nu\right] \tag{3.14}$$

$$= \gamma^5 \left[\sigma^{00} m(2m) + \sigma^{0j} m (p' + p)_j + \sigma^{i0} p_i (p' + p)_0 + \sigma^{ij} p_i (p' + p)_j\right] \tag{3.15}$$

$$= \gamma^5 \left[\sigma^{0i} m (p' + p - 2p)_i + \sigma^{ij} p_i (p' + p)_j\right] \tag{3.16}$$

$$= \gamma^5 \left[\sigma^{0i} m q_i + \sigma^{ij} p_i (p' + p)_j\right], \tag{3.17}$$

where $\sigma^{00} = 0$, and momentum conservation relation $p' - p = q$ was invoked. The spatial term in the relation above rotates the momenta since $\sigma^{ij} = \varepsilon^{ijk} R^k$, where R^k is the generator of the rotation group, and ε is the Levi-Civita symbol. The result of the rotations is $\sigma^{ij} = R^i \cdot (\vec{p} \times (\vec{p} + \vec{p'}))_i \ll m^2$, so this term can be discarded. After these reductions, the axial-Gordon identity is

$$\bar{u}(p') \gamma^5 u(p) = \bar{u}(p') \frac{i}{2m^2} \gamma^5 \sigma^{0i} m q_i u(p). \tag{3.18}$$

In the chiral basis,

$$\sigma^{0i} = -i \begin{pmatrix} \boldsymbol{\sigma}^i & 0 \\ 0 & -\boldsymbol{\sigma}^i \end{pmatrix}, \tag{3.19}$$

where the $\boldsymbol{\sigma}^i$ are Pauli matrices. Since $\gamma^5 = \left(\begin{smallmatrix} -\mathbf{1} & 0 \\ 0 & \mathbf{1} \end{smallmatrix}\right)$, the product

$$\gamma^5 \sigma^{0i} = i \begin{pmatrix} \boldsymbol{\sigma}^i & 0 \\ 0 & \boldsymbol{\sigma}^i \end{pmatrix} = 2iS^i, \tag{3.20}$$

where S^i is the spin operator. Finally, the desired non-relativistic reduction for the axial operator is

$$\bar{u}(p')\gamma^5 u(p) = \bar{u}(p') \left[\frac{i\vec{q} \cdot \vec{S}}{m} \right] u(p). \tag{3.21}$$

A list of non-relativistic reductions from interaction Lagrangians is shown in Table 3.1, adapted from Ref. [108].

3.1.5 Recoil Spectra

The energy spectra of the EFT operators can be calculated from their respective matrix elements in an identical manner to Chapter 1. Starting with the integral over phase space

$$d\sigma = \frac{m_\chi^2}{v E_\chi^i E_\chi^f} \frac{m_T}{E_T^f} \left[\frac{1}{2j_\chi + 1} \frac{1}{2j_N + 1} \sum_{\text{spins}} |\mathcal{M}|^2 \right] \frac{d^3 p'}{(2\pi)^3} \frac{d^3 k'}{(2\pi)^3} (2\pi)^4 \delta^4(p + k - p' - k'), \tag{3.22}$$

where $\mathcal{M}$ is the Galilean-invariant matrix element, the differential cross section can be expressed as

$$\frac{d\sigma(v, E_R)}{dE_R} = 2m_T \frac{1}{4\pi v^2} \left[\frac{1}{2j_\chi + 1} \frac{1}{2j_N + 1} \sum_{\text{spins}} |\mathcal{M}|^2 \right]. \tag{3.23}$$

Finally, this rate can be integrated over the galactic WIMP velocity distribution to obtain

$$\frac{dR_D}{dE_R} = N_T n_\chi \int_{v > v_{\min}} \frac{d\sigma(v, E_R)}{dE_R} v f_E(\vec{v}) d^3 v, \tag{3.24}$$

where N_T is the number of atoms in the detector, n_χ is the local number density of dark matter, and f_E is their velocity distribution.

$\mathcal{L}$	Non-relativistic reduction	$\sum_i c_i \mathcal{O}_i$
$\bar{\chi}\chi\bar{N}N$	$1_\chi 1_N$	$\mathcal{O}_1$
$i\bar{\chi}\chi\bar{N}\gamma^5 N$	$i\frac{\vec{q}}{m_N}\cdot\vec{S}_N$	$\mathcal{O}_{10}$
$i\bar{\chi}\gamma^5\chi\bar{N}N$	$-i\frac{\vec{q}}{m_\chi}\cdot\vec{S}_\chi$	$-\frac{m_N}{m_\chi}\mathcal{O}_{11}$
$i\bar{\chi}\gamma^5\chi\bar{N}\gamma^5 N$	$-\left[\frac{\vec{q}}{m_\chi}\cdot\vec{S}_\chi\right]\left[\frac{\vec{q}}{m_N}\cdot\vec{S}_N\right]$	$-\frac{m_N}{m_\chi}\mathcal{O}_6$
$\frac{P^\mu}{m_M}\bar{\chi}\chi\frac{K_\mu}{m_M}\bar{N}N$	$\frac{4m_\chi m_N}{m_M^2}1_\chi 1_N$	$\frac{4m_\chi m_N}{m_M^2}\mathcal{O}_1$
$\frac{P^\mu}{m_M}\bar{\chi}\chi\bar{N}i\sigma_{\mu\alpha}\frac{q^\alpha}{m_M}N$	$-\frac{m_\chi}{m_N}\frac{\vec{q}^2}{m_M^2}1_\chi 1_N - 4i\frac{m_\chi}{m_M}\vec{v}^\perp\cdot\left(\frac{\vec{q}}{m_M}\times\vec{S}_N\right)$	$-\frac{m_\chi}{m_N}\frac{\vec{q}^2}{m_M^2}\mathcal{O}_1 + \frac{4m_\chi m_N}{m_M^2}\mathcal{O}_3$
$\frac{P^\mu}{m_M}\bar{\chi}\chi\bar{N}\gamma_\mu\gamma^5 N$	$-\frac{4m_\chi}{m_M}\vec{v}^\perp\cdot\vec{S}_N$	$-\frac{4m_\chi}{m_M}\mathcal{O}_7$
$i\frac{P^\mu}{m_M}\bar{\chi}\chi\frac{K_\mu}{m_M}\bar{N}\gamma^5 N$	$\frac{4im_\chi}{m_M}\frac{\vec{q}}{m_M}\cdot\vec{S}_N$	$\frac{4m_\chi m_N}{m_M^2}\mathcal{O}_{10}$
$\bar{\chi}i\sigma^{\mu\nu}\frac{q_\nu}{m_M}\chi\frac{K_\mu}{m_M}\bar{N}N$	$\frac{m_\chi}{m_N}\frac{\vec{q}^2}{m_M^2}1_\chi 1_N + 4i\frac{m_N}{m_M}\vec{v}^\perp\cdot\left(\frac{\vec{q}}{m_M}\times\vec{S}_\chi\right)$	$\frac{m_N}{m_\chi}\frac{\vec{q}^2}{m_M^2}\mathcal{O}_1 - 4\frac{m_N^2}{m_M^2}\mathcal{O}_5$
$\bar{\chi}i\sigma^{\mu\nu}\frac{q_\nu}{m_M}\chi\bar{N}i\sigma_\mu\alpha\frac{q^\alpha}{m_M}N$	$4\left(\frac{\vec{q}}{m_M}\times\vec{S}_\chi\right)\cdot\left(\frac{\vec{q}}{m_M}\times\vec{S}_N\right)$	$4\left(\frac{\vec{q}^2}{m_M^2}\mathcal{O}_4 - \frac{m_N^2}{m_M^2}\mathcal{O}_6\right)$
$\bar{\chi}i\sigma^{\mu\nu}\frac{q_\nu}{m_M}\chi\bar{N}\gamma^\mu\gamma^5 N$	$-4i\left(\frac{\vec{q}}{m_M}\times\vec{S}_\chi\right)\cdot\vec{S}_N$	$-4\frac{m_N}{m_M}\mathcal{O}_9$
$i\bar{\chi}i\sigma^{\mu\nu}\frac{q_\nu}{m_M}\chi\frac{K_\mu}{m_M}\bar{N}\gamma^5 N$	$\left[i\frac{\vec{q}^2}{m_\chi m_M} - 4\vec{v}^\perp\cdot\left(\frac{\vec{q}}{m_M}\times\vec{S}_\chi\right)\right]\frac{\vec{q}}{m_M}\cdot\vec{S}_N$	$\frac{m_N\vec{q}^2}{m_\chi m_M^2}\mathcal{O}_{10} + \frac{4\vec{q}^2}{m_M^2}\mathcal{O}_{12} + \frac{4m_N^2}{m_M^2}\mathcal{O}_{15}$
$\bar{\chi}\gamma^\mu\gamma^5\chi\frac{K_\mu}{m_M}\bar{N}N$	$4\frac{m_N}{m_M}\vec{v}^\perp\cdot\vec{S}_\chi$	$4\frac{m_N}{m_M}\mathcal{O}_8$
$\bar{\chi}\gamma^\mu\gamma^5\chi\bar{N}\gamma_\mu\gamma^5 N$	$-4\vec{S}_\chi\cdot\vec{S}_N$	$-4\mathcal{O}_4$
$i\bar{\chi}\gamma^\mu\gamma^5\chi\bar{N}\frac{K_\mu}{m_M}\gamma^5 N$	$-4i\vec{v}^\perp\cdot\vec{S}_\chi\frac{\vec{q}}{m_M}\cdot\vec{S}_N$	$4\frac{m_N}{m_M}\mathcal{O}_{13}$
$i\frac{P^\mu}{m_M}\bar{\chi}\gamma^5\chi\frac{K_\mu}{m_M}\bar{N}N$	$-4i\frac{m_N}{m_M}\frac{\vec{q}}{m_M}\cdot\vec{S}_\chi\frac{\vec{q}}{m_M}\cdot\vec{S}_N$	$-4\frac{m_N^2}{m_M^2}\mathcal{O}_{11}$
$i\frac{P^\mu}{m_M}\bar{\chi}\gamma^5\chi\bar{N}i\sigma_{\mu\alpha}\frac{q^\alpha}{m_M}N$	$\frac{\vec{q}}{m_M}\cdot\vec{S}_\chi\left[i\frac{\vec{q}^2}{m_N m_M} - 4\vec{v}^\perp\cdot\left(\frac{\vec{q}}{m_M}\times\vec{S}_N\right)\right]$	$\frac{\vec{q}^2}{m_M^2}\mathcal{O}_{11} + 4\frac{m_N^2}{m_M^2}\mathcal{O}_{15}$
$i\frac{P^\mu}{m_M}\bar{\chi}\gamma^5\chi\bar{N}\gamma_\mu\gamma^5 N$	$4i\frac{\vec{q}}{m_M}\cdot\vec{S}_\chi\vec{v}^\perp\cdot\vec{S}_N$	$4\frac{m_N}{m_M}\mathcal{O}_{14}$
$\frac{P^\mu}{m_M}\bar{\chi}\gamma^5\chi\frac{K_\mu}{m_M}\bar{N}\mu\gamma^5 N$	$-4\frac{\vec{q}}{m_M}\cdot\vec{S}_\chi\frac{\vec{q}}{m_M}\cdot\vec{S}_N$	$-4\frac{m_N^2}{m_M^2}\mathcal{O}_6$

Table 3.1: List of relativistic Lagrangian densities, and the corresponding non-relativistic reduction.

The WimPyDD package was used to calculate the nuclear recoil spectra for the operators tested in this analysis [112]. Figure 3.3 shows the recoil spectra for the 14 isoscalar operators. The importance of a large energy window is demonstrated in Figure 3.4, which compares the integrated rates of O_1 and O_6 as a function of the high energy cutoff. While a high energy cutoff of 50 $\mathrm{keV_{nr}}$ is sufficient to capture all O_1 events, it barely captures any events of a 200 GeV WIMP with an O_6 interaction.

3.2 Detector conditions and response

3.2.1 First science run

The data used in the SR1 analyses constitute 60 ± 1 live-days, called the livetime of the search, which is defined as the time in which an interaction could have been recorded and reconstructed as a single scatter, after subtracting periods for calibrations, detector maintenance, dead time of the DAQ system, and excessively high trigger rates [5]. These data were collected from 23 December 2021 to 11 May 2022. The EFT analysis used the exact same set of runs as the WS, and so the detector conditions given below are the same as reported in Ref. [5].

The drift field was 193 kV/cm, while the extraction field in the gaseous xenon was 7.3 kV/cm. The temperature and pressure in the TPC were 174.1 K and 1.791 bar, respectively, and were found to be stable to within 0.2 %. Xenon was continuously purified at a rate of 3.3 t/day, and the electron lifetime was measured to be between 5000 μs and 8000 μs, which far exceeds the maximum drift time of 951 μs.

3.2.2 Calibrations

No additional source calibration data were obtained to model the detector response at high energies. In the SR1 WS, the TPC ER response was tuned by Greg Rischbieter to match the median

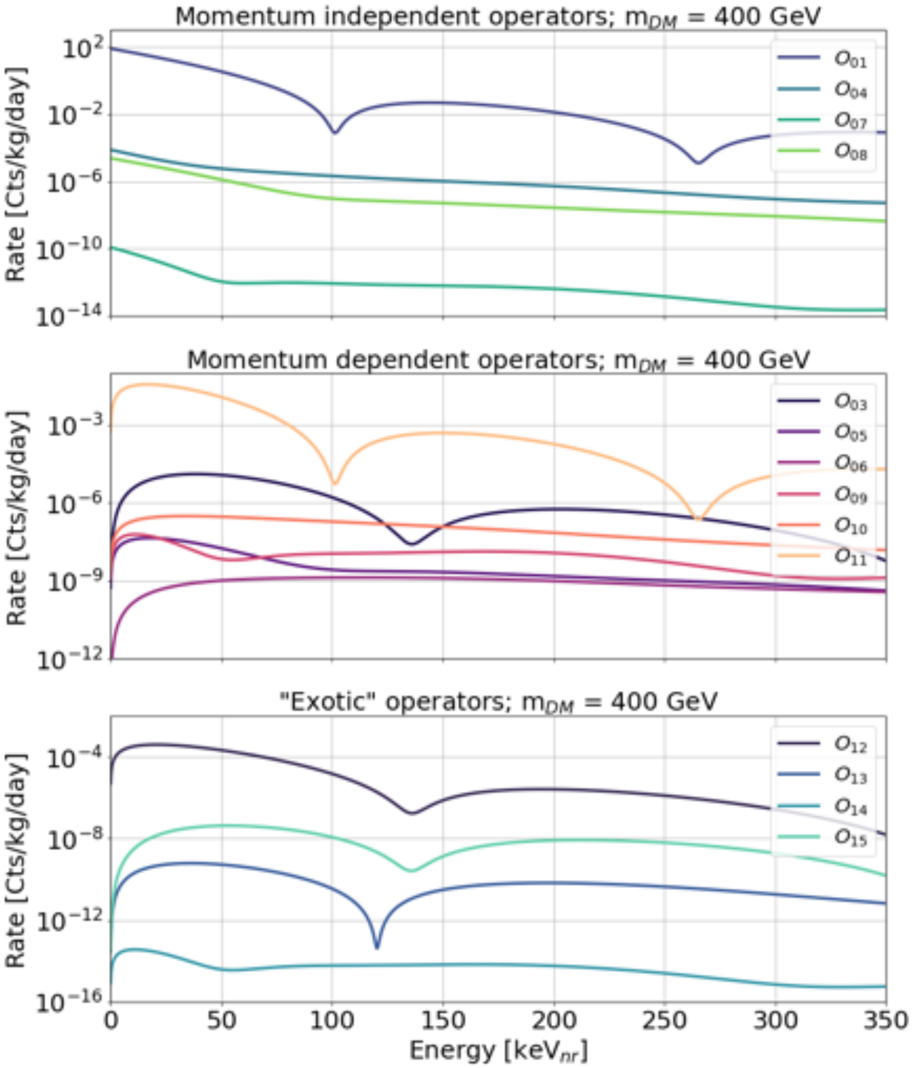


Figure 3.3: Differential event rates as a function of nuclear recoil energy deposited for 14 isoscalar EFT operators. The momentum-independent operators are shown in the top panel, momentum-dependent operators (with vanishing rate at zero energy) shown in the middle panel, while the bottom panel shows operators that cannot arise from a bosonic mediator but are presented for completeness.

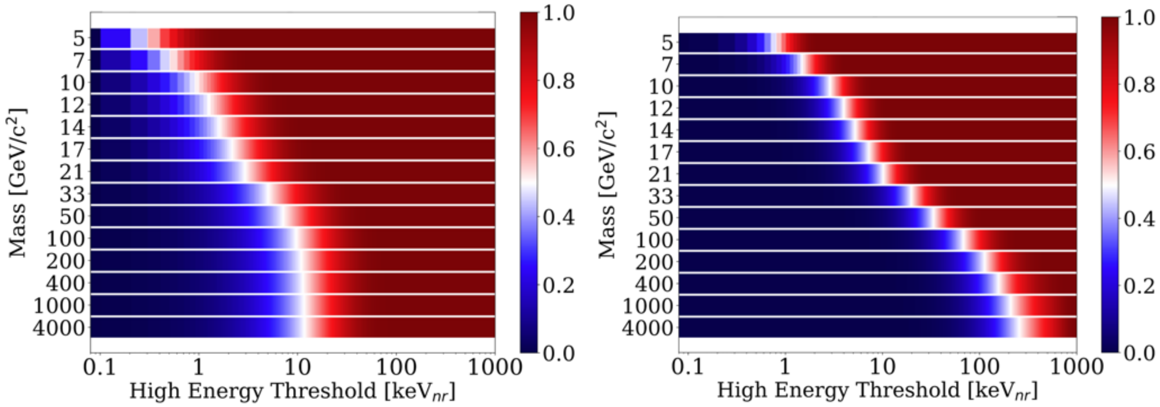


Figure 3.4: Integrated rates of $\mathcal{O}_1$ (left) and $\mathcal{O}_6$ (right) interactions at all WIMP masses considered, as a function of high energy cutoff. The z-axis color indicates the fraction of the total integrated rate below the cutoff energy.

and widths of the tritium calibration data, which diminishes in rate when S1c $>$ 100 phd. For the SR1 EFT search, the high energy continuation of the ER band was tuned to also match the ^{220}Rn calibration data, which consists of single scatter naked β decay events. The effects of any mismodeling manifesting as as ER leakage are negligible, as the NR and ER band are increasingly distant at higher energies. Figure 3.5 show the tritium, D-D neutron, and ^{212}Pb events used for calibration in SR1.

An ex-situ neutron calibration by the Livermore group provided data for the LXe light and charge yields up to 426 keV$_{\text{nr}}$ [113]. These measurements were found to be consistent with the extrapolation of the NEST NR model at the sub-percent level.

3.3 Analysis

A series of cuts was developed by the LZ collaboration for the WS analysis, and is reported in Refs. [5, 82]. These cuts targeted the WIMP search analysis region bounded in 3 phd $<$ S1c $<$ 80 phd. The EFT search extends this region of interest to be 3 phd $<$ S1c $<$ 600 phd, while placing a cut of $\log \text{S2c} < 4.5$ to remove ER events that are not useful to the search. The core cuts developed for the standard WIMP search are used for the EFT search. The main cuts in the

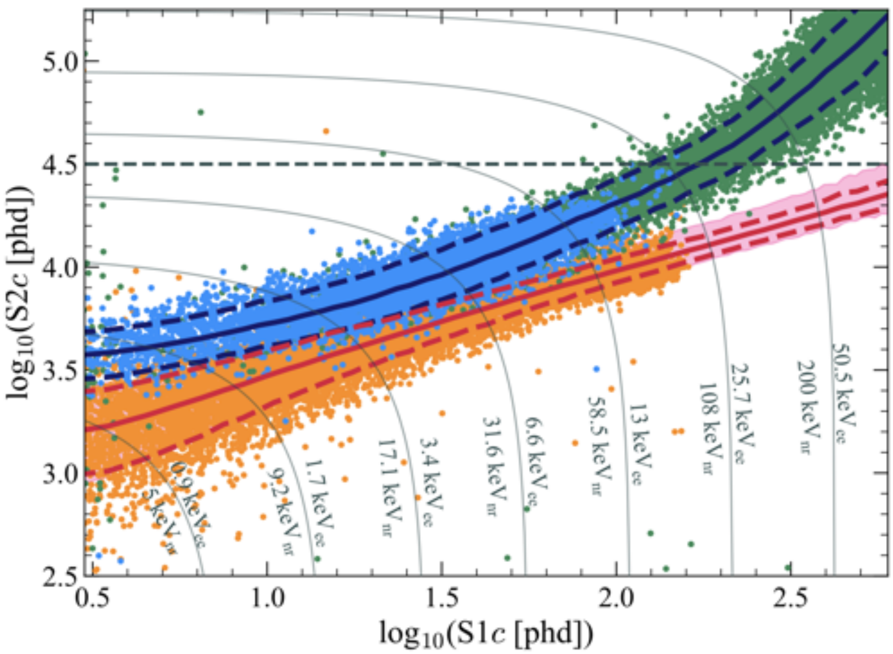


Figure 3.5: Calibration events from tritium (blue), D-D neutrons (orange), and ^{212}Pb (green) used to tune the LZ detector for the SR1 EFT search. The tuned NEST responses are shown with dark blue and red lines for ER and NR interactions, respectively, with the dashed lines corresponding to 90–10% confidence levels. The NR band was shifted from -1σ to 3σ (pink region) beyond the D-D energy endpoint to visually indicate the effect of any mismodelling. Plot created by Greg Rischbieter.

core cut sequence remove data with issues formed in the signal train, relating to the PMTs, DAQ, and HV systems, in addition to removing periods of time following a high energy event like a muon passing through the TPC. Data taken in a period of time after an event with a large S2 are also removed, as small pulses usually follow substantial ionization events. A data-driven fiducial cut is applied to remove high activity near the TPC walls, cathode and anode. Complementary to the fiducial cut are the skin and OD vetos, which remove multiply scattering γ rays and neutrons that scatter once in the TPC. Further, cuts based on pulse shapes and position are deployed to remove accidental backgrounds resulting from the random coincidence of S1–S2 pairs.

A major modification to the cut sequence in the EFT search is the inclusion of the γ-X cut, designed to remove a multiple scattering γ background that directly overlaps the high energy NR region. The development of the γ-X cut is the subject of Chapter 4. Figure 3.6 shows the positions of all events surviving the EFT cut chain, numbering 835, along with events removed by the skin veto, the OD veto, and the γ-X cut. Figure 3.7 shows the same set of data in S1c-$\log_{10}$S2c space.

The NR acceptance efficiency of the cut chain is shown in Figure 3.8, which classifies the efficiency of cuts on the region of interest, data quality, single scatter reconstruction, and γ-X.

3.3.1 Backgrounds

The EFT background model consists of 11 components, most of which are extensions of the models developed for the WS. Two primary departures from the WS background model are

1. the separation of the detector ER component from the flat ER, due to a few γ events from detector components with suppressed S2,
2. the inclusion of an ^{125}I background component.

The expectation and fit results of the background model to the data are shown in Table 3.2. Each background component is discussed below.

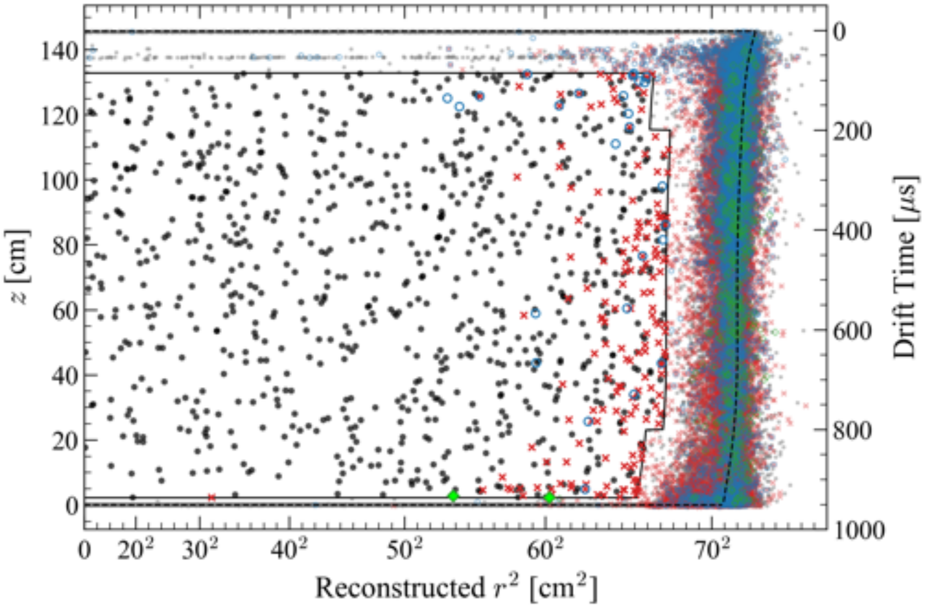


Figure 3.6: Positions of events passing all EFT cuts (black points) in the fiducial volume, events vetoed by the skin (red crosses), events vetoed by the OD (blue circles), and events removed by the γ-X cut (green diamonds). The fiducial volume is bounded within the black line, and events outside it that are removed by and survive the above three cuts are also shown. The concentration of events outside the fiducial volume correspond to the reconstructed position of the TPC wall.

Source	Expected Events	Fit Result
Flat ER	517.4 ± 82.8	604.6 ± 30.0
Detector ER	18.4 ± 9.2	22.3 ± 8.1
ν ER	55.3 ± 5.5	54.8 ± 5.5
^{124}Xe	8.2 ± 2.0	9.5 ± 2.0
^{127}Xe	20.5 ± 1.8	21.1 ± 1.8
^{136}Xe	55.1 ± 11.6	57.4 ± 11.2
^{125}I	8.9 ± 4.4	1.9 ± 4.2
^{8}B CEνNS	0.14 ± 0.01	0.14 ± 0.01
Accidentals	1.3 ± 0.3	1.3 ± 0.03
Subtotal	-	773.0 ± 34.2
^{37}Ar	[0, 288]	48.1 ± 9.4
Detector NR	$0.0^{+0.5}$	$0.0^{+1.8}$
Total	-	821.1 ± 35.3

Table 3.2: Number of events from sources in the 60 d x 5.5 t exposure. The middle column shows the expected numbers of background events, with error bars used to constrain the nuisance parameters in the statistical fit.

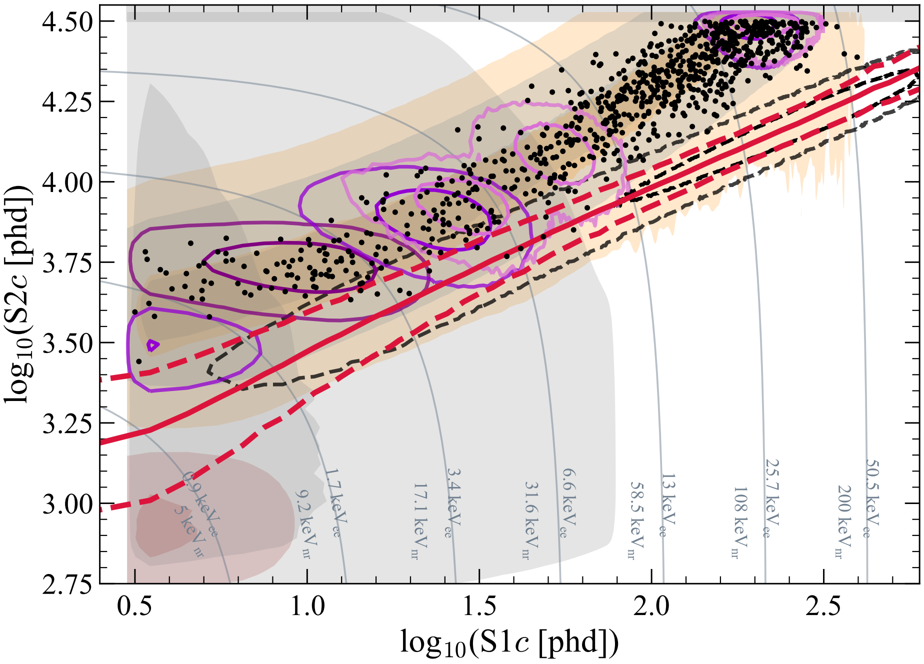


Figure 3.7: The SR1 EFT search data (black points) in the S1c-$\log_{10}$S2c space after all cuts. Also shown are the 1σ and 2σ expectations of the flat ER and accidentals background (gray shaded), the detector ER (orange shaded), monoenergetic peaks from ^{37}Ar and activated xenon products (purple contours), and ^{8}B CEνNS (red). The median NR band is shown in the solid red line, while the dotted red lines show the 90–10% confidence levels. The black dashed line shows the expectation for an $\mathcal{O}_6$ WIMP of mass 1,000 GeV. Plot created by Billy Boxer.

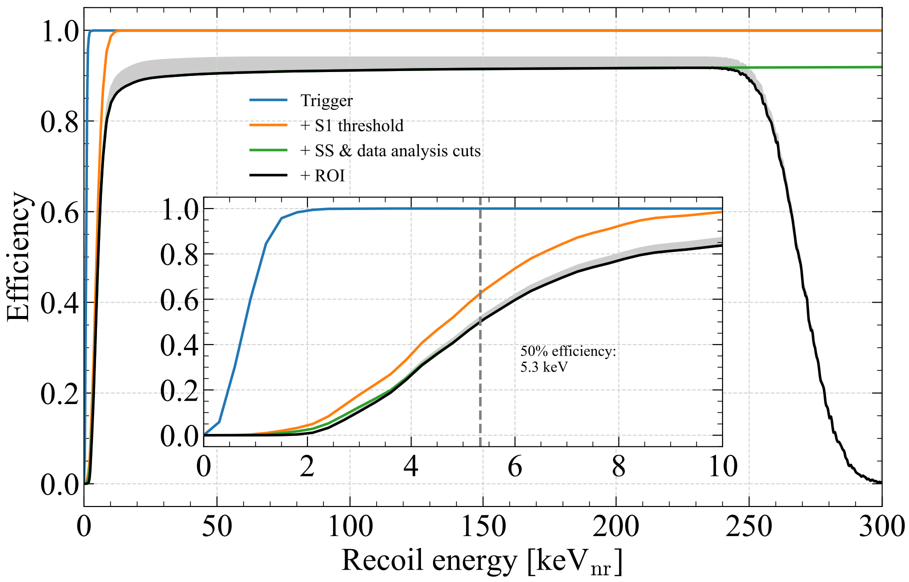


Figure 3.8: NR acceptance efficiency as a function of NR energy for the trigger (blue), the 3-fold PMT coincidence requirement and the 3 phd threshold (orange), single scatter reconstruction and analysis cuts (green), and the EFT search ROI (black). The inset shows the NR acceptance at low energies, with the dotted line at 5.3 keV indicating the 50% efficiency. The error band (gray) is evaluated using AmLi and tritium calibration data.

Flat ER – Continuous β emissions by radioactive sources dissolved in LXe produce flat energy spectra that are aggregated as a flat ER background component. The largest contribution is from the naked (unaccompanied by a γ ray) β decay of ^{214}Pb, which is a daughter of ^{222}Rn. Other sources of naked β decays are ^{212}Pb produced in the ^{220}Rn chain, and ^{85}Kr, a trace element in natural xenon. The concentrations of ^{85}Kr and ^{39}Ar (found to be negligible for the SR1 exposure) were reduced via charcoal chromatography at SLAC.

Detector ER – γ radiation from detector components and the cavern rock Compton scatter in the TPC to produce an ER background. The sources of γ rays are the trace primordial elements ^{60}Co, ^{40}K, ^{238}U, and ^{232}Th present in detector components, in addition to ^{238}U, ^{232}Th, and ^{40}K from the cavern walls. This background component was modeled separately from the flat ER to account for events below the ER band that seem to suffer from charge loss to the TPC walls.

ν *ER* – Electroweak interactions of solar neutrinos with atomic electrons of the xenon produce low energy ER backgrounds. The resulting ER energy spectrum is flat in the EFT region of interest. Neutrinos can also undergo coherent elastic ν-nucleon scattering (CEνNS), producing NR events. The largest contribution to the CEνNS process is from ^{8}B neutrinos, which only create 0.14 events in the EFT ROI.

^{124}*Xe* – This isotope undergoes double electron capture, first measured in the XENON experiment [114]. The combination of K, L, and M-shell decays produce ER events in the EFT region ROI.

^{127}*Xe* – The electron capture decay of ^{127}Xe produces x-ray deposits that are prevalent for both the WS and the EFT search. Further, ^{127}Xe decays near the boundaries of the detector produce events with an NR-like S1/S2 ratio, known as γ-X events.

^{136}*Xe* – The double β decay of this isotope produces a continuum ER spectrum. While a background for the WS and EFT searches, ^{136}Xe is an important candidate to observe the 0ν2β decay process.

^{125}*I* – A new addition to the background model of SR1, ^{125}I decays via electron capture with a γ ray of 35.5 keV, in addition to x-rays. This energy was too large to affect the WS, but the extended energy range of the EFT search necessitates accounting for ^{125}I. This isotope is produced by the neutron activation of ^{124}Xe to ^{125}Xe, which subsequently inverse β decays to ^{125}I.

^{8}*B CEνNS* – Produced by the $^8B \rightarrow ^8\mathrm{Be} + e^+\nu_e$ reaction in the sun, ^8B neutrinos have energies up to 15 MeV. They are the only CEνNS background in the WS and EFT search, dominant over the rates of the diffuse supernova (DSN) and atmospheric neutrinos.

^{37}*Ar* – The electron capture decay of ^{37}Ar produces a 2.82 keV K-shell x-ray that made up a significant portion of the SR1 backgrounds. The ^{37}Ar was produced by cosmogenic spallation during the transport and storage of xenon on the surface of Earth [115]. The importance of this isotope is diminished for future science runs due to its 35 day half life.

Accidentals – The random coincidence of S1 and S2 pulses in an event window produce this background, which is prevalent at the low energy side of the energy window. Accidental backgrounds arise from various sources from various components of the detector, and a data driven approach was taken to model this component of the background model.

Detector NR – The sources of neutrons in LZ are the spontaneous fission of ^{238}U and the (α, n) reactions of lighter nuclei in detector components. The OD veto enables the identification of neutron events, and was used to predict the absence of neutrons in the WS and EFT search regions of interest.

3.4 Results

No evidence of an excess is found among the EFT operators in the isospin basis for both elastic and inelastic WIMP interactions. The details of the statistical test are given in subsection 3.4.1. A total of 56 relevant models were tested within the extended energy window. Upper limits on the coupling strengths for each EFT interaction are presented in subsection 3.4.2 for elastic scatters and subsection 3.4.3 for DM upscattering to a heavier state.

3.4.1 Statistical method

The search for dark matter using LZ data is formalized as a frequentist statistical test, known as the profile likelihood ratio (PLR) method. When the test is performed for the discovery of a new signal, the null hypothesis to be rejected H_0^D describes only background data, while the alternative hypothesis H_1^D includes both background and signal [116]. When setting upper bounds on the WIMP–nucleon cross section, as performed here, the null hypothesis H_0^L is a model with signal and background, and the alternative hypothesis H_1^L is background-only.

The likelihoods of the two hypotheses given the data can be quantified and compared using a test statistic. One advantage of the PLR is that all experimental uncertainties are taken into account in the test statistic as nuisance parameters $\boldsymbol{\theta}$. The quantity known as the PLR quantifies the agreement between the data and the hypotheses, and expresses the agreement as the ratio

$$\lambda(\mu) = \frac{\mathcal{L}(\mu, \hat{\boldsymbol{\theta}})}{\mathcal{L}(\hat{\mu}, \hat{\boldsymbol{\theta}})}, \tag{3.25}$$

where $\mathcal{L}$ is the likelihood function given below, and μ represents the signal strength. The variables $\hat{\mu}$ and $\hat{\boldsymbol{\theta}}$ maximize the likelihood, while $\hat{\boldsymbol{\theta}}$ maximizes the likelihood conditionally for a given signal strength μ. The ratio $\lambda(\mu)$ is close to 1 for a good agreement between the data and the given value of μ.

The likelihood function $\mathcal{L}$ is the probability of observing the data as a function of the model, parameterized by the signal strength μ and nuisance parameters $\boldsymbol{\theta}$. The LZ likelihood function is

$$\mathcal{L} = \left[\mathrm{Pois}(n_{\mathrm{obs}}|\mu_{\mathrm{tot}}) \prod_{i=1}^{n_{\mathrm{obs}}} \frac{1}{\mu_{\mathrm{tot}}} \left(\mu_{\mathrm{sig}} f_{\mathrm{sig}}((x)_i) + \sum_{b \in \mathrm{bkg}} \mu_b f_b(\boldsymbol{x}_i) \right) \right] \prod_{p \in \mathrm{NPs}} \mathrm{Gaus}(\mu_p | \nu_p, \sigma_p), \tag{3.26}$$

where n_{obs} is the total observed number of events, μ_{tot} is the summed expectation of background counts, μ_{sig} is the signal strength tested, μ_b is the expectation of each background component, $f_{\mathrm{sig}}(\boldsymbol{x}_i)$ and f_{b} are the integrated counts of events weighted by the signal and background PDFs, respectively, μ_p represents the strength of the nuisance parameters, and where ν_p and σ_p are the mean and standard deviation for the corresponding nuisance parameter (NP). There are three components of the likelihood function. The Poisson term extends the likelihood to the expectation of total number of events observed. The first product is the core component of the likelihood function, and compares the observed data in two dimensions (S1c, $\log_{10}$S2c) to the expectations of the PDFs. Finally, the Gaussian term constrains the nuisance parameters.

The test statistic t_μ contains the same information as λ, but is defined for convenience as

$$t_\mu = -2 \log \lambda(\mu), \tag{3.27}$$

which ranges from 0 to infinity. For a dark matter search, in which a signal can only increase the number of events, the test statistic is modified to ensure that the minimum value of the estimator

$\hat{\mu}$ is zero. The modified test statistic is

$$\tilde{t}_{\mu} = \begin{cases} -2\log\frac{\mathcal{L}(\mu,\hat{\hat{\boldsymbol{\theta}}})}{\mathcal{L}(\hat{\mu},\hat{\boldsymbol{\theta}})}, & \text{if } \hat{\mu} \geq 0 \\ -2\log\frac{\mathcal{L}(\mu,\hat{\hat{\boldsymbol{\theta}}})}{\mathcal{L}(0,\hat{\boldsymbol{\theta}}(0))}, & \text{if } \hat{\mu} < 0, \end{cases} \tag{3.28}$$

where negative values of $\hat{\mu}$ are still allowed for $\mu > 0$ [116].

The p value of the test statistic distribution quantified the amount of disagreement between the data and the particular value of μ being tested, and is calculated from

$$p_{\mu} = P(\tilde{t}_{\mu} \geq \tilde{t}_{\mu,\text{obs}}|\mu) = \int_{\tilde{t}_{\text{obs}}}^{\infty} f(\tilde{t}_{\mu}|\mu) d\tilde{t}_{\mu}, \tag{3.29}$$

where $\tilde{t}_{\mu,\text{obs}}$ is the test statistic calculated from the observed data, and where $f(\tilde{t}_{\mu}|\mu)$ is the probability density function for $\tilde{t}_{\mu}$, typically generated from toy simulations. A two-sided interval bounding the signal strength μ is constructed by setting the p value equal to a value α set by convention in the dark matter community to be 0.1, producing confidence levels of 90 %. This procedure is carried out for each WIMP mass, for each operator-isospin combination.

Both in the standard WIMP search and the EFT search, there are background fluctuations which cause the confidence interval construction to produce upper limits corresponding to signal strengths too small for the detector to be sensitive to. In these cases, the upper limit is power constrained such that the power of the alternate hypothesis is $\pi_{\text{crit}} = 0.16$, following the convention of the SR1 WS [5]. This corresponds to a restriction for the limits falling below the -1σ band.

3.4.2 Elastic results

The 90% upper limits for the coupling strengths of WIMPs scattering elastically via the operators $O_{1,3-15}$ are shown in Figure 3.9 for the pure isoscalar scenario, and in Figure 3.10 for the pure isovector scenario. We follow Ref. [109] in reporting the dimensionless coupling strength $(c_i^N \times m_W^2)^2$, where $m_W = 246.2$ GeV is the Higgs's vacuum expectation value. The observed limits are power constrained for WIMP masses between 17 GeV and 30 GeV such that the minimum observed limit does not fall below the -1σ band of the expected limit. The majority of the observed limits are compatible with observations to within 1σ, with a handful operator–mass combinations producing limits weaker than expected by less then 2σ, as shown in Figure 3.11. The weaker limits correspond to $\mathcal{O}_{3,13,15}$ at masses above 30 GeV, where the highly peaking spectra overlap with ER leakage.

Operators $\mathcal{O}_1$ and $\mathcal{O}_4$ allow for a consistency check with the SI and SD operators that couple solely to the total charge and total spin, as reported in the first LZ result [5]. These two operators are unaffected by the $\vec{v}^{\perp}$ and $\vec{q}$ degrees of freedom, resulting in unsuppressed event rates that lead to the most stringent constraints of the entire set of EFT operators with the exceptions of $\mathcal{O}_{11}$ and $\mathcal{O}_{12}$. The operator $\mathcal{O}_{11}$ only differs from the standard SI interaction due to a momentum-dependence that reduces its rate at low recoil energies; the nuclear response is identical. $\mathcal{O}_{12}$ is an example of an semi-coherent operator whose rate is enhanced by a positive parity coupling involving the nucleon velocity, leading to contributions from each nucleon when summed over the nucleus [109]. Further, while most WIMP couplings are excluded most strongly at WIMP masses of 30–50 GeV, interactions that are suppressed by several powers of $\vec{q}$, such as $\mathcal{O}_6$ and $\mathcal{O}_{15}$, are constrained most tightly at WIMP masses of 200-300 GeV. The highly suppressed operators benefit the most from the extended energy window.

Comparisons with results from XENON100 [117], LUX [119], and PandaX [118] analyses are also provided for reference in Figures 3.9 and 3.10, having accounted for differing conventions for the coupling strength normalization. Many results are not shown for comparison, owing to their use of the proton–neutron basis instead of the isospin basis in reporting results. Although LZ demonstrates the best sensitivity for all EFT interactions considered, previous analyses have used outdated one-body nuclear density matrices that lead to a stronger limit than LZ for some operators such as $\mathcal{O}_{13}$ [120].

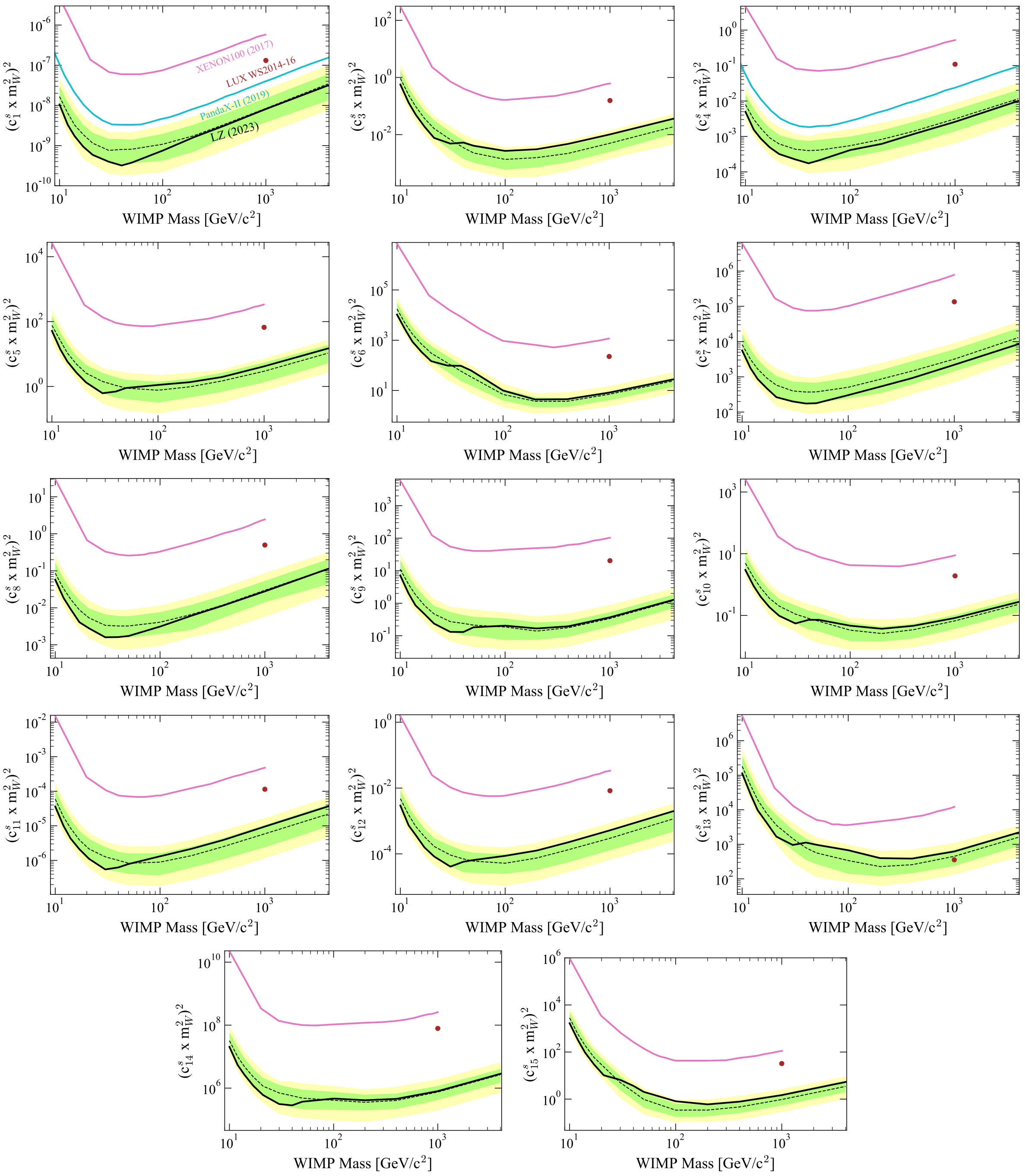


Figure 3.9: Elastic isoscalar cross section upper limits (black), with the 1σ and 2σ bands shown in green and yellow, respectively. The median sensitivity based on the LZ background model is shown in dotted black. Also shown are the XENON100 results (magenta) [117], PandaX-II (blue) [118], and a point from the LUX inelastic analysis (brown) [119].

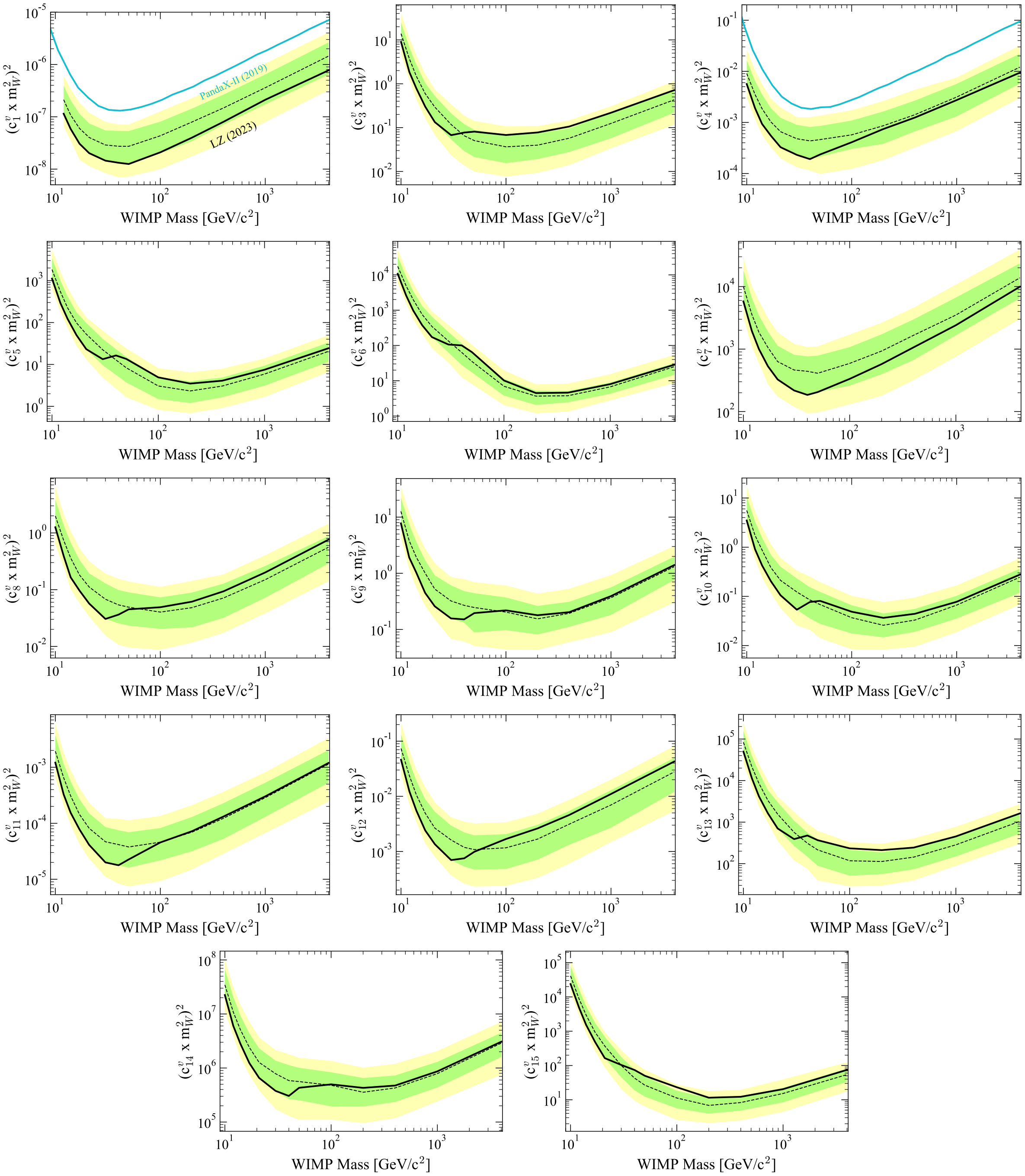


Figure 3.10: Elastic isovector cross section upper limits (black), with the 1σ and 2σ bands shown in green and yellow, respectively. The median sensitivity based on the LZ background model is shown in dotted black. The PandaX-II result is shown in blue [118].

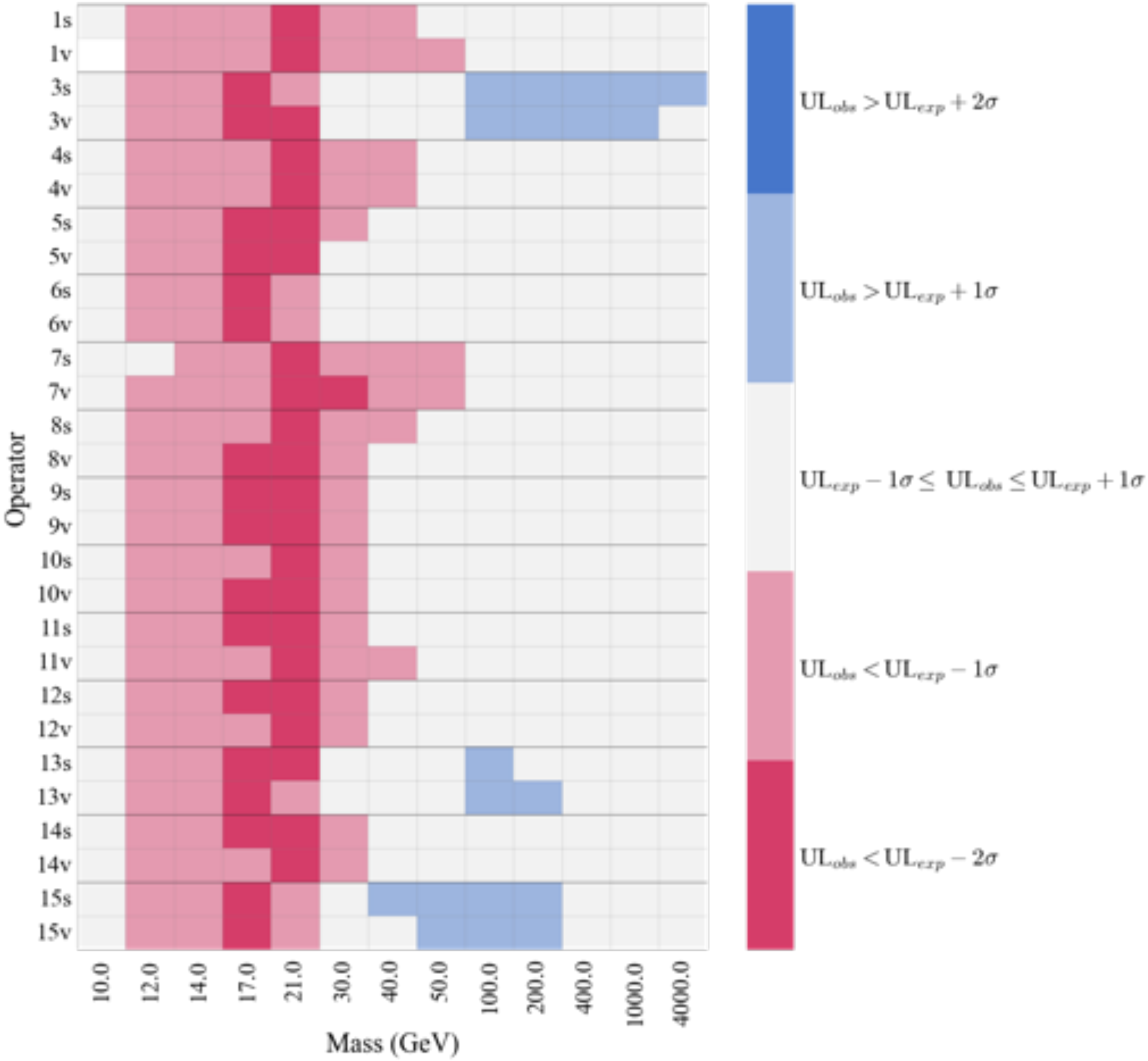


Figure 3.11: Statistical significance of observed upper limit as compared to the expectation for each operator–isospin combination at the tested masses. Red shades indicate a stronger limit than expected, with a power constraint applied to the darker shades, while the blue shades indicate a weaker limit than expected.

3.4.3 Inelastic results

The elastic results were generalized to consider a WIMP that scatters to a heavier excited state with a mass splitting δ up to 250 keV [121]. This kinematic effect is incorporated into the Hermitian basis vectors by accounting for δ in the energy conservation condition. The perpendicular velocity in Eq. 3.4 is modified such that

$$\vec{v}^{\perp}_{\text{inelastic}} = \vec{v} + \frac{\vec{q}}{2\mu_N} + \frac{\delta}{|\vec{q}|^2}\vec{q} = \vec{v}^{\perp} + \frac{\delta}{|\vec{q}|^2}\vec{q}. \tag{3.30}$$

The differential event rate in Eq. 3.24 is then reduced since the minimum WIMP velocity required to deposit a certain energy is higher, due to some energy going towards the excitation of the WIMP.

90% upper limits for the coupling strengths of WIMPs scattering inelastically via the operators $O_{1,3-15}$ are shown in Figure 3.12 for the purely isoscalar scenario, and in Figure 3.13 for the purely isovector scenario. In these figures, mass splittings of the excited WIMP up to $\delta = 250$ keV are considered. As in the case of the elastically scattering WIMP, no significant excesses are observed, and observed limits are always within 2σ of the expectations as shown in Figures 3.14, 3.15, and 3.16 for WIMPs of mass 400 GeV, 1,000 GeV, and 4,000 GeV, respectively.

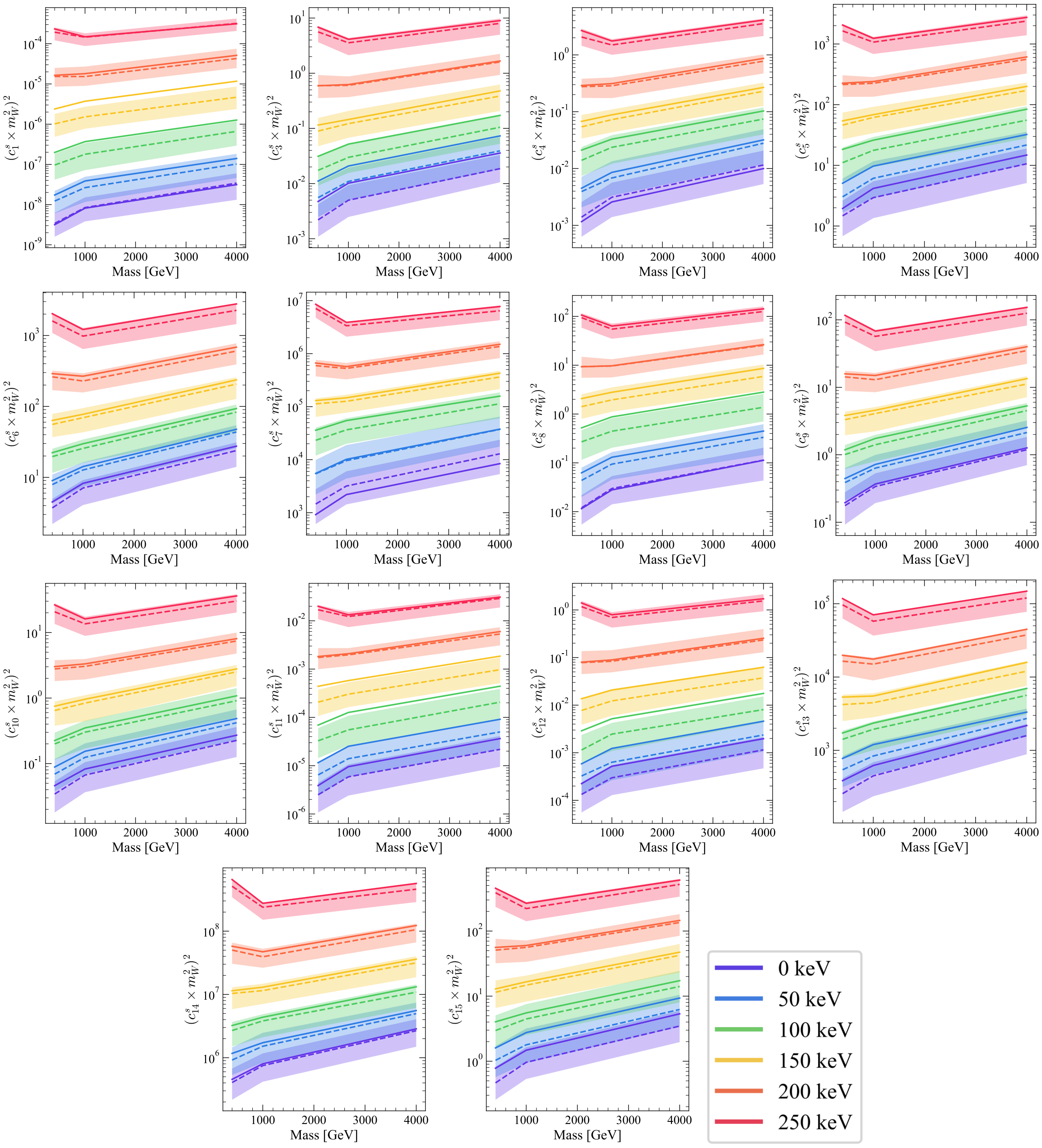


Figure 3.12: Inelastic isoscalar upper limits for the fourteen WIMP-nucleon interactions (solid lines) with the median expectation (dashed line) and 1σ band (shaded). Upper limits are evaluated for WIMP masses of 400 GeV, 1,000 GeV, and 4,000 GeV for δ = 0 keV (purple), 50 keV (blue), 100 keV (green), 150 keV (yellow), 200 keV (orange), and 250 keV (red).

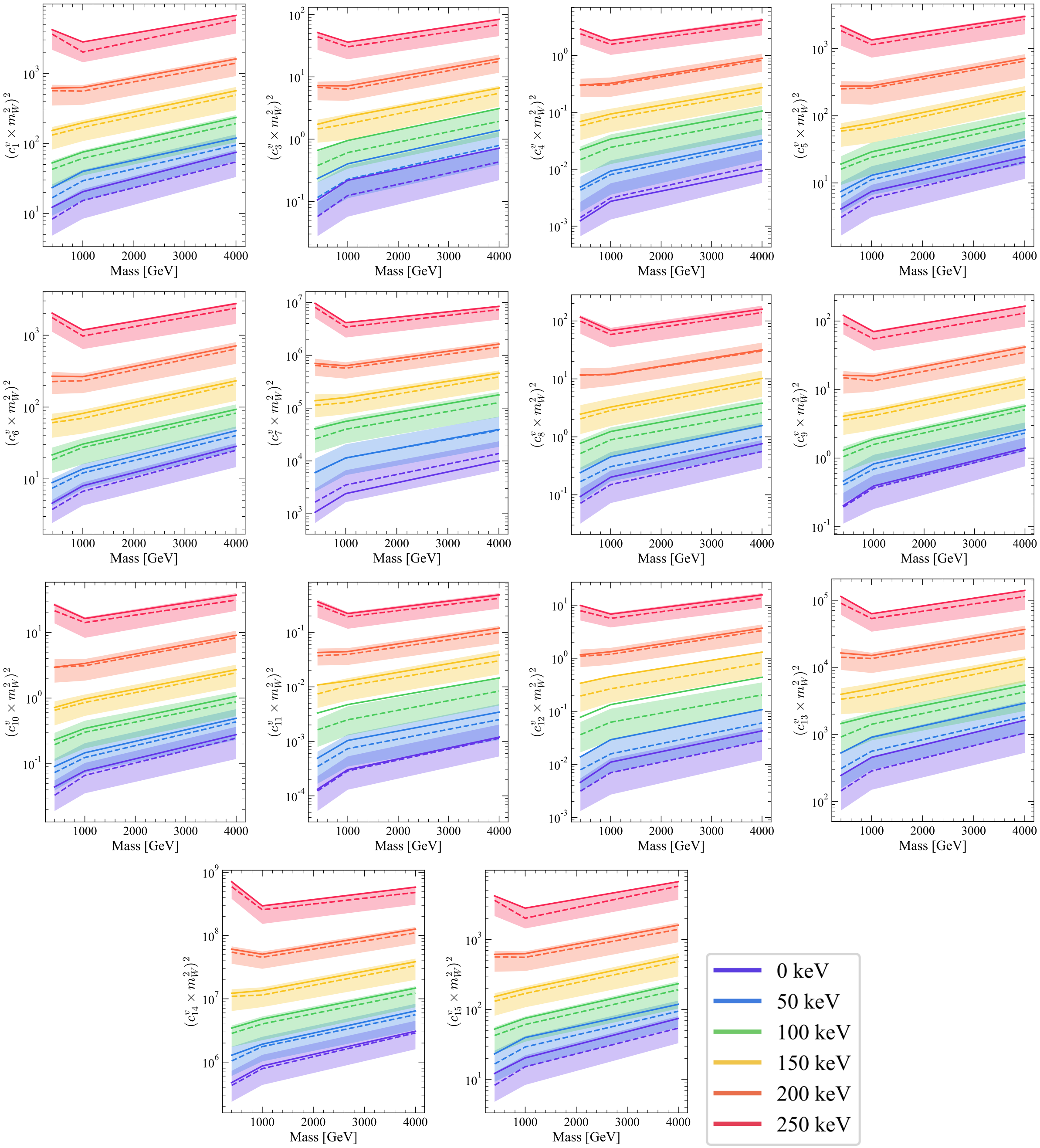


Figure 3.13: Inelastic isovector upper limits for the fourteen WIMP-nucleon interactions (solid lines) with the median expectation (dashed line) and 1σ band (shaded). Upper limits are evaluated for WIMP masses of 400 GeV, 1,000 GeV, and 4,000 GeV for $\delta = 0$ keV (purple), 50 keV (blue), 100 keV (green), 150 keV (yellow), 200 keV (orange), and 250 keV (red).

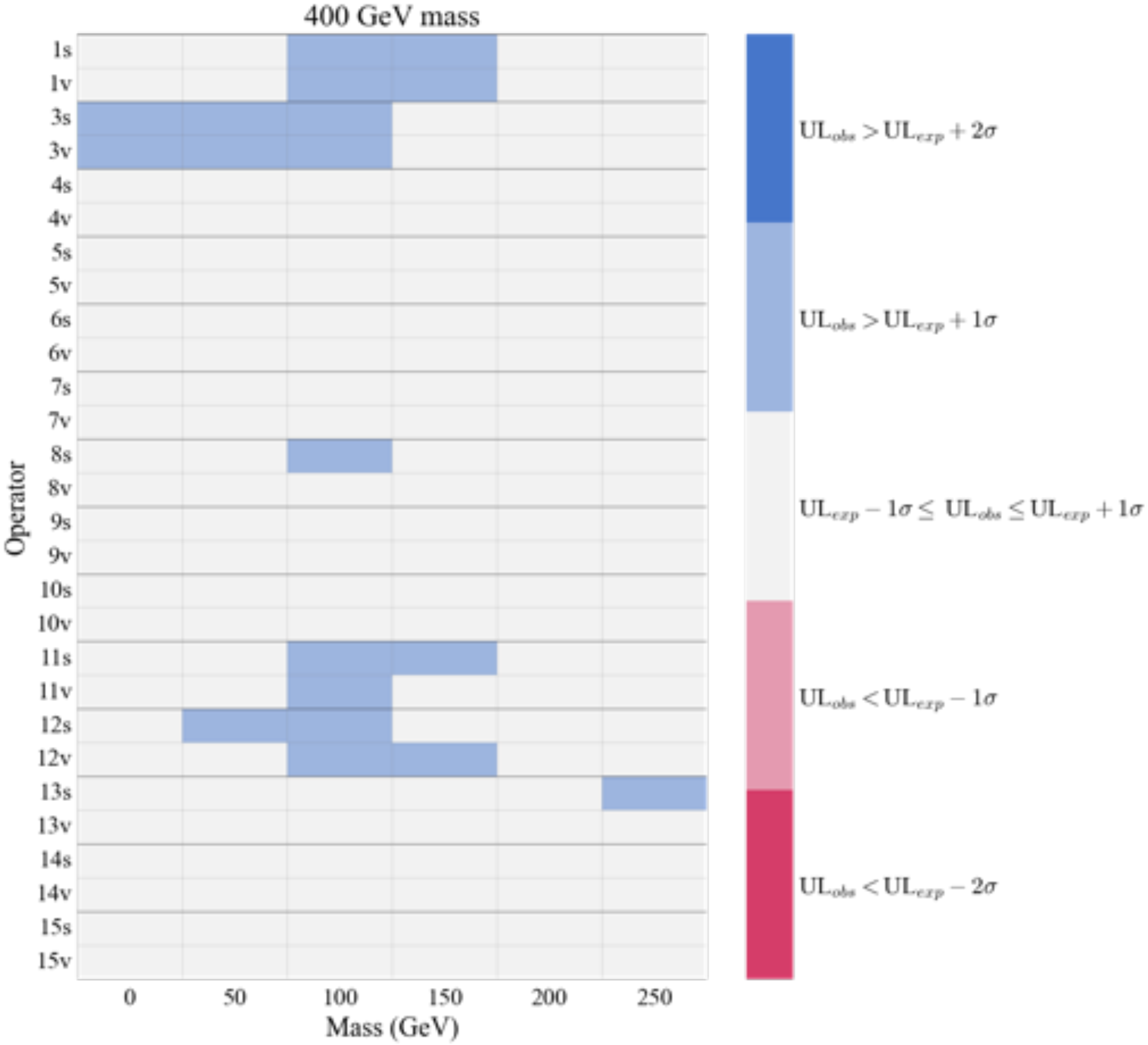


Figure 3.14: Statistical significance of observed upper limit as compared to the expectation for a 400 GeV WIMP.

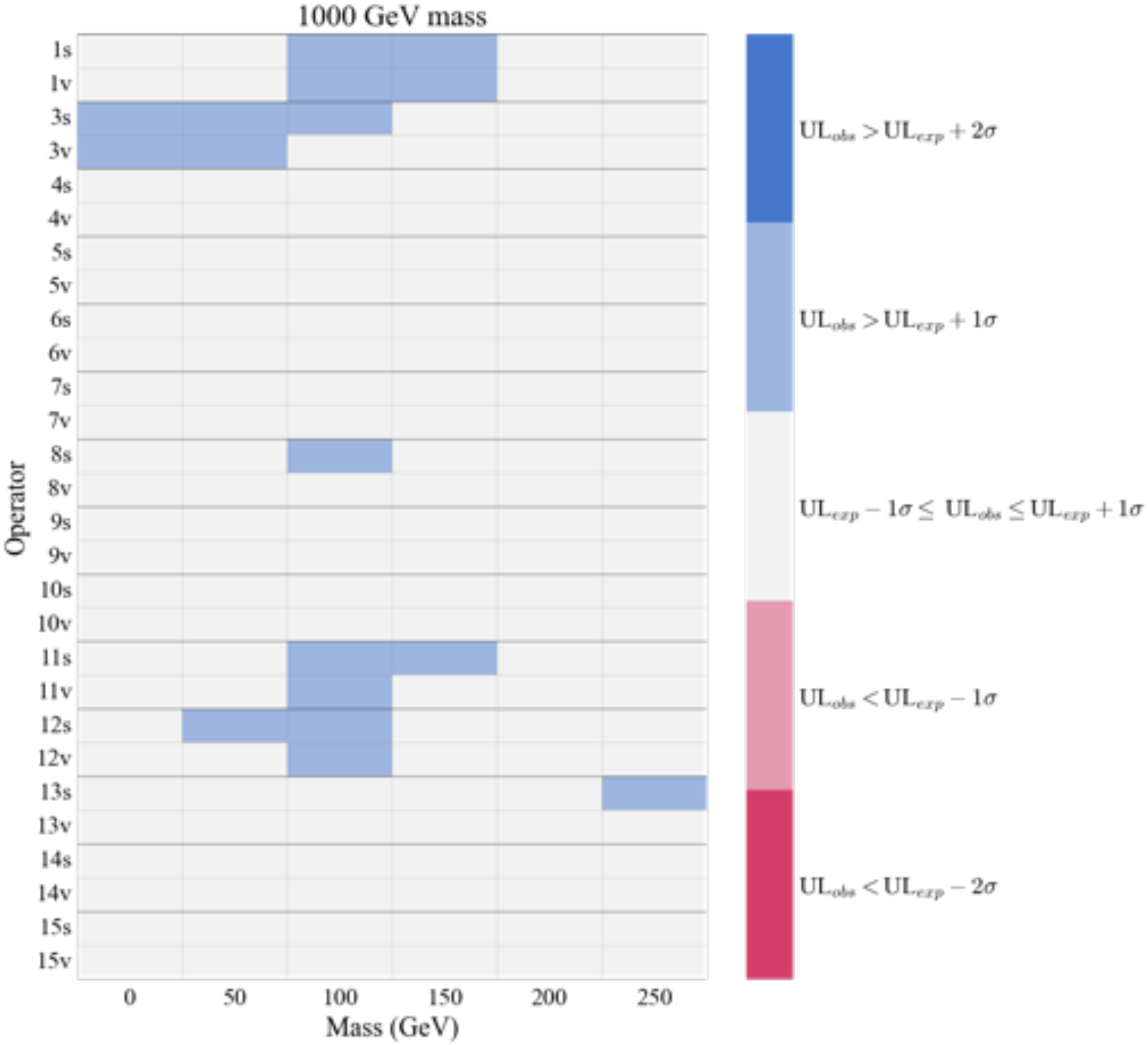


Figure 3.15: Statistical significance of observed upper limit as compared to the expectation for a 1,000 GeV WIMP.

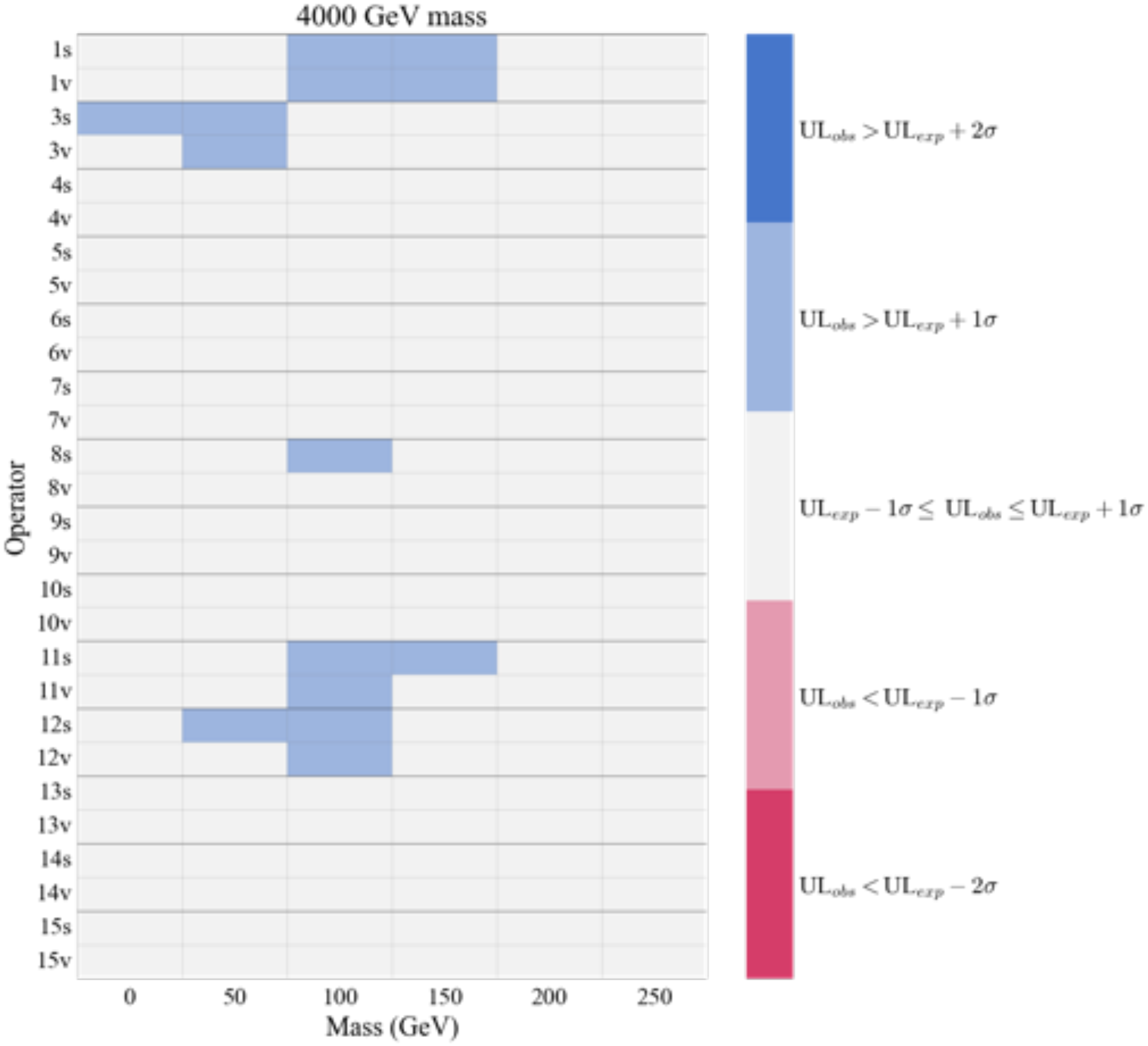


Figure 3.16: Statistical significance of observed upper limit as compared to the expectation for a 4,000 GeV WIMP.

Chapter 4

γ-X Background and Classification

Nothing is more frightening than a fear you cannot name.

Cornelia Funke, Inkheart

The primary challenge in extending the energy region of interest to conduct an EFT search is a new source of background that only manifests at higher energies, called the γ-X background. LZ's interaction finder algorithm classifies γ-X events as single scatters, even though this background is produced by γ rays scattering multiple times near the boundaries of the TPC. This is because despite the multiple scatter, only one S2-producing interaction site is observed.

In γ-X events only one scatter occurs in the fiducial volume and the rest may occur either i) below the cathode in the reverse field region (RFR), or ii) near the TPC wall. The electric field in the RFR points in the opposite direction to the drift field, and the electric field near the TPC wall deviates from uniformity due to the proximity of field shaping rings. An energy deposit in either of these regions may not produce an ionization signal due to the lack of an uniform electric field that drifts electrons towards the gate, but will produce scintillation light. Therefore, the misclassification of γ-X events as single scatters occurs because they have one apparent S1 pulse created by the summation of scintillation pulses from multiple energy deposits (which cannot be individually resolved), and one S2 pulse from the interaction in the fiducial volume. The γ-X background directly overlaps the region occupied by high energy nuclear recoil events in S1–S2 space, necessitating a comprehensive model and a cut to remove them.

The signature of this background is complex. Several characteristics of the data have been identified to detect γ-X events, including geometric information of the hit pattern that the scintillation light creates on the PMT arrays, and distributions of S1 and S2 pulse areas. These discriminant characteristics have been summarized in 7 variables, detailed in subsection 4.1.1. Simple cuts on these variables do not discriminate γ-X events cleanly, so a boosted decision tree (BDT) model was trained on simulations to identify this background in real data. The advantage of the BDT is that it processes the feature space holistically and provides a γ-X cut that incorporates information from the 7-dimensional data, a task that is difficult for a human to perform manually.

The performance of the BDT was evaluated on i) simulations, ii) data from ER and NR calibration runs, and iii) a high-energy sideband from the SR1 background run. The EFT analysis requires the BDT to have a near perfect acceptance of NR single scatters, so that bona-fide EFT WIMP events are not cut out. The next requirement in order of importance is a high rejection of γ-X events. A complete removal of γ-X events, while desirable, is not required, due to the existence of a γ-X model that can account for unremoved events in the statistical analysis. The result of the BDT evaluation, detailed in section 4.3, demonstrates that the BDT meets these requirements when run on simulated and experimental data.

4.1 Description of γ-X background events

γ-X events can originate from several sources. The three types of γ-X events with the highest rate were chosen to be studied in detail, and are shown in Figure 4.1.

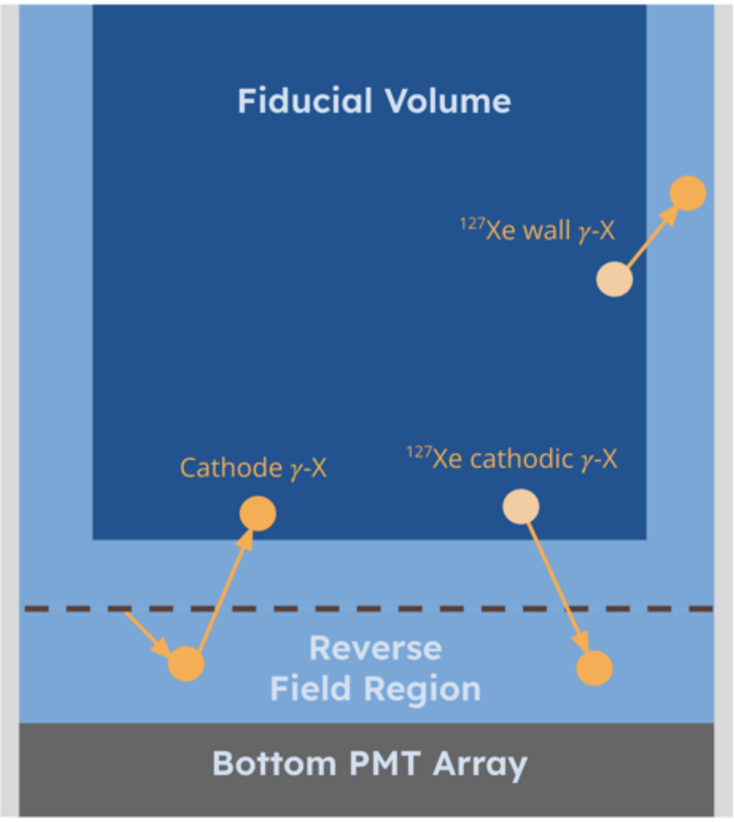


Figure 4.1: Schematic of the TPC showing three different types of γ-X events, produced by: a γ ray emitted from the cathode grid, a ^{127}Xe decay near the cathode, and near a wall.

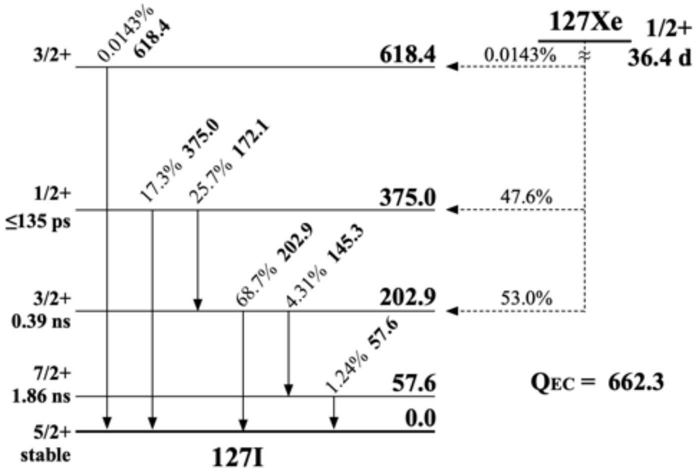


Figure 4.2: Decay scheme of ^{127}Xe in units of keV. THIS IS DQ's FIGURE. REPLACE WITH MY OWN.

1. The first type is called cathode γ-X, referring to the origin of the γ rays from rare-earth decays in cathode grid wire impurities. A γ ray emitted from the cathode can take one of two paths to create a γ-X event; either scattering first in the RFR and then in the fiducial volume as shown in Figure 4.1, or vice versa.

2. ^{127}Xe is an isotope of xenon created from activation that decays via electron capture to ^{127}I, releasing a γ ray and an x-ray, each with characteristic energies. A decay diagram is shown in Fig. 2. When a ^{127}Xe electron capture decay occurs, the x-ray deposits its energy in the immediate vicinity of the decay, while the γ ray traverses up to 10 cm before Compton scattering or being totally absorbed. A cathodic ^{127}Xe γ-X event occurs if the decay occurs in the fiducial volume and the γ traverses into the RFR. The γ deposit will contribute only S1 light, while the x-ray deposit contributes S1 and S2 light.

3. If the electron capture occurs near the wall of the TPC, the γ ray may traverse into a region of the wall where, due to electric field non-uniformities, the resulting energy deposit only produces a partial ionization signal. In this case, a wall ^{127}Xe γX event occurs.

The rate of γ-X events in SR1 has been estimated from simulations, even though the rarity of these events makes direct simulation difficult. For every 1,000 ^{127}Xe electron capture decays in the TPC, only one γ-X event is created. Since direct simulation is difficult, a models of this background have been created using a custom simulation framework that takes into account the full LZ geometry, 3D Compton scattering, and the initial energy distributions of γs emitted from impurities in the cathode, and characteristic energies of ^{127}Xe γs and x-rays. This generator provides distributions of γ-X events in S1–S2 space that cannot be identified completely with either the NR or ER bands. In addition, the energy deposits from this simulation are fed into a program that converts them to corrected signal variables, in addition to calculating quantities relevant to the discrimination. This custom γ-X generator has been validated using LZ data with ^{127}Xe activation peaks. Integrating the generated model in S1–S2 over the analysis region of interest, normalized by the expected rate of ^{127}Xe electron capture decays in the FV, provides an estimate of the expected number of γ-X events for the EFT analysis. Up to S1 $= 600$ phd, we expect 1.6 γ-X events with the ^{127}Xe cathodic component dominating the expected rate.

4.1.1 Characteristic features of γ-X events

The 7-dimensional data provided as inputs to the classifier can be categorized into PMT hit pattern features, position features, and signal features, and are summarized here.

1. **Cluster Size**

   Cluster size refers to the size of the 'splash' the S1 light makes on the bottom PMT array. It is quantified as the distance of each PMT from the centroid of the splash, weighted by the PMT channel area A_i^{ch}, such that

$$\text{Cluster Size} = \frac{\sum_i^{\mathrm{bPMTs}} A_i^{\mathrm{ch}}(r_i - r_{\mathrm{centroid}})}{\sum_i^{\mathrm{bPMTs}} A_i^{\mathrm{ch}}}. \tag{4.1}$$

   In a cathodic γ-X event, the scattering vertices in the reverse field region produce scintillation light close to the bottom PMT array, shrinking the cluster size, while leaving the drift

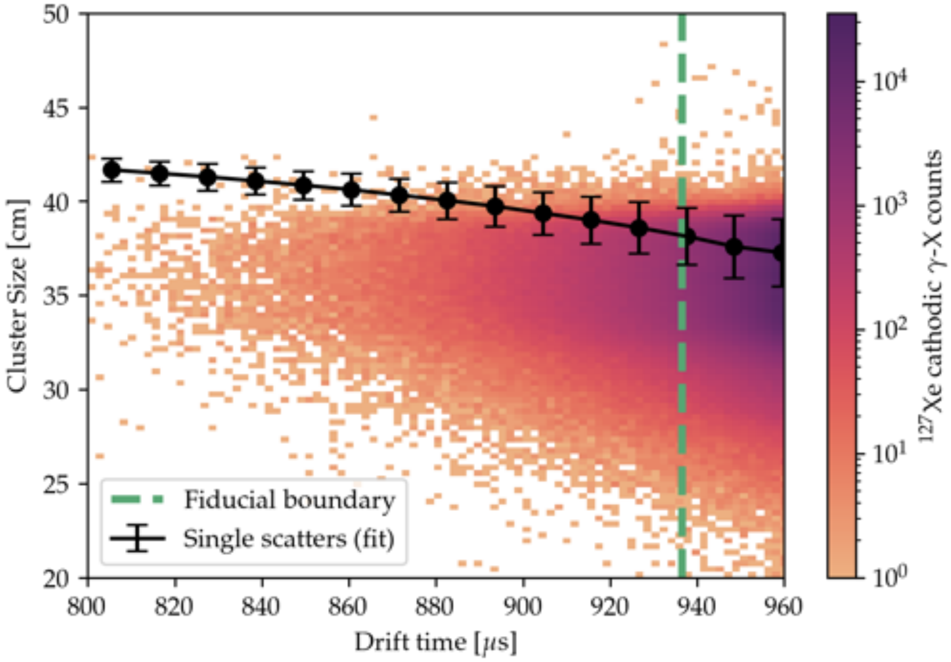


Figure 4.3: Cluster size versus drift time distribution of ^{127}Xe cathodic γ-X events, shown with the trend of single scatter events (solid black). The fiducial boundary for the SR1 EFT search is shown by the green dashed line. All data shown are simulated.

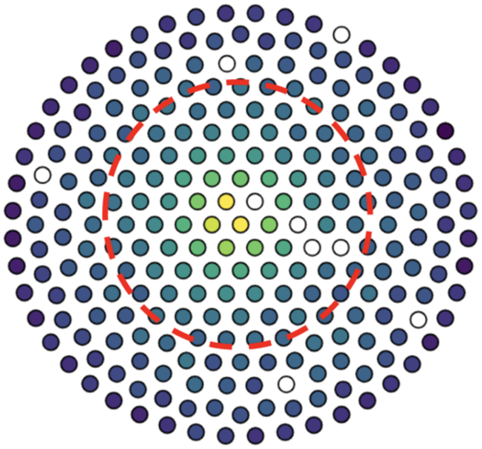


Figure 4.4: The bottom PMT hit pattern for an example event. Each solid circle shows the location of a PMT, and the color indicates the amount of light it collected: no light (purple) and largest amount of light in the event (yellow). Several PMTs were turned off for SR1 and are shown as circles with no color. The calculated cluster size (40.6 cm) is shown in the red circle.

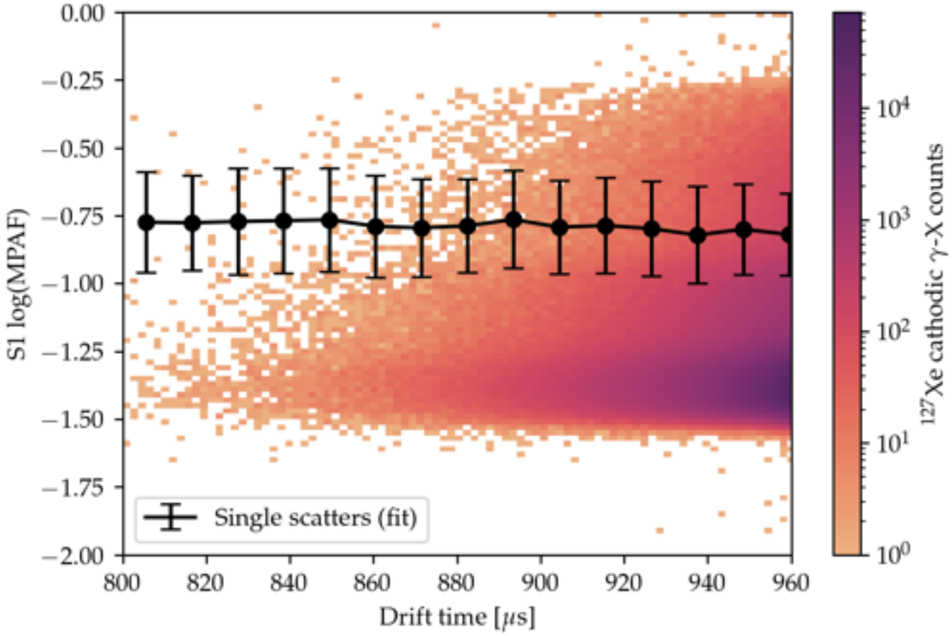


Figure 4.5: Cluster size versus drift time distribution of ^{127}Xe cathodic γ-X events, shown with the trend of single scatter events (solid black). The fiducial boundary for the SR1 EFT search is shown by the green dashed line. All data shown are simulated.

time unchanged. The cluster size distribution of simulated ^{127}Xe cathodic γ-X background is shown with drift time in Fig. 4.3, along with simulated single scatters. The single scatter data are sliced in drift time and the set of cluster size histograms are fitted to Gaussian functions for a clear comparison with the γ-X events. Fig. 4.4 shows the hit pattern for an example event in the bottom PMT array, with the cluster size radius indicated in red.

2. **Max Log Peak Fraction (MPAF)**

   The MPAF is defined as

$$\mathrm{MPAF} = \frac{\max\{A_i^{\mathrm{ch}}\}}{\mathrm{sum}\{A_i^{\mathrm{ch}}\}}, \tag{4.2}$$

   and is the most amount of light a single bottom PMT collected from the S1, expressed as a fraction of the total light the bottom PMT array recorded. The distribution is shown in Fig. 4.5.

3. **Radius**

The model uses a cut on the radius to perform a rough selection of candidate events for wall γ-X.

4. **Drift time**

   Candidate events for cathodic γ-X are selected using a cut on drift time. Drift time can provide additional classification power when cut sequentially with the S1 TBA.

5. **Top Bottom Asymmetry (S1 TBA)**

   This is the magnitude of the asymmetry of total S1 light seen by the top and bottom PMT arrays, given by

   $$\mathrm{TBA} = \frac{\sum_i^{\mathrm{tPMTs}} A_i^{\mathrm{ch}} - \sum_i^{\mathrm{bPMTs}} A_i^{\mathrm{ch}}}{\sum_i^{\mathrm{tPMTs}} A_i^{\mathrm{ch}} + \sum_i^{\mathrm{bPMTs}} A_i^{\mathrm{ch}}}. \tag{4.3}$$

   While the S1 TBA is an excellent proxy variable for the drift time of a single scatter event, it is not so for a cathodic γ-X event. This is because the drift time of a cathodic γ-X event corresponds solely to the interaction in the fiducial volume that produced the ionization signal, while the TBA encodes information about the interaction closer to the bottom PMT array as well. Therefore cathodic γ-X events are expected to have more negative S1 TBA than single scatters at the same apparent depth. This is shown in Figure 4.6 using simulated single scatter and ^{127}Xe cathodic γ-X events. The single scatter data are sliced in drift time and Gaussian functions are fitted to the set of resulting TBA distributions for a clear comparison with the cathodic γ-X distribution.

6. **S1**

   γ-X events, especially those produced by ^{127}Xe decays, occupy distinctive regions in S1–S2 space. The size of the S1 pulse can be used with the size of the S2 pulse to roughly isolate γ-X events.

7. **S2**

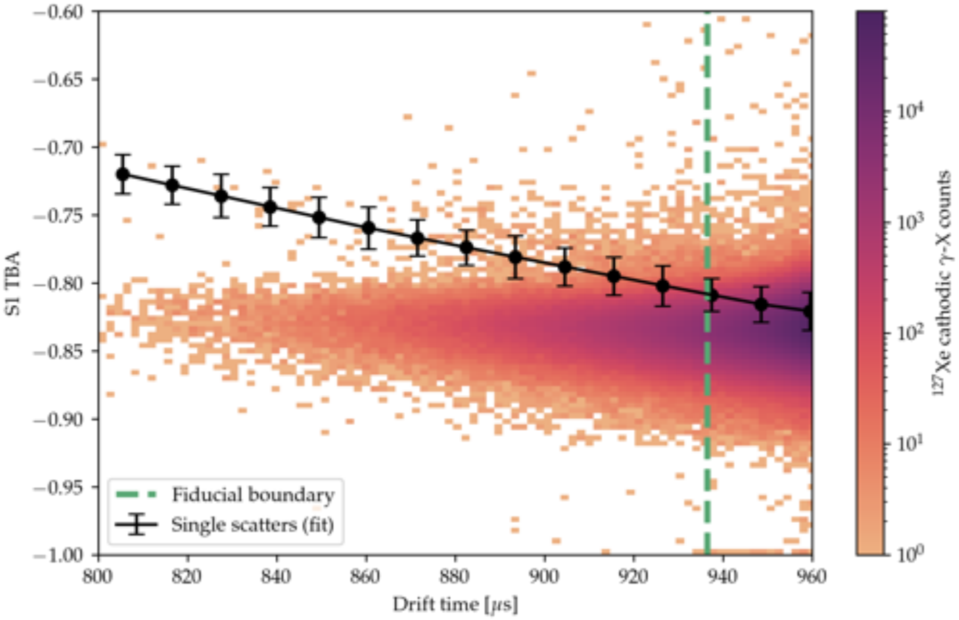


Figure 4.6: TBA versus drift time distribution of ^{127}Xe cathodic γ-X events, shown with the trend of single scatter events (solid black). The fiducial boundary for the SR1 EFT search is shown by the green dashed line. All data shown are simulated.

γ-X events produced by ^{127}Xe decays occupy bands in the space of all S2s. The locations of these bands are set by the characteristic X-ray energies, at 32.2 keV, 5.2 keV, 1.1 keV and 186 eV. The S1-S2 distributions of the three sources of γ-X events are shown in Figure 4.7.

4.2 Description of the γ-X classifier

Due to the complex nature of the γ-X background, the classification task was performed using a BDT. BDTs belong to the area of machine called supervised learning, where the learning task is carried out by training the model on data for which truth information already exists, called the training set. In this case, the training set consists of events with a 7-dimensional data vector (components described earlier) that belong to the single scatter and γ-X classes. Once trained, the BDT is able to make predictions for data without truth information. Since algorithms like BDTs are often construed as black boxes, in this section we review the theory behind BDTs and show that the training process is almost identical to a task familiar to physicists: performing a fit.

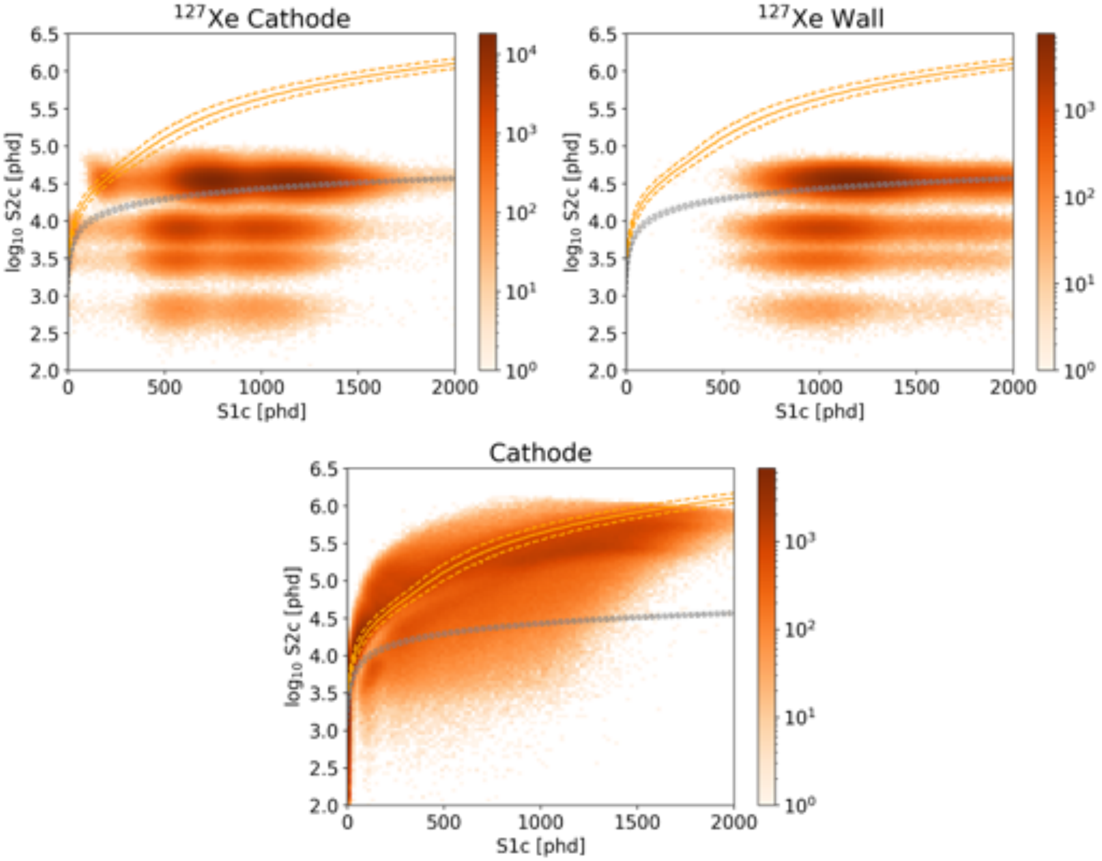


Figure 4.7: S1-S2 distributions of the cathodic ^{127}Xe model (top left), the wall ^{127}Xe model (top right), and cathode γ-X model (bottom). The detector β band and NR band are shown in orange and gray, respectively.

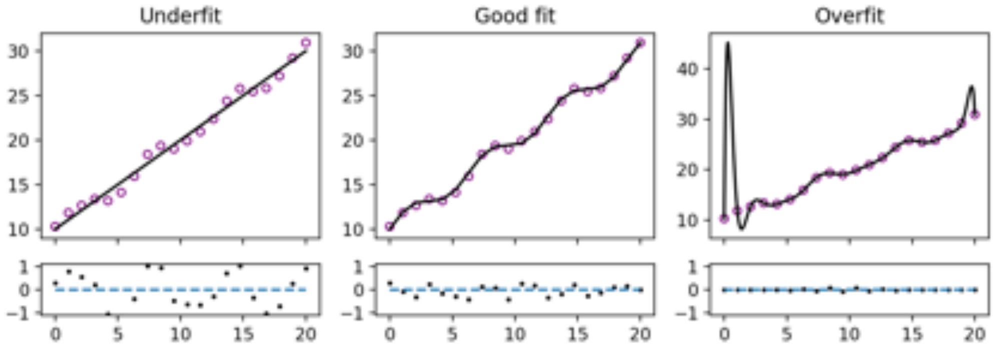


Figure 4.8: An example of an underfit with a linear function (left), a good fit (middle), and an overfit with a high order polynomial (right). Residuals (data - model) are shown in the bottom row of plots. The purple data are from the function $y = 10 + x + \sin(x)$ with some random noise added. The data in all three plots are the same.

4.2.1 Gradient boosting

The strategy used to train a BDT is known as gradient boosting. The principle of gradient boosting is to add up an ensemble of simple models f_g, or weak learners, to create a large model F_g that performs a complicated task. Fig. 4.8 shows how one might manually apply this strategy with the task of approximating an unknown function to fit some data. The weak learners in this case are simple functions, but for BDTs, they are decision trees. In the simplified example, the data is quite well described by the summation of 3 simple functions, but in practice, the number of weak learners is much larger. The number of weak learners M is an instance of a hyperparameter, a parameter of the model that can be adjusted to increase performance. For instance, with $M = 2$ our approximation is only a linear function $10 + x$ (left plot of Fig. 4.8), and we will not be able to predict the sinusoidal trend that seems apparent in the data. On the other hand M can be increased to describe the training data arbitrarily well (right plot of Fig. 4.8) at the expense of poorly describing new data sampled from the same underlying distribution.

Gradient boosting only works if each simple model can be summed to produce the target variable. In other words, the predictions of the ensemble is required to be additive. To demonstrate this consider the task of approximating the scalar output of a function $y = f(\mathbf{X})$ that takes an m-dimensional input vector $\mathbf{X}$. Suppose that n of these vectors are each associated with a truth variable (the target) to make up the training set $\bar{\mathbf{X}}_\mathrm{t}$. That is, each element of the training set will be a vector $\mathbf{X}$ and a truth variable y. A model $F_G := (\mathbb{R}^n \times \mathbb{R}^m) \to \mathbb{R}^n$ that approximates f can be built using a sequence of weak learners $f_g : (\mathbb{R}^n \times \mathbb{R}^m) \to \mathbb{R}^n$ using the recurrence relation

$$F_g(\bar{\mathbf{X}}_\mathrm{t}) = F_{g-1}(\bar{\mathbf{X}}_\mathrm{t}) + f_g(\bar{\mathbf{X}}_\mathrm{t}), \tag{4.4}$$

where the first function f_0 is a starting guess for the prediction, typically the average of target values in the training set $\mathbf{X}_\mathrm{t}$, such that

$$f_0(\mathbf{X}_\mathrm{t}) = \frac{1}{n} \sum_{y \in \mathbf{X}_\mathrm{t}} y. \tag{4.5}$$

In practice, the influence of new learners is reduced by scaling their predictions down by the learning rate, η, before adding them to the predictions of the model

$$F_g(\mathbf{X}_\mathrm{t}) = F_{g-1}(\mathbf{X}_\mathrm{t}) + \eta f_g(\mathbf{X}_\mathrm{t}). \tag{4.6}$$

The primary function of the learning rate is to reduce overfitting in the model, and allow future weak learners to also contribute their results to the prediction.

Given that the first element f_0 is a guess for the scalar to be predicted, the desired effect of adding subsequent elements, or boosts, is to add the residual between y and the first approximation $\hat{y}$ in order to nudge the prediction towards y, the true target value. Therefore each weak learner f_g is trained to reproduce the the residual $(y - \hat{y})$ before it is added to the model.

The process of training each weak learner is specific to the implementation of the gradient boosting algorithm. In this case, the weak learners are small decision trees that split the training data on a single feature at each node. Choices have to be made in how deep to grow the tree, what feature to split on, and how to calculate the numerical value to split on. These choices are other hyperparameters of the BDT, and can be adjusted for better performance.

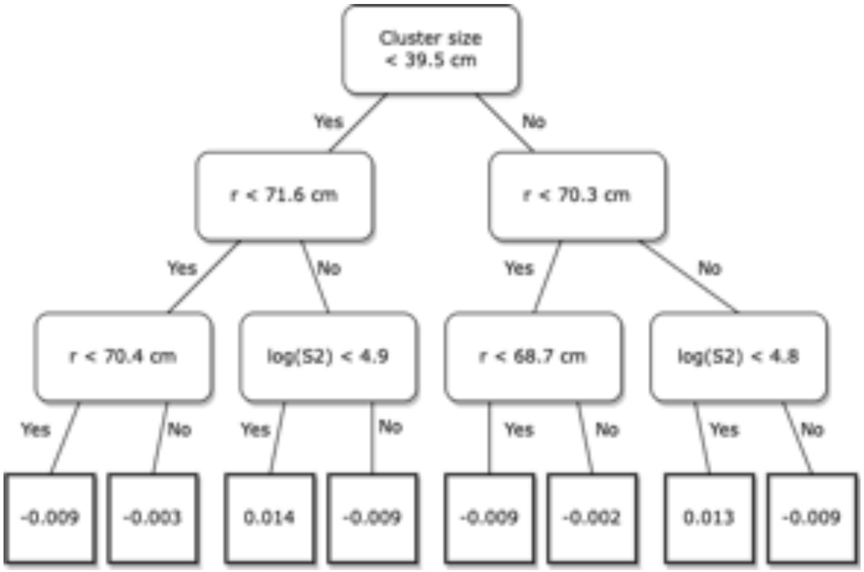


Figure 4.9: A single decision tree from an early version of the γ-X classifier with a maximum depth of 3. Each leaf (blocks at the bottom) contains a residual, calculated by averaging the residuals from the older trees. The numerical values of the splits are chosen to minimize the variance among the data in the leaves.

4.2.2 Decision trees

Fig. 4.9 shows an example of a tree in an early γ-X classifier. All of the trees in this model were grown to a maximum depth of 3, whereas the final model grows trees to a max depth of 6. Inspecting the decision paths in the tree, it is no surprise that it is a weak learner because it cannot partition the data finely enough over the full 7-dimensional distribution necessary to make accurate decisions. However, it is sufficient (and also necessary) that the individual tree isolates γ-X events only slightly better than random. This is achieved for each tree f_g by choosing numerical values for the split at each node such that the variance of the residuals (calculated from the preceding sequence of trees F_{g-1}) in each leaf of f_g is minimized. In each leaf, new residuals are calculated by taking the average of the residual vector and applied to the predictions of the data in the leaf, nudging the predictions towards the actual values.

4.2.3 Gradient boosting as gradient descent

It can be shown that gradient boosting is nearly identical to gradient descent, an iterative optimization algorithm used widely in fitting routines, for example to find the minimum in a loss function L. The two techniques are subtly different because in boosting, the minimization is carried out in the space of predictions, rather than in the space of parameters, as is the case for gradient descent. To demonstrate this, suppose that from an initial starting location $\mathbf{X}$, the algorithm calculates the gradient of the loss function $\nabla L(\mathbf{X})$ and updates the initial point in the resulting direction such that

$$\mathbf{X}_i = \mathbf{X}_{i-1} - \nabla L(\mathbf{X}_{i-1}), \tag{4.7}$$

where the negative sign facilitates the descent into the minima of L rather than the maximum.

The loss function in gradient boosting is defined as a measure of discrepancy between the predictions $\hat{\mathbf{y}}$ and the true values $\mathbf{y}$. A common choice of loss function is the mean squared error (MSE), defined for a model F_G as

$$L(\mathbf{y}, F_G(\bar{\mathbf{X}})) = \frac{1}{n}\sum_{i=1}^{n}(y_i - F_G(\mathbf{x}_i))^2. \tag{4.8}$$

Writing $\hat{\mathbf{y}}^g = F_g(\bar{\mathbf{X}})$ as the model predictions (using the superscript as an index, not an exponent), the loss function L can be minimized with respect to a single prediction $\hat{y}_j^g$ such that

$$\begin{aligned}
\frac{\partial L}{\partial \hat{y}_j^g} &= \frac{\partial}{\partial \hat{y}_j^g}\left[\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i^g))^2\right] && (4.9)\\
&= \frac{1}{n}\frac{\partial\left[(y_j - \hat{y}_j^g)^2\right]}{\partial \hat{y}_j^g} && (4.10)\\
&= \frac{-2(y_j - \hat{y}_j^g)}{n}, && (4.11)
\end{aligned}$$

which is the residual of the jth event, modulo a factor of $-2/n$. Since the residual is the quantity that each weak learner predicts in its leaves, the recurrence relation describing the model update can be rewritten as for the entire vector of predictions as:

$$
\begin{aligned}
\hat{\mathbf{y}}^{g} &= \hat{\mathbf{y}}^{g-1} + f_{g-1}(\bar{\mathbf{X}}) && (4.12)\\
&= \hat{\mathbf{y}}^{g-1} - (-f_{g-1}(\bar{\mathbf{X}})) && (4.13)\\
&= \hat{\mathbf{y}}^{g-1} - (-\mathbf{r}^{g-1}) && (4.14)\\
&= \hat{\mathbf{y}}^{g-1} - \nabla L(\mathbf{y}, \hat{\mathbf{y}}^{g-1}), && (4.15)
\end{aligned}
$$

where $f_{g-1}(\bar{\mathbf{X}})$ is identified as the residual vector $\mathbf{r}^{g-1}$ obtained from minimizing L. Comparing this result to Eq. 4.7, gradient boosting can be interpreted as gradient descent acting on the space of predictions. Even though the MSE error was used to demonstrate this fact, the result of Eq. 4.12 is generalizable; the use of any differentiable loss function in boosting will perform gradient descent on the space of predictions.

Despite the similarities between gradient descent and gradient boosting, there are core differences in how the models using each of these strategies are trained. A neural network is an example of a model that uses gradient descent to update its parameters: the weights and biases of each of its nodes. The nodes of a network are already established before the training starts, and with each training step, the weights and biases are updated using a loss function that is minimized with respect to the very same parameters. The minimization directly tweaks the parameters of the network, and there is no notion of boosting. In contrast to starting with an established architecture, the BDT uses gradient boosting as a method to build the model. New trees are added sequentially that have the effect of minimizing the loss function over the space of predictions. That is, using the partially correct results of the previous trees, new trees are built to improve the overall predictions.

4.3 Implementation of the γ-X classifier

The model chosen for this task was the XGBoost implementation of the BDT algorithm [122]. The difference between the XGBoost model and other implementations of the algorithm is that XGBoost is developed with optimizations to increase the efficiency of the computation, resulting in the implementation being widely used for data science.

The γ-X BDT was designed to perform multi-class classification of the three types of γ-X events and the two types of single scatters (ER and NR) modeled in simulations. The MSE loss function discussed in the previous section (Eq. 4.8) can be modified to accommodate a binary classification task as opposed to a regression task (estimating the value of a multivariate function). The implementation of the loss function in the trees can be further modified to accommodate a multi-class classification task, like the one desired in this chapter.

4.3.1 Loss function

For a simple binary classification task, the target variable to be predicted $\hat{y}$ is interpreted as a probability between 0 and 1. In this case the loss function has to penalize incorrect predictions from both classes equally, for which a log-loss is more optimal than MSE from a standpoint of not having to compute the squares of floating point numbers between 0 and 1. A log-loss function for binary classification is defined as

$$\text{Log-loss}_B = \frac{-1}{n}\sum_{i}^{n}\left[y_i\log(\hat{y}_i) + (1-y_i)\log(1-\hat{y}_i)\right], \tag{4.16}$$

which evaluates to zero for a perfect classifier.

To fully accommodate a multi-class classification task, we need to i) allow for the possibility of multiple classes in the loss function, and ii) implement this loss function in the trees such that the most likely class can be chosen. One way to meet both of these requirements is to build a binary BDT for each unique pair of classes that we are interested in classifying. A second way to meet these requirements, the one implemented in XGBoost, is to devise a loss function that can handle multiple classes and use it for each class in each tree, essentially vectorizing the loss function. The first approach was tried and the resulting model did not perform as well as XGBoost, although it may show promise by reducing the computational cost if the classification task does not require every pair of classes to be tested. The loss function used in XGBoost's multi-class classification is based on the softmax function

$$\text{softmax}(\hat{y}_c) = \frac{e^{\hat{y}_c}}{\sum_k^K e^{\hat{y}_k}}, \tag{4.17}$$

where the index c denotes the class. The softmax function transforms class predictions $\hat{y}_c$ into probabilities of K possible outcomes. The actual loss function for class c is

$$\text{Log-loss}_c = -\log\left[\text{softmax}(\hat{y}_c)\right] = \log\left[\sum_j^K e^{\hat{y}_j}\right] - \hat{y}_c, \tag{4.18}$$

where in each tree of the model, this loss function is evaluated once for each class.

4.3.2 Data preparation

The single scatter events were simulated using LZ's fast chain simulation package. This contains a GEANT4-based BACCARAT simulation that produces energy deposits from ER and NR events uniformly distributed in the LZ TPC. These energy deposits are then processed with LZLAMA, which applies a NEST-based parameterized detector response model. The result is a data file containing core quantities like S1, S2 and position, but lacking more granular information that would be produced in an actual data file, like individual channel areas.

Directly simulating γ-X events using BACCARAT is extremely inefficient. Instead they were generated by a γ-X generator written by Greg Rischbieter [123], which samples the photon mean free path and uses the Klein-Nishina formula to calculate the energy deposits due to Compton scattering. This solution is more efficient because it avoids the costly time-steps of a GEANT4-based simulation.

The data produced by the custom Monte Carlo and LZLAMA require further processing to calculate the cluster size and MPAF quantities. Since these data files do not contain individual PMT information, a direct calculation is not possible. A procedure called lightloading was used to approximate the PMT hits from the position of the simulated events. The lightloading procedure uses detailed auxiliary simulations to generate a light map that can be used to fill PMT information in the data file. The cluster size and log MPAF features can then be directly calculated.

4.3.3 BDT training and tuning

5×10^6 events each from the five data classes were simulated: 1) ER SS, 2) NR SS, 3) 127Xe Cathodic γ-X, 4), 127Xe Wall γ-X, 5) Cathode γ-X. The BDT was trained on 70% of the data, while 30% (the testing set) was reserved to test its performance at each step of the training process. Between training different BDTs, several of their aspects were adjusted to maximize performance on the training data. Once a well performing model was obtained, these parameters were further tuned for performance on actual data, discussed in subsection 4.3.5. The result of optimizing the various hyperparameters are shown in Table 4.1. The BDT using these parameters was considered the final model, to be deployed in the EFT search region.

4.3.4 Evaluation on simulated data

To quantitatively assess the performance of the BDT on simulations, a procedure called a k-fold cross validation was done. The simulated data was split into 10 groups while preserving the balance of classes in each group. Ten BDTs that use the final configuration were then trained, each using a different group of data as a testing set. Table 4.2, produced by averaging the results of the ten BDTs, is a confusion matrix that shows the percentage of each class that was predicted

Hyperparameter	Value	Optimized
Number of trees	30,000	
Max tree depth	6	✓
Learning rate	0.016	✓
Training data subsample rate	0.5	✓
Feature subsample rate	0.5	✓
Minimum loss at split	2.79	✓
L1 regularization	1.05	✓
L2 regularization	1.00	✓

Table 4.1: Optimized hyperparameters used in the final model, calculated using a surrogate BDT model that takes the hyperparameters as input, trained to regress on a summary variable characterizing performance on simulated data. The final model was found to be relatively insensitive to most hyperparameters except the max tree depth and learning rate.

↓ P, → T	SS NR	SS ER	Cathode γ-X	^{127}Xe cathode γ-X	^{127}Xe wall γ-X
SS NR	$\mathbf{99.54 \pm 0.03}$	0.12 ± 0.01	0.001 ± 0.001	0.075 ± 0.008	0.23 ± 0.02
SS ER	0.22 ± 0.01	$\mathbf{99.75 \pm 0.02}$	0.27 ± 0.02	0.006 ± 0.003	0.000 ± 0.001
Cathode γ-X	0.003 ± 0.001	0.12 ± 0.01	$\mathbf{97.85 \pm 0.06}$	3.62 ± 0.07	0.004 ± 0.002
^{127}Xe cathode γ-X	0.035 ± 0.008	0.007 ± 0.002	1.87 ± 0.05	$\mathbf{96.28 \pm 0.07}$	0.006 ± 0.004
^{127}Xe wall γ-X	0.20 ± 0.02	0.001 ± 0.001	0.004 ± 0.002	0.012 ± 0.005	$\mathbf{99.76 \pm 0.02}$

Table 4.2: Confusion matrix from a 10-fold cross validation run showing the correct identification rate (%) across the five classes on the diagonal, and the misclassification rate on the off-diagonals. The true classes are shown in the columns, while the predicted classes are shown in the rows.

as the other classes. The rows of the matrix represent the predicted classes (and add to one), while the columns represent the true classes. The high values on the diagonal indicate that the BDT is performing well.

The multi-class result is useful for knowing the misclassification rates for each type of γ-X event, but ultimately the γ-X cut is binary. Table 4.3 shows the same confusion matrix reduced to the single scatters and γ-X classes. To gauge how the BDT will perform in the EFT search region, the fiducial volume and the region of interest (S1c$<$ 600 phd) cuts have to be applied to the data being evaluated. The confusion matrix resulting from these cuts is shown in Table 4.4. At least in simulated data, γ-X events are rejected with high efficiency while the single scatter acceptance is near perfect.

↓ P, → T	SS	γ-X
SS	**99.81 ± 0.01**	0.194 ± 0.009
γ-X	0.19 ± 0.01	**99.806 ± 0.009**

Table 4.3: Reduced confusion matrix from a 10-fold cross validation run showing the correct identification rate (%) across the two classes on the diagonal, and the misclassification rate on the off-diagonals.

↓ P, → T	SS	γ-X
SS	**99.997 ± 0.005**	0.4 ± 1.2
γ-X	0.003 ± 0.005	**99.6 ± 1.2**

Table 4.4: Reduced confusion matrix after FV and ROI cuts from a 10-fold cross validation run showing the correct identification rate (%) across the two classes on the diagonal, and the misclassification rate on the off-diagonals.

The acceptance of single scatter events remains high even near the lower boundary of the fiducial volume, as shown in Fig. 4.10.

4.3.5 Evaluation on LZ validation data

The BDT was evaluated on several sources of actual data to ensure that it was not cutting out single scatters in the EFT region of interest, and to ensure that it was removing at least some γ-X events. However, we do not have truth knowledge about whether an event in the real data is γ-X or not. Nevertheless, the actual data can be used to build confidence that the BDT's decision making is aligned with the physical differences apparent between true single scatters and γ-X events. The tool used to make these judgments is the γ-X model.

As an example, there are stratified bands of events below the NR band in the SR1 data with S1c $\geq$ 1000 phd. These events are consistent with ^{127}Xe wall and cathodic γ-X events, and there is no other mechanism to our knowledge that could create such striations in this region. Therefore we can assume that these events are γ-X and look at if the BDT cut them out.

The datasets used are:

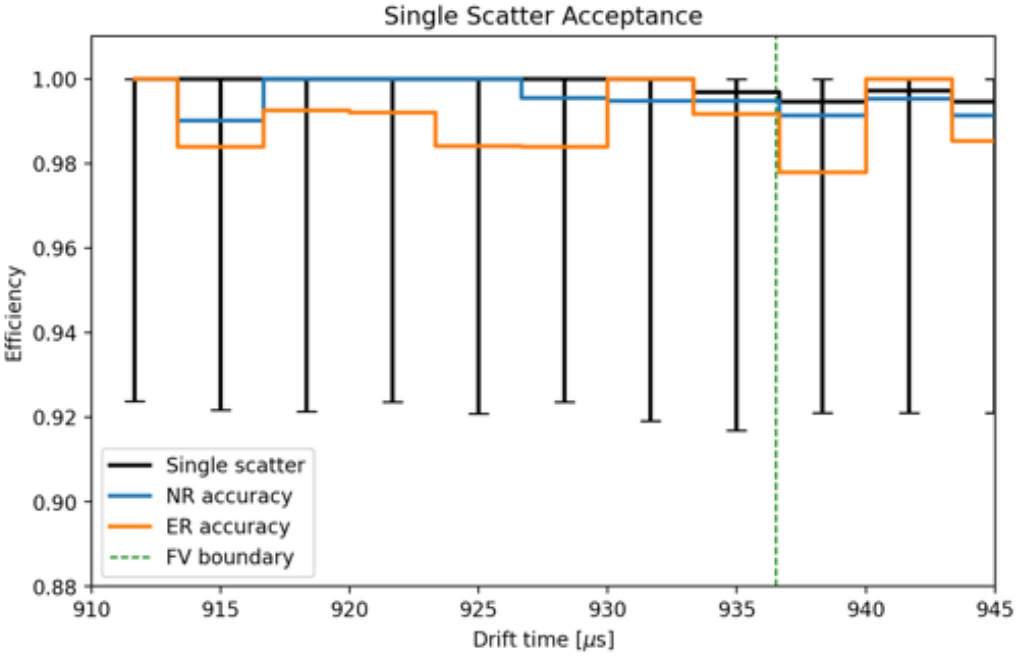


Figure 4.10: Efficiency of accepting single scatters (black) as a function of drift time. The correct classification rates of NR events (blue), and ER events (orange) are also shown, but are lower than the general single scatter acceptance. The BDT sometimes confuses ER and NR events since the only distinction it can draw between them is in S1–S2 space, which exhibits ER leakage.

1. ^{220}Rn calibration - A broad band of beta events was used to cross check the single scatter acceptance in the EFT ROI. We expect very few γ-X events in this dataset, as most of the events are ^{220}Rn daughter beta decays. However, the BDT identified several events to be γ-X. The results are shown in Figure 4.11. These events could be misclassified, but there is also the possibility that some of them are indeed true γ-X events produced by radioactivity from detector components. The detector response model is not tuned above S1= 600 phd. The resulting divergence in this regime of the β band in simulations (that the BDT was trained on) and the actual data could contribute to potential misclassifications.

2. DD calibration - A perfect nuclear recoil acceptance is a critical performance requirement for the BDT. No nuclear recoil events are expected to be cut from this dataset, and none were, as seen in Fig, 4.12. However, the neutron conduits that the DD gun shot through were not near the bottom of the detector, so this run likely did not produce any candidates for misclassification. Regardless, the DD calibration serves as a sanity check.

3. AmLi calibration - The AmLi neutron calibration provides datasets with nuclear recoil events at three locations corresponding to the height of the CSD. Therefore, a population of near the bottom of the detector is available to check the BDT acceptance on, at the expense of γ backgrounds produced by the AmLi source. The result of the BDT is shown in Fig. 4.13. The BDT did not remove any events on the NR band (except one event between the β and NR bands), but did remove what appear to be multi-scattering neutrons with a similar event topology to γ-X events.

4. High energy sideband - The BDT was tested on the SR1 background data with S1c$\geq$ 1000phd, a region where a lot of γ-X events are expected. The BDT removed the majority of events consistent with the γ-X model on and below the NR band as shown in Fig. 4.14. Still, there is a fraction of events identified as single scatter events remaining on the NR band that are almost certainly misclassifications. The relatively large rate of misclassifications (compared to the ^{220}Rn evaluation) can be attributed to the fact that this energy

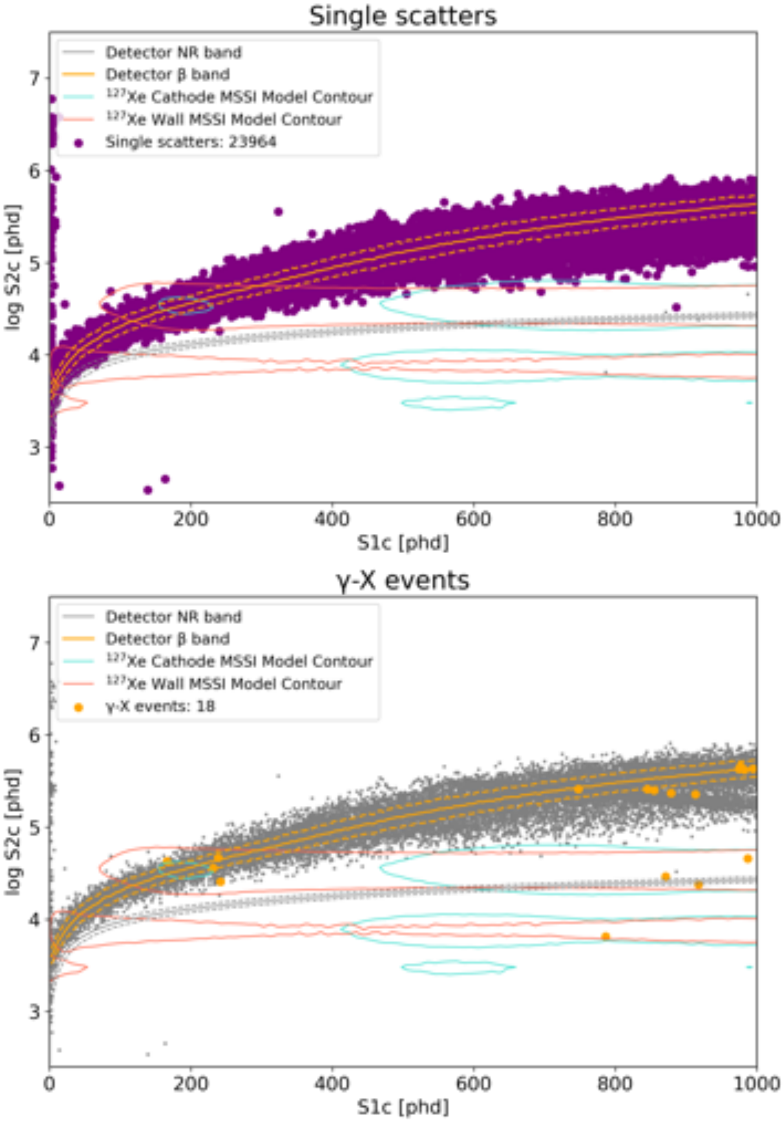


Figure 4.11: BDT classification results on ^{220}Rn calibration data. Events classified as single scatters are colored purple in the top plot, and events classified as γ-X are colored orange in the bottom plot. Underlying the colored points are all the data (grey). Also shown are γ-X model contours encapsulating 90% of counts expected in the plot axis limits for cathodic ^{127}Xe γ-X (cyan), and wall ^{127}Xe γ-X (red).

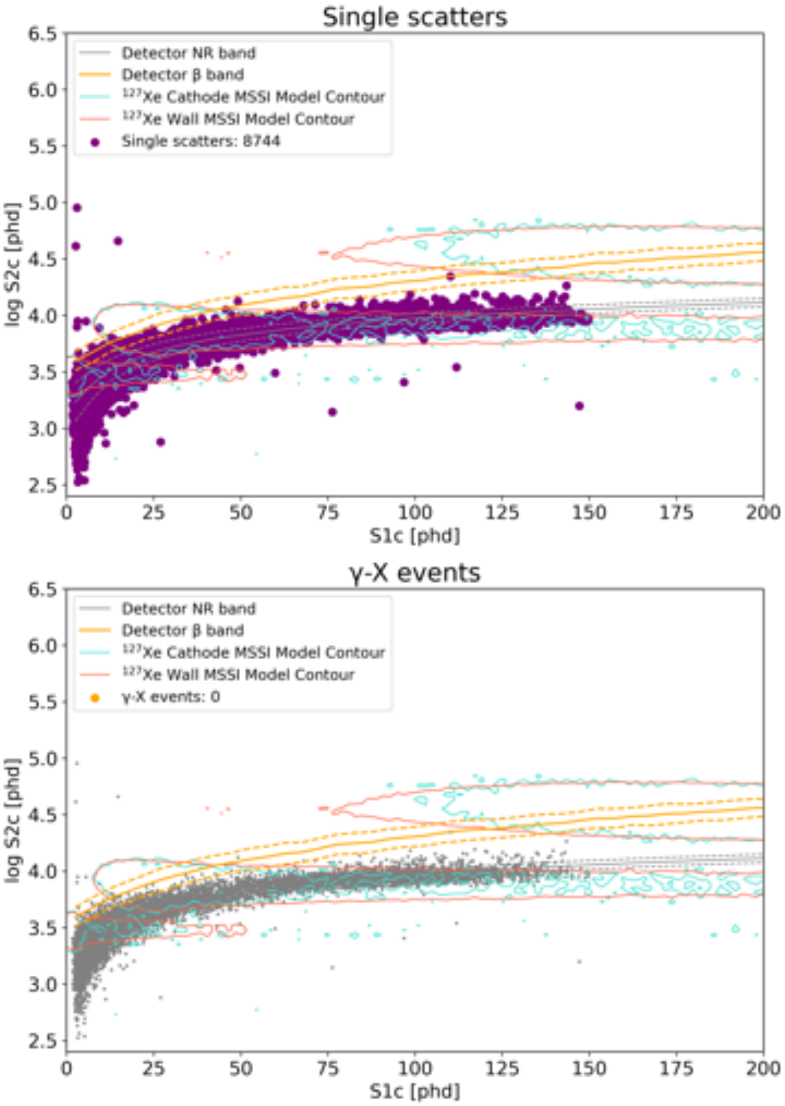


Figure 4.12: BDT classification results on DD calibration data. Events classified as single scatters are colored purple in the top plot, and events classified as γ-X are colored orange in the bottom plot. Underlying the colored points are all the data (grey). Also shown are γ-X model contours encapsulating 90% of counts expected in the plot axis limits for cathodic ^{127}Xe γ-X (cyan), and wall ^{127}Xe γ-X (red).

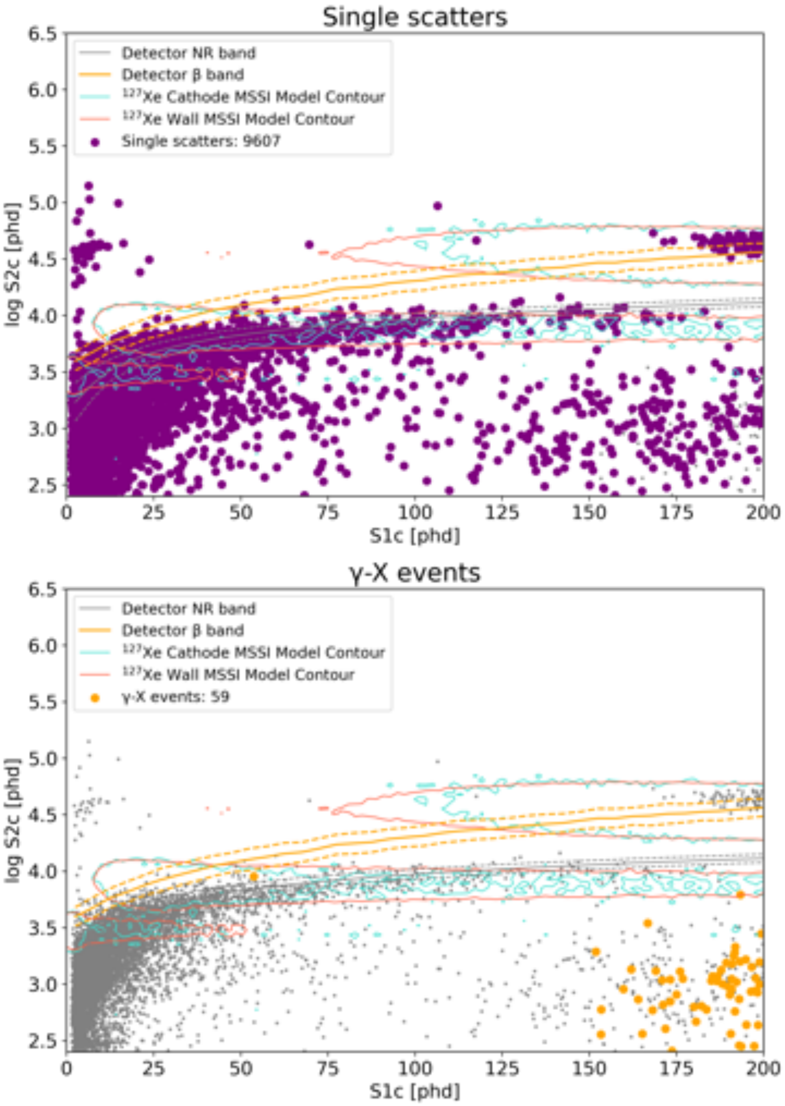


Figure 4.13: BDT classification results on AmLi calibration data. Events classified as single scatters are colored purple in the top plot, and events classified as γ-X are colored orange in the bottom plot. Underlying the colored points are all the data (grey). Also shown are γ-X model contours encapsulating 90% of counts expected in the plot axis limits for cathodic ^{127}Xe γ-X (cyan), and wall ^{127}Xe γ-X (red).

sideband is much higher than the maximum energy the that the detector response model is to. We chose not to evaluate the BDT in the $600 < \mathrm{S1} < 1000$ phd region to avoid biasing ourselves before this search region is unblinded in a future SR2 analysis.

4.4 Application of γ-X classifier on search region

After the BDT was trained, tuned, and evaluated on simulated testing data, calibration data, and the SR1 high energy sideband, it was deployed in the SR1 EFT analysis region of interest. The result in the space of S1–S2 is shown in Fig. 4.15. Of 843 events in the search region, two γ-X events were identified, consistent with the expectation (1.6 events) from the γ-X models. In signal space they are consistent with ^{127}Xe cathodic γ-X events, one from the L-shell and the other from the M-shell. In support of this hypothesis, Fig. 4.16 shows the physical location of the γ-X events to be near the cathode.

Cluster size, log(MPAF), and TBA distributions of events in the search region are shown in Fig. 4.17, along with the γ-X events. These attributes of the γ-X fall within typical values observed in simulations.

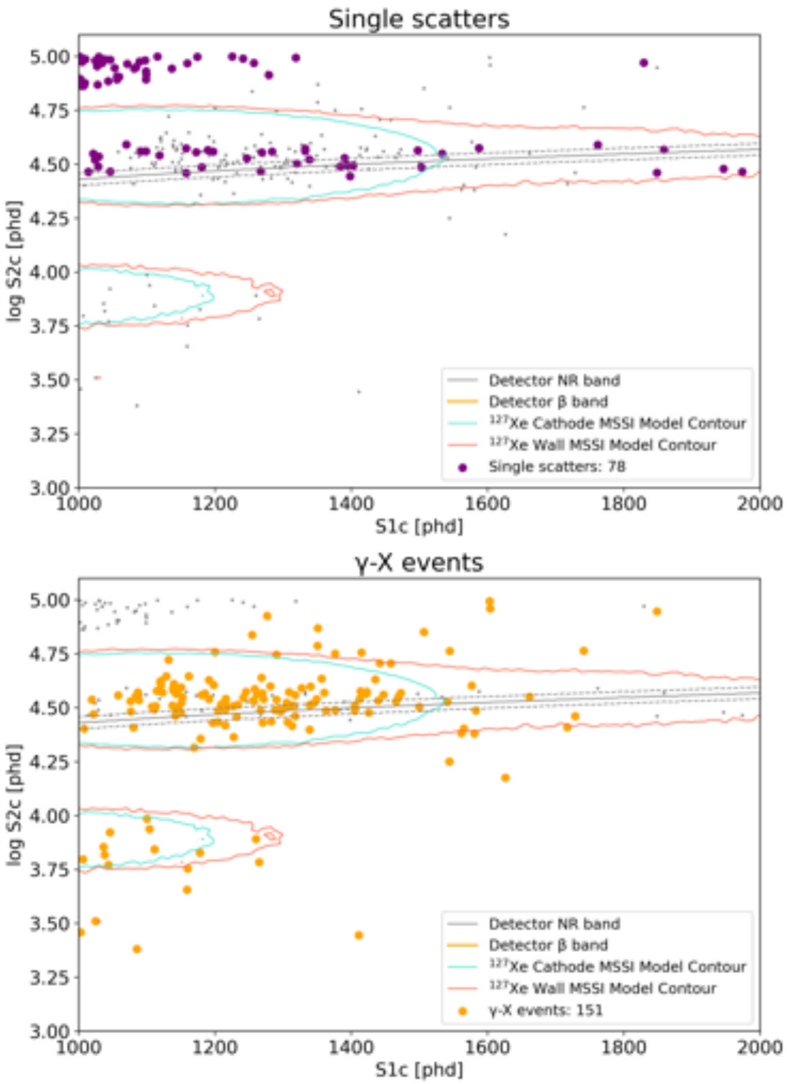


Figure 4.14: BDT classification results on high energy sideband data of SR1. Events classified as single scatters are colored purple in the top plot, and events classified as γ-X are colored orange in the bottom plot. Underlying the colored points are all the data (grey). Also shown are γ-X model contours encapsulating 90% of counts expected in the plot axis limits for cathodic ^{127}Xe γ-X (cyan), and wall ^{127}Xe γ-X (red).

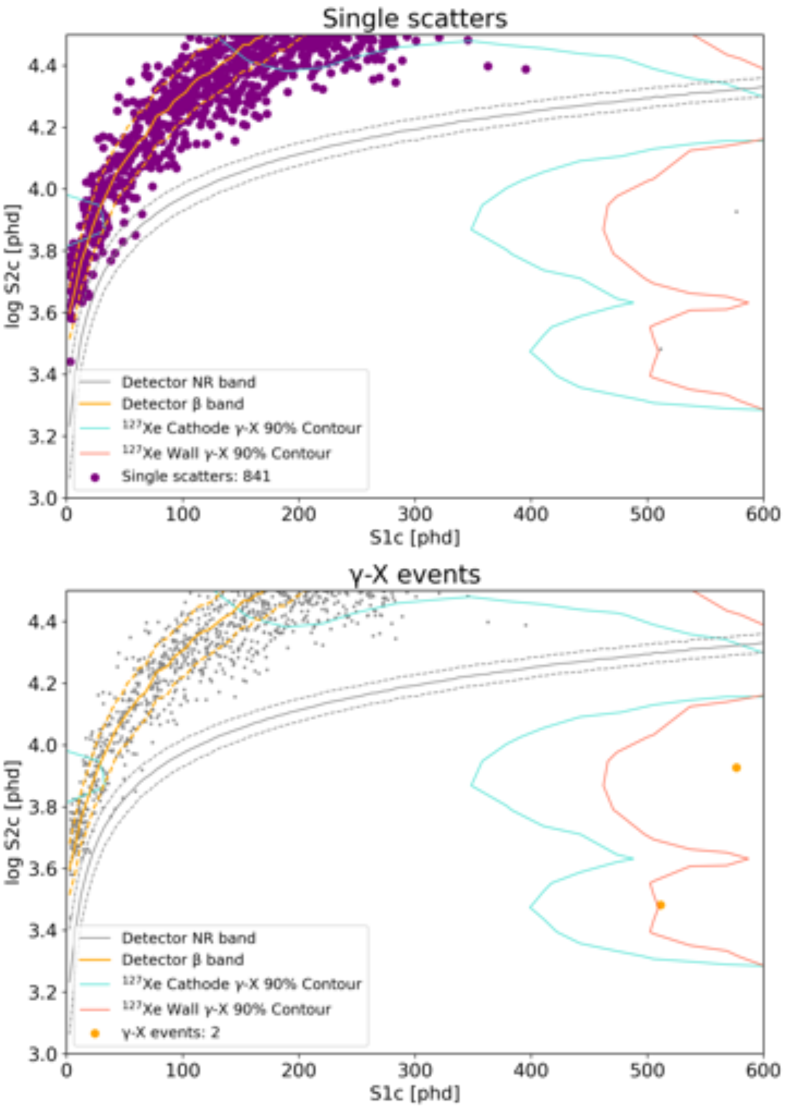


Figure 4.15: BDT classification results on the SR1 EFT search region. Events classified as single scatters are colored purple in the top plot, and events classified as γ-X are colored orange in the bottom plot. Underlying the colored points are all the data (grey). Also shown are γ-X model contours encapsulating 90% of counts expected in the plot axis limits for cathodic ^{127}Xe γ-X (cyan), and wall ^{127}Xe γ-X (red).

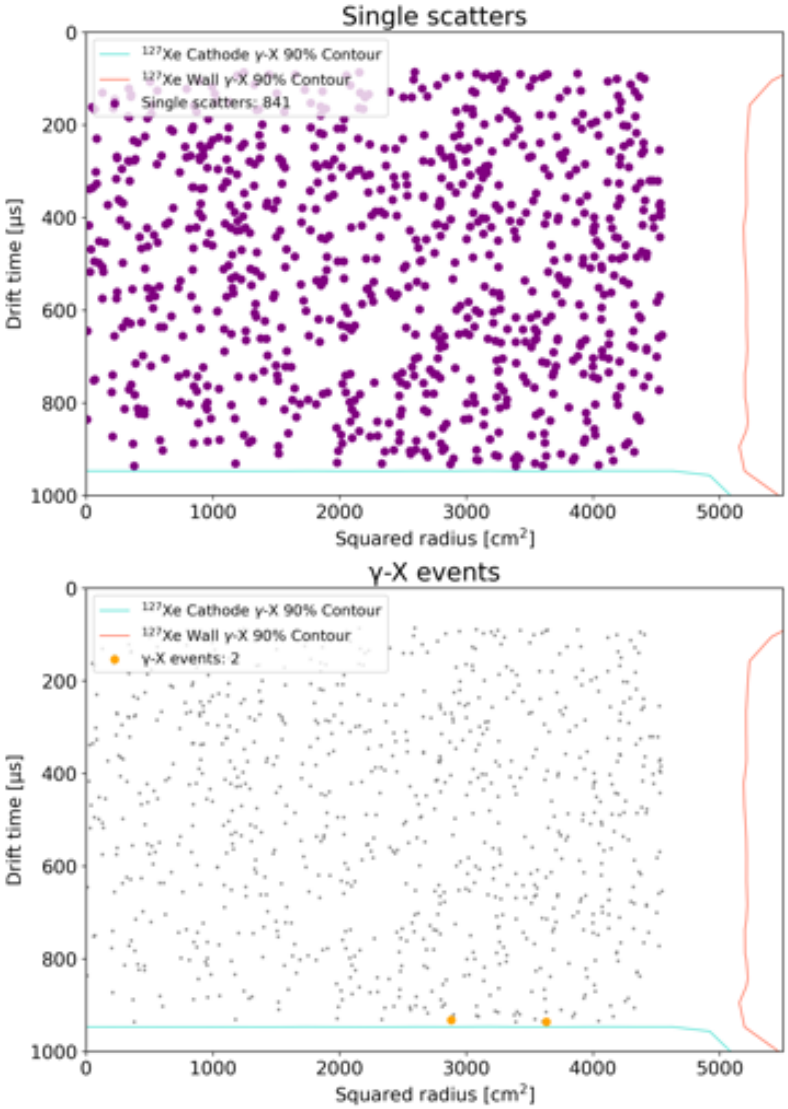


Figure 4.16: BDT classification results on the SR1 EFT search region in position space. Events classified as single scatters are colored purple in the top plot, and events classified as γ-X are colored orange in the bottom plot. Underlying the colored points are all the data (grey). Also shown are γ-X model contours encapsulating 90% of counts expected in the plot axis limits for cathodic ^{127}Xe γ-X (cyan), and wall ^{127}Xe γ-X (red).

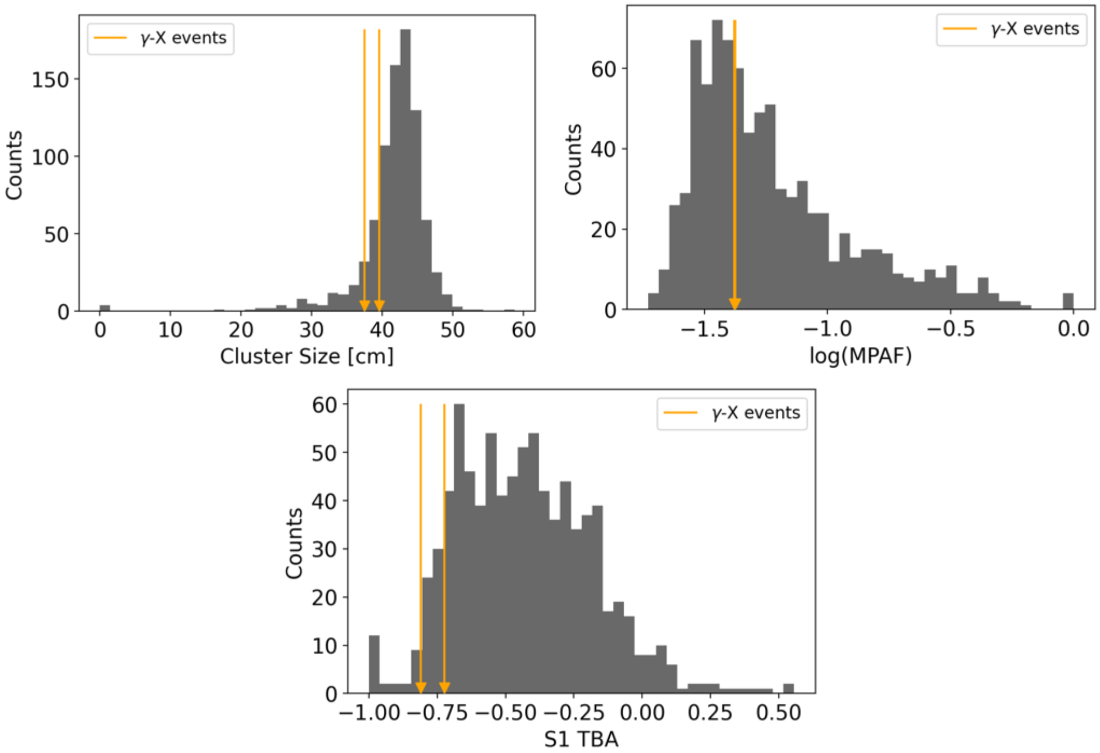


Figure 4.17: γ-X events identified by the BDT (yellow arrows) in comparison to the rest of the events in the search region, compared in distributions of cluster size (top left), log(MPAF) (top right), and TBA (bottom).

Chapter 5

Anomaly Detection in LZ Using Machine Learning

These are the sort of things people ought to look at. Things without pretensions, satisfied to be merely themselves.

Aldous Huxley, The Doors of Perception

It is easy to underestimate the pace of the machine learning revolution in experimental physics. It runs the gamut from on-chip hardware [124] to event reconstruction algorithms [125] to higher level statistical inference [126]. This chapter focuses on the uses and extensions of machine learning tools for anomaly detection applied to LZ simulations and data.

The anomaly detectors discussed in this chapter perform classification tasks but not all of them require training on auxiliary data, as in the γ-X BDT discussed in Chapter 4. Two paradigms of machine learning are represented with these anomaly detectors, *unsupervised* and *supervised* learning. The γ-X BDT is an example of supervised learning, because the BDT was trained with a large set of simulated data with labels indicating the type of event (γ-X or single scatter). Tools that leverage unsupervised learning do not require labeled datasets, but are instead deployed directly on datasets where classification is desired.

Supervised and unsupervised anomaly detection

In general, anomalies are defined as deviations of a quantity from expected values. The number of events that deviate from the expectations and the magnitude of the deviation are important factors that go into the classification decision, which differ among various tools. Unsupervised anomaly detectors are effective in identifying salient features of datasets such as outliers or clusters, which are subject to change with the addition or removal of data. Thus the classification decisions of unsupervised anomaly detectors are a function of the dataset they are deployed on. On the other hand, supervised anomaly finders are immutable once trained, e.g. the prediction for a single datum does not depend on neighboring data. Further, the availability of choice in training data offers flexibility in targeting different anomaly populations. Figure 5.1 shows an example of how the anomaly detection strategy changes depending on whether an unsupervised or supervised approach is taken.

The nature of the anomalies themselves are diverse. They can range from shape deformations of S1 and S2 pulse waveforms to unexpected correlations in a high-dimensional datasets with reconstructed quantities. They may originate from data processing inefficiencies or physical effects. Further, anomalies may already be known or lie in wait to be discovered. As a result, the vagueness of the context surrounding an anomaly makes looking for them challenging. To address this the tools presented here have a degree of interpretability that provides information on why a datum or a group of data are anomalous.

5.1 Supervised anomaly detection

Supervised learning is defined by its use of labeled datasets to train models. For example, the γ-X BDT in Chapter 4 was trained on simulated data that consisted of seven features, in addition to a label encoding the type of event. The trained classifier was then used to predict the labels of previously unseen data. If training data corresponding to anomaly distributions exist, an appro-

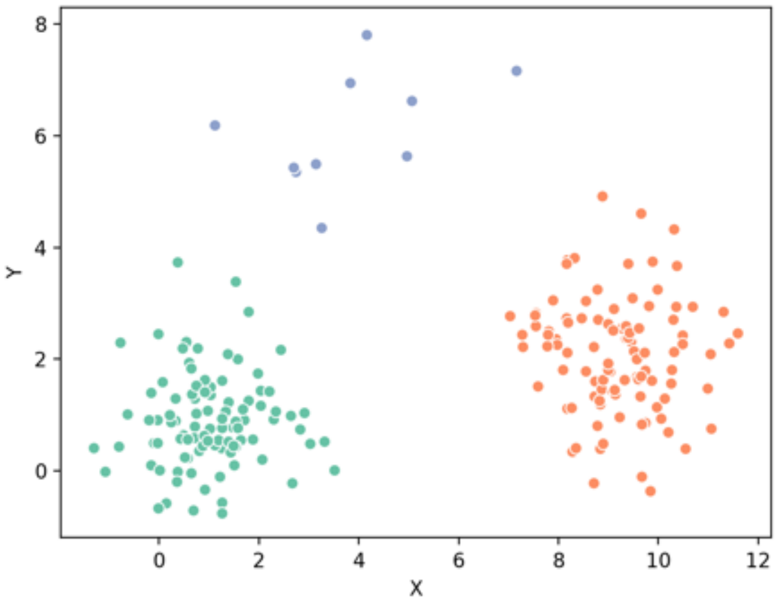


Figure 5.1: This example dataset has two main distributions (green and orange) points, and a smaller set of points (purple) shown in two arbitrary dimensions x and y. An unsupervised learning tool would be able to recognize patterns in the dataset such as the anomalous purple points and the two distinct clusters. With supervised learning, classifying either the two main populations or the sparse purple population would require the tool to be trained with many samples from the underlying distributions, which may not always be available. An alternative method to identify the purple anomalous points is to train a tool to reconstruct the green and orange distributions, leaving the anomalies to be intentionally misreconstructed. Each of these methods is discussed in the text.

priate anomaly detector can be trained, as in the first example of waveform anomalies presented here. A second technique for anomaly detection using supervised learning takes advantage of the fact that anomalies are misreconstructed in networks that are trained to replicate their inputs, and is presented in subsection 5.1.2.

5.1.1 Convolutional neural networks for pulse anomalies

The properties of events used in the analysis are reconstructed from the areas, shapes, and timings of pulses in the PMT waveforms [5]. These pulse characteristics could be distorted by bugs in the data processing chain or by physical effects in the detector, leading to errors in reconstructed quantities. Therefore, anomalies of a physical origin could be identified by the deformations in pulse shape. The prospects of using machine learning as a tool for waveform classification is explored here, using γ-X events as a test case.

γ-X S1 pulse

Chapter 4 identified seven features that could be used to discriminate γ-X events. Here the focus is on the S1 pulse waveform, which for a γ-X event is the summation of two separate scintillation pulses, as opposed to the singular scintillation S1 of a single scatter. The time difference between the multiple scatters of a γ ray are too small to be resolved by the LZ digitizers, but pulse shape of a γ-X S1 pulse is expected to be different from that of a single scatter because one of the scatters occurs outside the TPC where the electric field is different.

A convolutional neural network (CNN) was used to gauge the classification power afforded by the differences in S1 pulse shapes between γ-X and single scatter events. The shape differences alone were found to be insufficient to perform this classification. However, by providing features identified in Chapter 4 as additional inputs to the CNN, a similar performance to a BDT is obtained. Figure 5.2 shows the comparison in performance between the CNN with waveforms only, CNN augmented with features, and a BDT trained only on features.

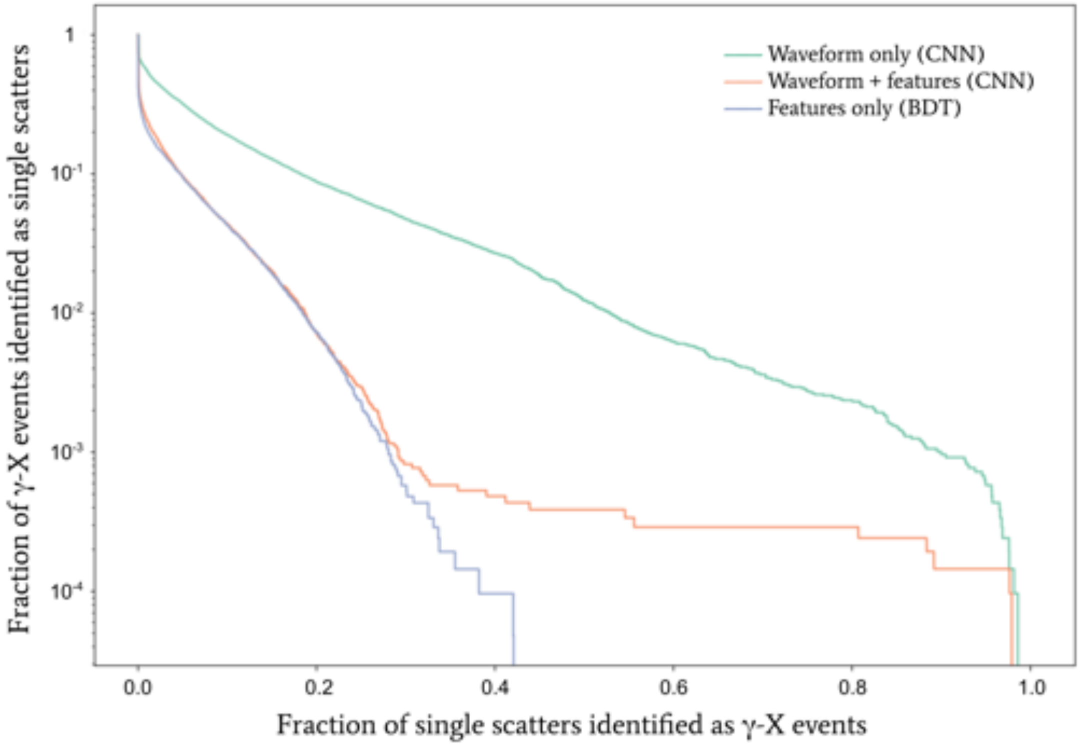


Figure 5.2: Misclassification rates of γ-X events and true single scatters from the waveform only CNN (green), feature augmented CNN (orange), and a BDT using only features (purple). The S1 pulse shapes alone are insufficient for classification, as indicated by the poorly performing waveform only CNN. With the addition of the features, the performance of the CNN matches the BDT in the regime of low single scatter misclassification.

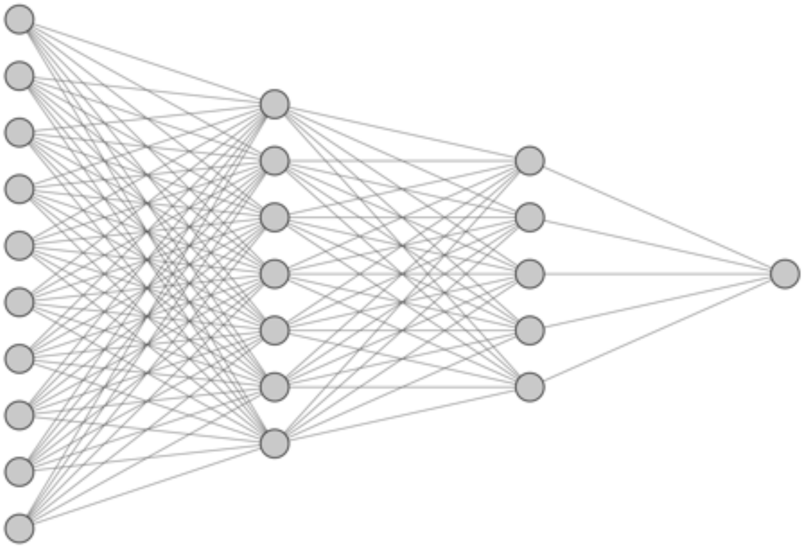


Figure 5.3: An example neural network architecture with 10 nodes in the input layer, 7 and 5 nodes in the two inner layers, and a single node on the output layer. All nodes on adjacent layers are connected, hence this is a *fully connected* neural network.

Convolutional neural network description

CNNs are a class of artificial neural network, which in their simplest form are functions that map an input vector of data to predictions. An example of an neural network that takes in a vector of length 10 and outputs a scalar is shown in Figure 5.3. Each node is an artificial neuron, which takes in a weighted sum of inputs from the previous layer and maps them to a scalar output via an activation function with adjustable parameters that control the strength of the neuron's output[1]. Artificial neural networks are trained with labeled datasets, i.e. a set of vectors and their corresponding labels, during which the network parameters are adjusted via a training algorithm to minimize the error between the network predictions and the labels.

A CNN is designed to performs classification on vectors or matrices with correlated patterns such as time-series data (waveforms) or pixel images (photos). To reduce the computational cost of processing detailed inputs, CNNs possess an interface that reduces the size of the input vectors or matrices while retaining the correlations [127]. The interface consists of convolutional

---

[1] The inputs and outputs of an artificial neuron are analogous to dendrites and axons of biological neurons.

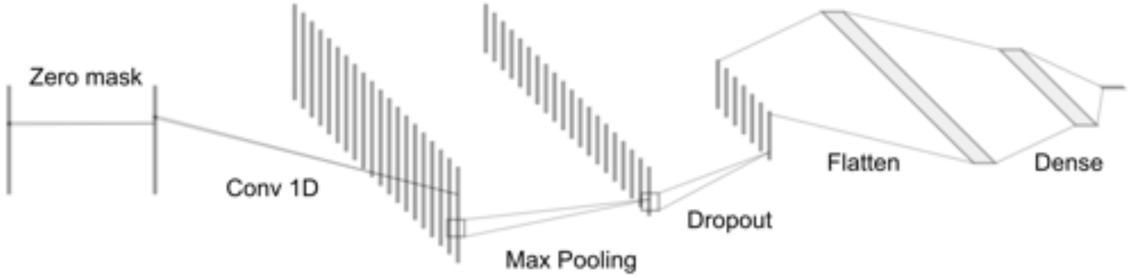


Figure 5.4: Schematic of a 1D CNN. The input waveform is truncated to 50 samples (zero mask) before being convolved with a set of kernels (first stack of vertical lines). Each kernel output is pooled, where the size of the data is reduced by combining the output of adjacent nodes. The size is further reduced by randomly dropping outputs before flattening the nodes into a regular fully connected neural network with a scalar output.

layers and pooling layers, after which the information is typically transmitted to a fully connected neural network. A schematic of a CNN is shown in Figure 5.4.

For one-dimensional waveform inputs, the convolutional layer convolves the input vector with kernel vectors that are smaller than the size of the input, such that abstracted features of the waveforms are extracted. In the case of the γ-X CNN, each S1 pulse was truncated to be 50 samples long (500 ns), and a typical kernel was 7 samples long. An example kernel is $[0,0,1,1,1,0,0]$, which has a smoothing effect when run over the waveform. 250 different kernels are run over each input vector, and the resulting layers are passed to a pooling layer which reduces the size of the data by combining the outputs of neuron clusters within each convolved waveform before they are provided to the fully connected network. The pooling layer employed ‘maximum pooling’, where only the maximum of adjacent neurons was taken as the input to the next layer. The final step before the regular fully connected network is the dropout layer, which randomly drops 30% of the neuron outputs to further reduce size and prevent overfitting [127].

Training and results

Training data was prepared using full-chain simulations that produced waveforms. Uniformly distributed 70 keV electrons were simulated in the reverse and forward field regions, which result in interactions with different light yields due to the difference in electric field magnitude. The γ-X S1 pulses were created by lining up and adding a reverse field region S1 pulse and a forward field region S1 pulse. A single scatter S1 was just taken to be a forward field region S1. All pulses were normalized to unit area, and truncated at 50 samples from the start of the pulse. The γ-X events were assigned a value of 1 in prediction space, and the single scatters were assigned 0.

The result of the waveform CNN on simulated evaluation data is shown in Figure 5.2. While it performs better than a random classifier, it is unusable for any useful classification. No hyperparameter or architecture optimization was conducted. Even though optimizing the model may increase the performance, it is likely that there are no differences between the pulse classes as simulated that provide discrimination power. It is even possible that the better-than-random classification exhibited by the CNN is driven by pathologies of the γ-X pulses that were introduced during the waveform preparation. However, the γ-X CNN provided a starting point for the introduction of additional features while retaining waveform information.

Multi-headed convolutional neural networks

The inputs to the waveform CNN were extended to include cluster size, TBA, and the maximum area fraction of the bottom and top PMT arrays. These features were introduced in the fully connected region of the CNN, bypassing the data reduction interface, as shown in Figure 5.5.

The multi-headed CNN provides a way to incorporate both waveforms and reconstructed quantities in a classification task. The performance of the multi-headed CNN exceeds that of the waveform only CNN, as seen in Figure 5.2, confirming that these quantities are important. The scalar output of the multi-headed CNN is shown in Figure 5.6, indicating good separation of the two pulse classes. The confusion matrix of the multi-headed CNN, with the CNN output cut at 0.5, is presented in Table 5.1. However, the multi-headed CNN does not outperform the BDT, which did not use waveforms. The reasons for this are unclear, as they are two different models

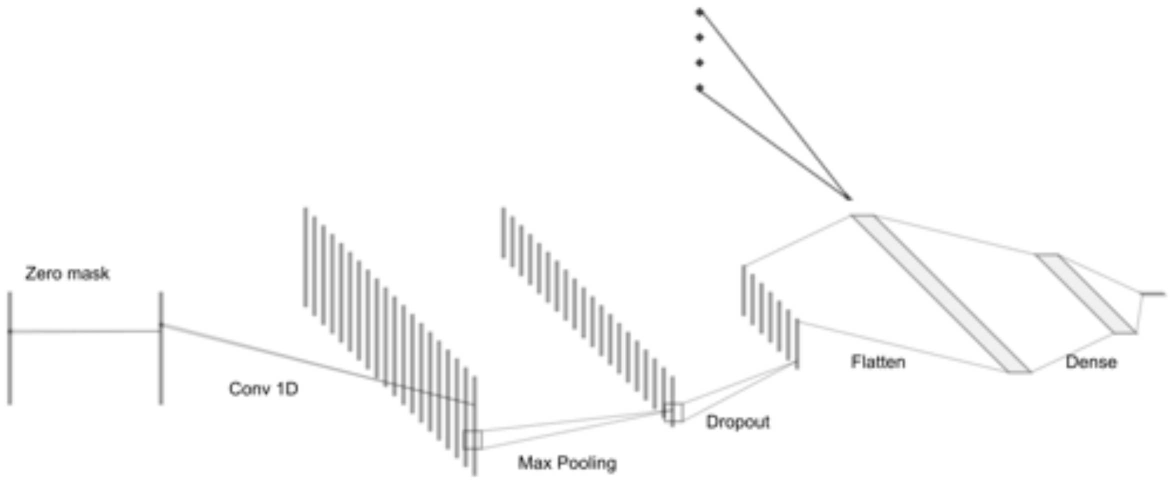


Figure 5.5: Schematic of a waveform CNN with four additional features introduced in the fully connected region.

↓ P, → T	SS	γ-X
SS	**93.3**	7.3
γ-X	6.7	**92.7**

Table 5.1: Confusion matrix from the multi-headed CNN showing the correct identification rate (%) across the two classes on the diagonal, and the misclassification rate on the off-diagonals.

trained with different aspects of the data. We surmise that this may be due to the BDT's relative ease of training, and that a proper optimization of the CNN would improve it. On the other hand, waveforms simply might not contain relevant information for identifying γ-X events, and their inclusion confused the CNN by introducing noise.

Inner layer clustering

The inner layers of the waveform CNN were found to contain characteristic features of the pulse shapes that provided useful classification information. Figure 5.7 shows the result of applying various convolutional kernels on an example waveform; activations are prominent near the start and end of the pulse. These features of the waveform, once propagated through the pooling and dropout layers, are expressed as neuron outputs in the fully connected region of the CNN.

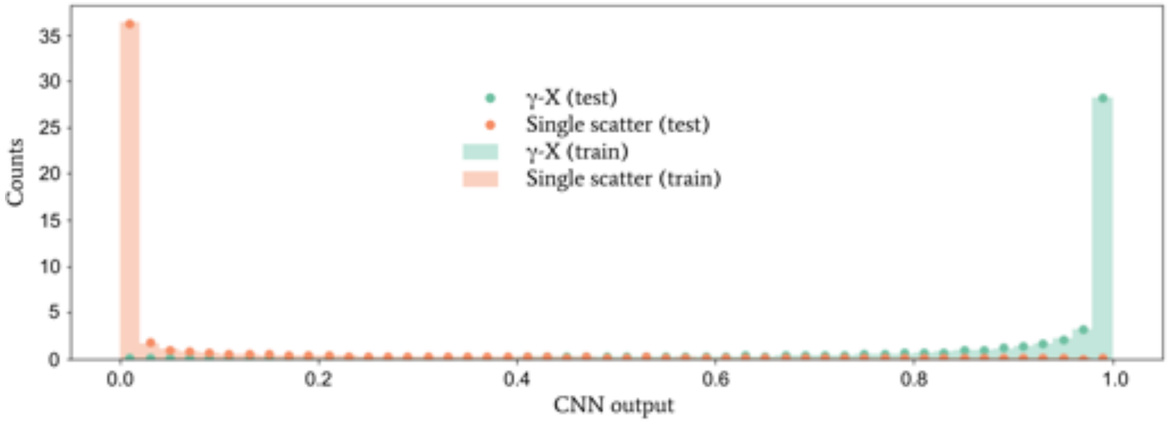


Figure 5.6: Distribution of multi-headed CNN model output, with the true γ-X events in green, shaded for the training set, and points for the evaluation set. Similarly in orange are the single scatter events.

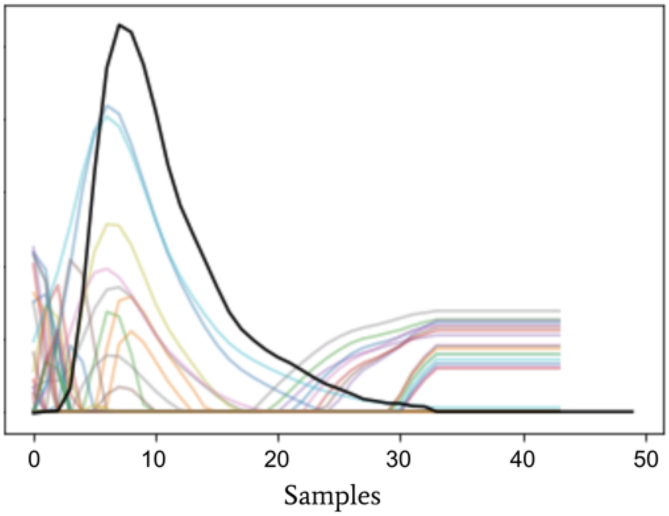


Figure 5.7: The original S1 waveform is shown in black, and the result of convolution from 10 different kernels are shown in the various colors.

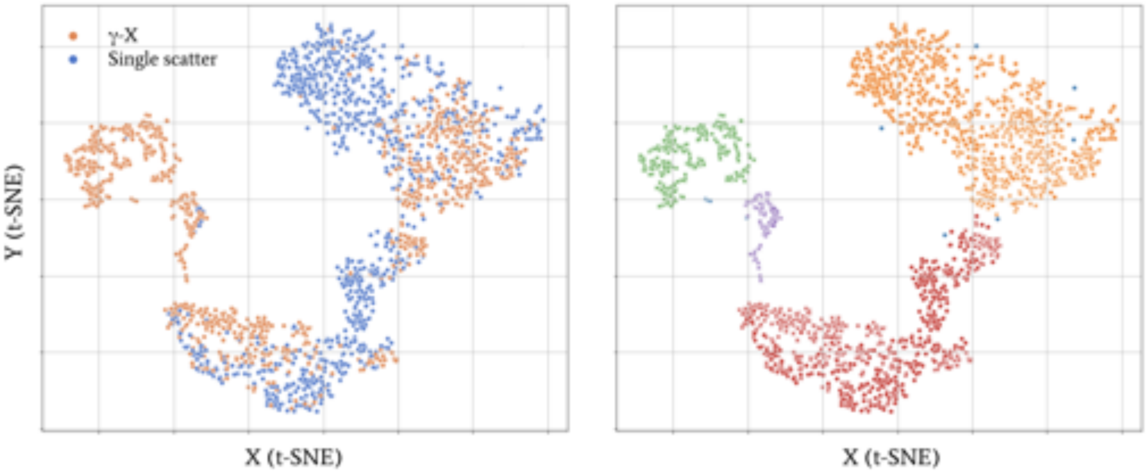


Figure 5.8: Distributions of the waveform CNN's intermediate layer output reduced to two dimensions using the t-SNE algorithm. Both plots show the same data points and the axes are arbitrary, set by the t-SNE algorithm. On the left plot the points corresponding to γ-X events are colored orange and the single scatters are colored blue. The green, purple, orange, and red points on the right plot correspond to four distinct groupings of the data, clustered using the DBSCAN algorithm. Outliers identified by DBSCAN are shown in blue.

The inner fully connected layers of the CNN still contain thousands of neurons, which is difficult to analyze and visualize. Therefore, the vector of inner layer activations is reduced in size from a length of 5,000 to 2, using a tool called t-distributed stochastic neighbor embedding (t-SNE) that was used in Maris Arthurs' thesis [82]. T-SNE embeds data into lower dimensions such that nearby points in the lower dimensions correspond to similar points in the full feature set [128]. The distribution of the t-SNE compressed inner layer activations are shown in Figure 5.8, where points are identified in two ways. First, the true class labels are applied to each point, indicating that there is not much discrimination between γ-X and single scatter events at this stage of the CNN. Secondly, a clustering based on population density is applied using the DBSCAN algorithm, demarking clear groupings of points [129].

While the clusters did not cleanly separate γ-X events and single scatters, inspecting the waveforms corresponding to each grouping yielded categories of pulse morphologies represented in the rows of in Figure 5.9. Among the morphologies are typical S1 shapes, pulses with boundaries that either start too early or too late, noisy pulses, and mergers that were classified as single

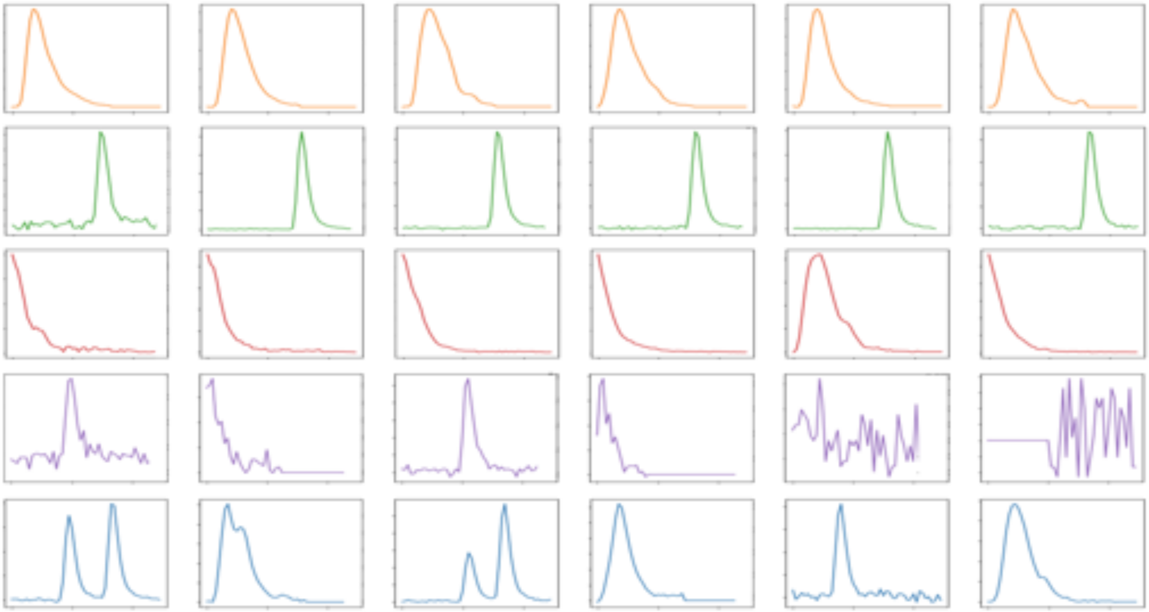


Figure 5.9: Each row shows representative pulses that correspond to the colored clusters in the right plot of Figure 5.8. The orange pulses are typical of a normal S1, the green pulses have a boundary that begins too early, the red pulses have a boundary that begins too late, the purple pulses generally have noise, and the blue waveforms are merged pulses.

pulses.

The clustering evident in the intermediate layers of the CNN is facilitated by the extraction of waveform features by the convolutional, pooling, and dropout layers. Additional work is required to explore this method as a tool for identifying waveform anomalies. Specifically, it is unknown precisely what amount of feature extraction is optimal for the reduction to well-separated clusters. Finally, unlike some of the unsupervised anomaly detectors explored in this chapter, the anomalous populations targeted with a dimensionally reduced CNN layer need not be rare.

To sum up, a CNN was trained with the intention of classifying the S1 pulses of γ-X events from those of single scatters. A CNN trained only on waveforms performed only slightly better than random, but was greatly improved with the addition of reconstructed quantities that are sensitive to γ-X events. Such multi-headed CNNs, which take both waveforms and reconstructed quantities as inputs, could be utilized in future pulse classification tasks. Further, the inner layers of the waveform CNN were found to be an appropriate space where various pulse morphologies can be identified and clustered with the aid of a dimensionality reduction tool such as t-SNE, indicating an avenue for waveform anomaly detection.

5.1.2 Anomalies as misreconstructions in autoencoders

An alternative method to perform supervised anomaly detection is with a variant of artificial neural networks called the autoencoder. An example autoencoder architecture is shown in Figure 5.10. These networks are designed to replicate the input; the number of output nodes is the same as the number of input nodes. The first half of the autoencoder encodes the input vector into a latent space smaller than the size of each event, represented by the nodes in the innermost layer, and the second half decodes this information [127]. During training the network parameters are adjusted to minimize the difference between the input and output vectors. Since it learns to replicate only the most represented features in the training data, anomalies in the evaluation data are reconstructed poorly.

An autoencoder was trained on simulated single scatter reconstructed quantities, described in Table 5.2. It was evaluated on LZ Mock Data Challenge 3 (MDC3) background simulations with only single scatter and fiducial cuts applied, along with two other anomaly detection algorithms presented in the next section. The reconstruction error is given by

$$E_{\text{autoencoder}} = \sum_{i=0}^{N} (x_i^I - x_i^O)^2, \tag{5.1}$$

where N is the length of the input data vector, x_i^I and x_i^O are the input and reconstructed components of the vector, respectively. The ability to calculate the reconstruction error for each feature i provides interpretability to the autoencoder; it allows a determination of exactly what features contributed to the error. The resulting anomalies from the autoencoder (along with two unsupervised algorithms) are shown in S1c–log(S2c) space in Figure 5.11, and position space in Figure 5.12.

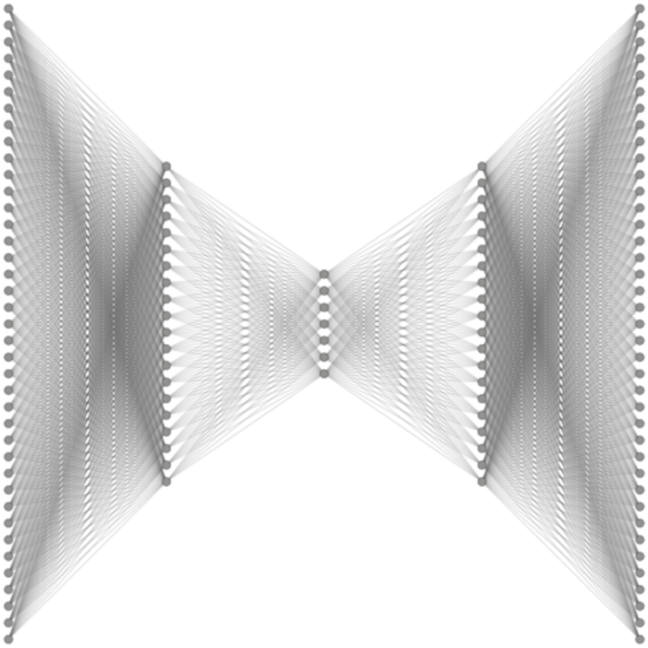


Figure 5.10: Schematic of an autoencoder with seven innermost layers. The number of input and output nodes are the same, reflecting its design to reconstruct inputs.

Feature	Description
Radius	Radial position of event reconstructed from the S2 light pattern on top PMT array.
Drift time	Vertical position of event, measured as the time difference between the S1 and S2.
TBA (S1)	Top bottom asymmetry of S1 light (defined in Eq. 4.3).
MPAF (S1)	Maximum fraction of light seen by PMTs on the bottom and top arrays (defined in Eq. 4.2).
Cluster size (S1)	Area-weighted average of PMT distances from the centroid of the PMT hit pattern (defined in Eq. 4.1).
Prompt fraction (S1)	Fraction of S1 area in the first 50 ns of the pulse, relative to total pulse area.
Area fraction time (S1, S2)	Times when pulse area reaches $\{1, 5, 10, 25, 50, 75, 90, 95, 99\}$ % of total pulse area.
Pulse area fraction (S1, S2)	Pulse area fraction within fixed time windows of $t \in \{-50\text{ns}, 50\text{ns}, 100\text{ns}, 200\text{ns}, 500\text{ns}, 1\mu\text{s}, 2\mu\text{s}, 5\mu\text{s}\}$ after pulse start.
Adjacent pulse areas	Areas of pulses immediately before and after S1s and S2s.

Table 5.2: Features of events and pulses used in unsupervised anomaly finding.

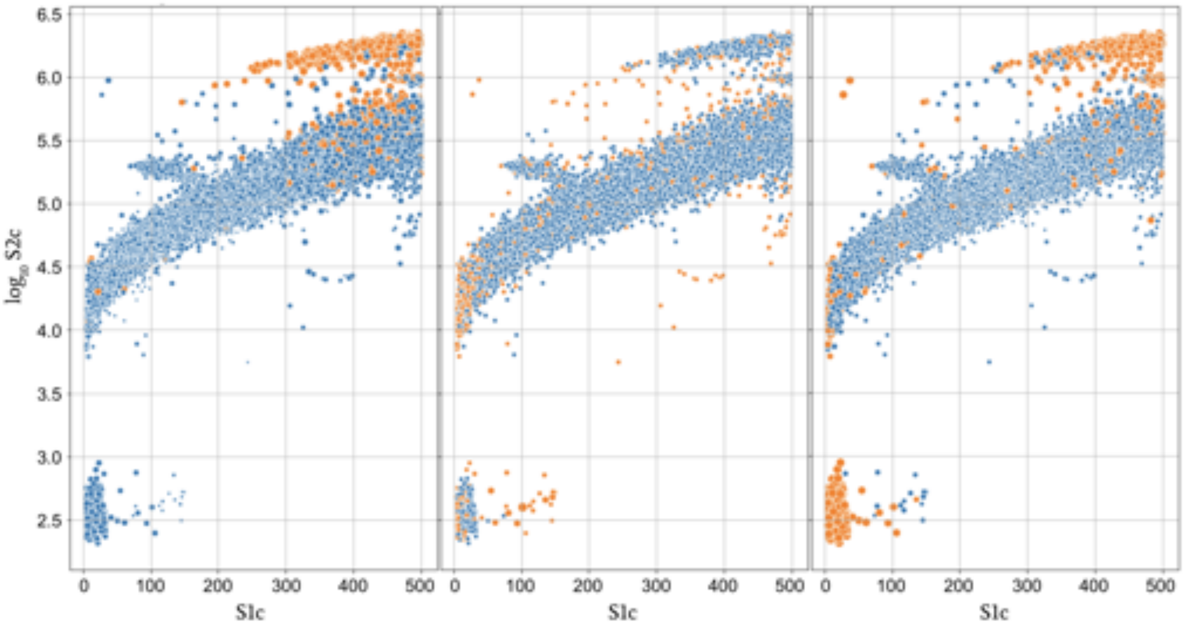


Figure 5.11: Anomalies (orange) in MDC3 background simulations in S1c–log S2c space identified by the autoencoder (left), isolation forest (center), and the local outlier factor (right). The dataset consists of 15,000 events after single scatter and fiducial cuts. Roughly 1,350 of the most anomalous points identified by each anomaly detector are highlighted here. The size of point corresponds to the anomaly score, a metric that is specific to each algorithm. Various populations of events other than the expected ER band are visible in the data; some have not been removed at this stage of the cut chain, and others are due to simulation bugs.

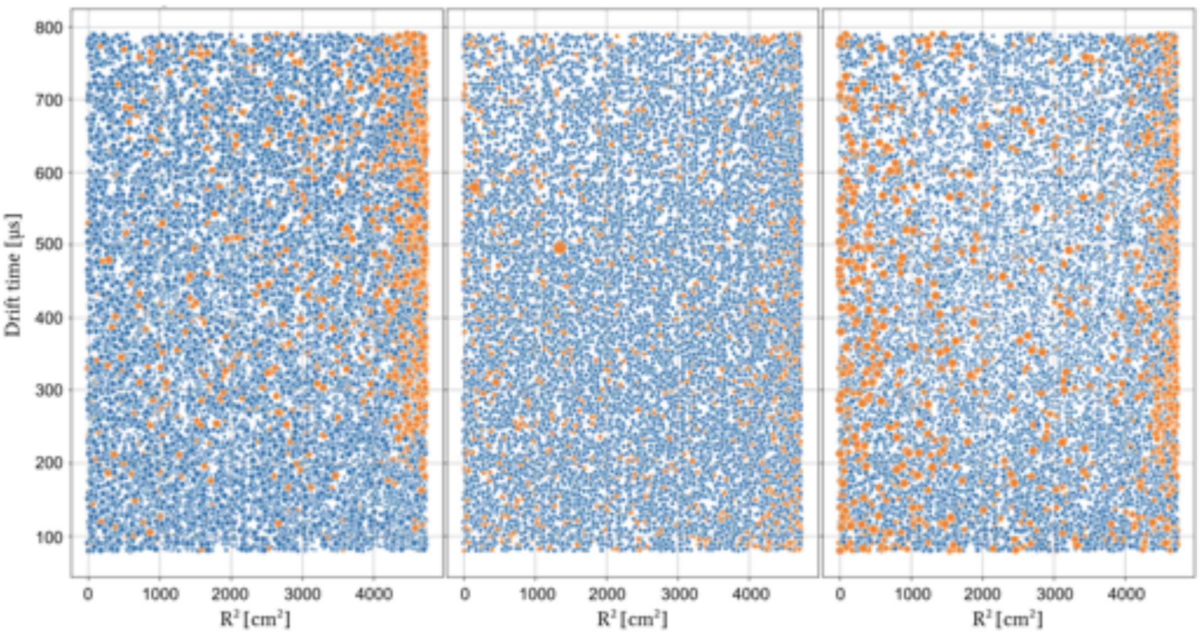


Figure 5.12: Anomalies (orange) in MDC3 background simulations in position space identified by the autoencoder (left), isolation forest (center), and the local outlier factor (right). These are the same anomalies described in Figure 5.11.

A variant that is currently under study is the convolutional autoencoder, which has an interface between the fully connected parts of the autoencoder and the input/output nodes that is designed to extract and reconstruct waveforms. However, further work is required to assess its use on LZ waveforms.

5.2 Unsupervised anomaly detection

Unsupervised learning is defined by the lack of labeled data used in training. For most unsupervised algorithms, including ones discussed here, the classification decisions are only relevant in the context of the dataset. In other words, the anomalies identified are subject to change with the dataset. The unsupervised techniques explored here were chosen for their speed and ability to process large amounts of data. The two methods are the local outlier factor [130], and the isolation forest (IF) [131]. The local outlier factor was not considered further due to the relative difficulty in interpreting results. The IF was studied further in several case studies and a procedure to improve its interpretability is introduced here.

Figures 5.11 and 5.12 show anomalies in S1c–log(S2c) and position space identified by the autoencoder, isolation forest, and local outlier factor (LOF) algorithms when they are run over the same set of data with features described in Table 5.2.

5.2.1 Isolation forest

The IF is designed to identify anomalies that are few and far between. It relies on a series of random cuts on randomly chosen features to build a binary decision tree. Anomalies are identified as the data that are isolated with the fewest number of cuts, as illustrated in Figure 5.13 in two dimensions. The anomaly score $s(h,n)$ for a point x is related to the number of cuts $h(x)$ by

$$s(h,n) = 2^{-\frac{E(h(x))}{c(n)}}, \tag{5.2}$$

where n is the number of events in the dataset, $E(h(x))$ is the number of cuts needed to isolate x averaged over an ensemble of trees, and the normalization factor $c(n)$ is given by

$$c(n) = 2H(n-1) - \frac{2(n-1)}{n}, \tag{5.3}$$

where $H(i) \approx \ln(i) + 0.577$ is the harmonic number [131].

A typical distribution of anomaly scores is shown in Figure 5.14, corresponding to S1 pulse shapes in an SR1 commissioning dataset. The tail of the distribution contains the anomalous events, which can be selected with a user-defined cut.

There are several features of the IF that make it appropriate for detecting anomalies in LZ data. Firstly, since the feature to be cut on at each decision point is chosen at random, there is no additional computing cost incurred by adding features, allowing for a high-dimensional dataset with redundant features. Secondly, the algorithm is fast because no distance or density measure

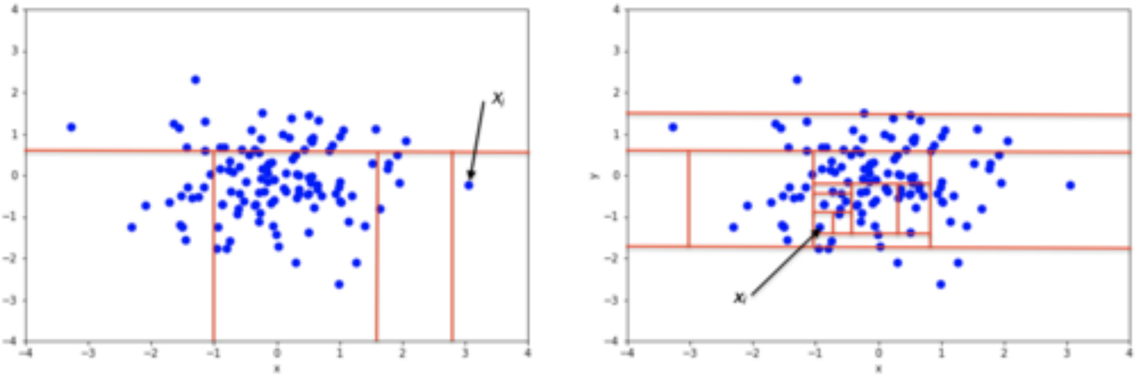


Figure 5.13: The IF recursively cuts (red lines) on the data until each datum is isolated. An outlier (left) requires a fewer number of cuts to isolate, in contrast to a point within a prominent distribution (right). Figure by Sal Borrelli.

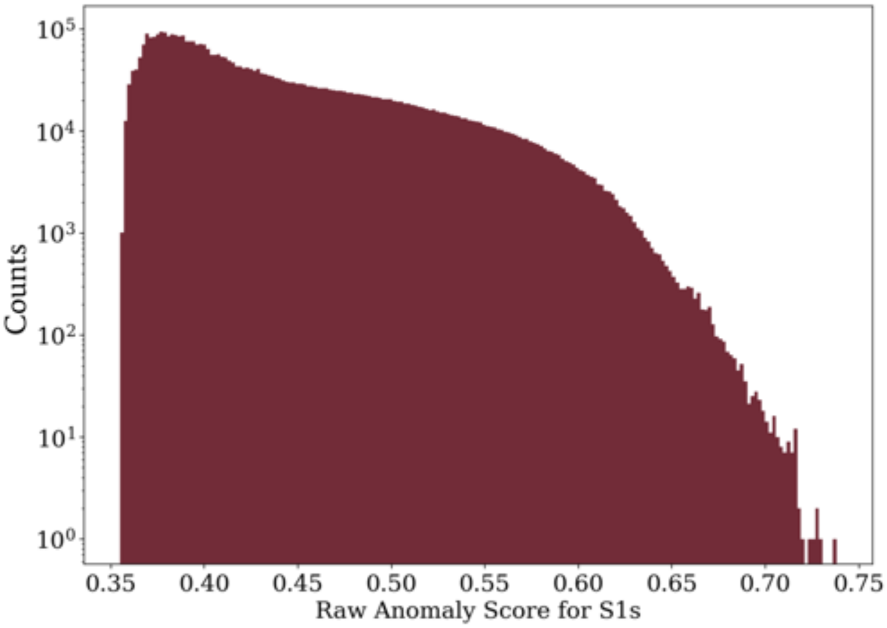


Figure 5.14: Distribution of anomaly scores for S1 pulses in an early commissioning dataset. Anomalous pulses may be identified by setting a user-defined threshold on this distribution.

has to be calculated for each point. In fact, it has a linear time complexity [131]. Finally and most importantly, the cuts within the decision trees can be assessed, providing knowledge of the features that contributed most to short decision paths, i.e. anomalies.

Feature decorrelation

The features chosen for the IF heavily impact the resulting anomalies. This choice is often up to the user and their use case. If a known or suspected anomalous population is targeted, the user is likely to know the relevant features to use. If the IF is deployed as a catch-all anomaly detector, then the user may start with a comprehensive list of reconstructed quantities and prune the list based on the results.

Typically it is best to use uncorrelated quantities, as otherwise the IF is biased to place extra weight on the correlated quantities, diluting the contribution of other features. For example, there are highly correlated variables that summarize the shape of a pulse such as the Area Fraction Times (AFTs), described in Table 5.2. The time taken to reach 25% and 50% of a pulse's area, respectively, are necessarily correlated. Figure 5.15 shows Kendall rank correlation coefficients, which measure ordinal correlations, for families of quantities that summarize pulse shapes. These correlations were reduced before the data was provided to the IF. For instance, the AFTs were decorrelated by taking the differences of adjacent AFTs, e.g. time taken to reach 50% minus time taken to reach 25% of the pulse area. Figure 5.16 shows the reduced correlation map for the pulse summary quantities.

Importance ranking

The anomalies identified by the IF, or anomaly detectors in general, can be explained with the aid of external tools. One such method is to use an innately interpretable surrogate model, a BDT for instance, that is intentionally overtrained on the output of the IF. BDTs provide feature rankings that allow the origin of anomalies to be explained. Dedicated explainer tools like SHapley

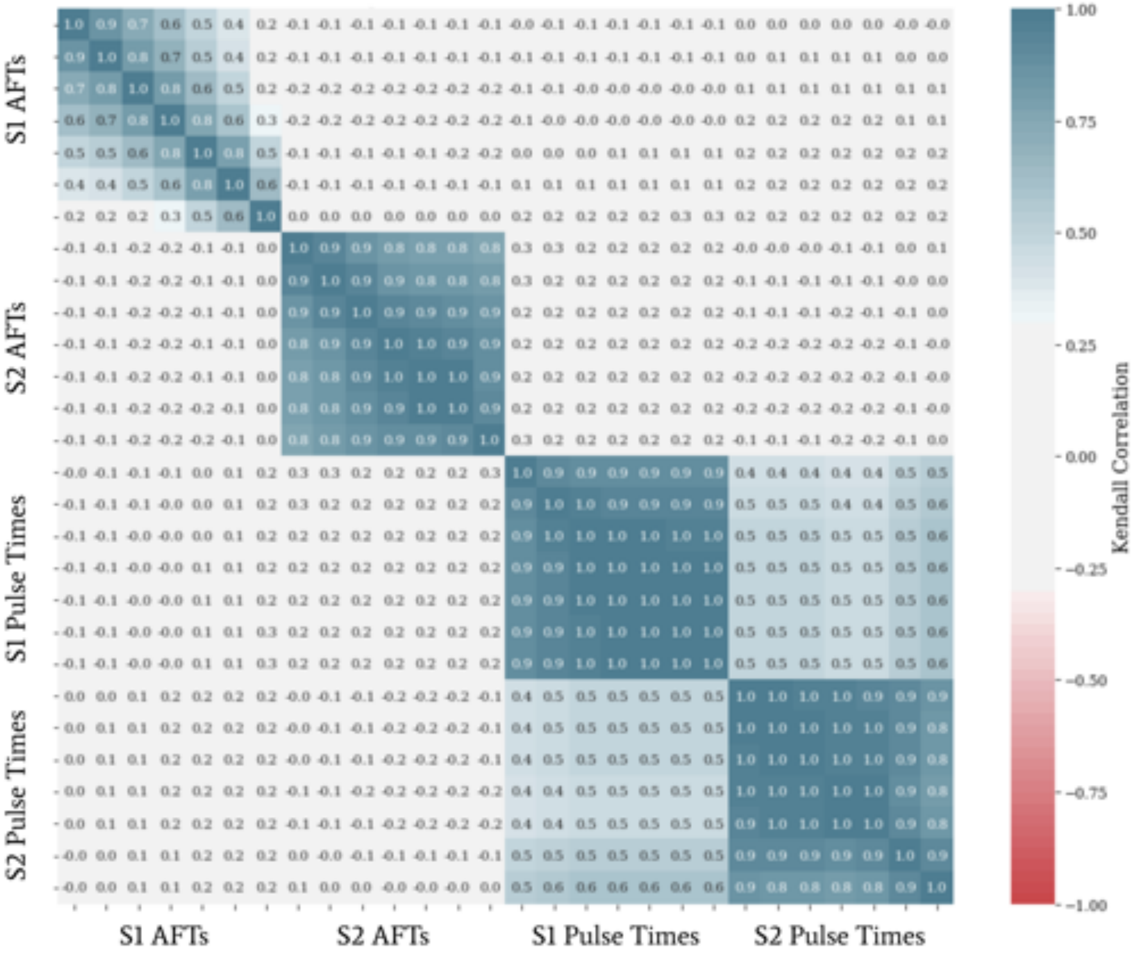


Figure 5.15: Kendall rank correlation coefficients (defined in Ref. [132]) for the AFT and pulse time families, defined in Table 5.2. Strong positive correlations exist within each family.

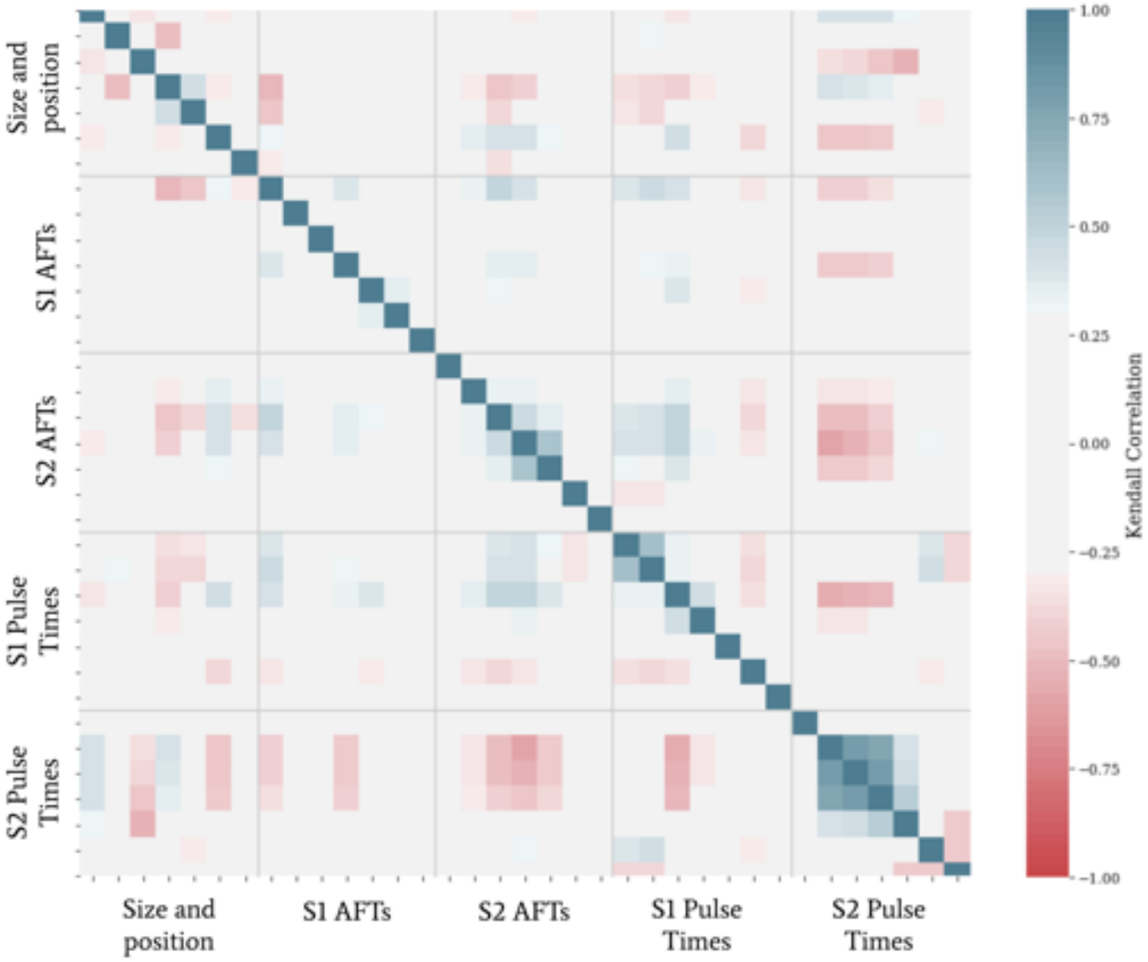


Figure 5.16: Kendall rank correlation coefficients (defined in Ref. [132]) for the decorrelated AFT and pulse time families, in addition to several features related to pulse size and event position.

Additive exPlanations (SHAP) [133], utilizes game theory to compute the contribution of each feature to the anomaly score. However, these two approaches often gave contradictory results.

It is possible to sidestep the external tools. The decision making of the IF is innately explainable if its underlying principle is taken advantage of; that outliers have short decision paths on average over all the trees in the IF. The features that were cut on in these short paths are more important than those cut on in long decision paths. An algorithm to calculate feature importances is given below.

1. Determine a threshold for decision path length. The maximum depth of trees in the IF given by $\log_2(N)$ where N is the number of events used to grow a tree. The recommended value is $N = 256$, resulting in a maximum tree depth of 8 [131]. Thus an appropriate path length threshold is 3 or 4.
2. For each event, loop through the trees in the IF and select paths shorter than the threshold path length.
3. Count which features are being cut on at each node, the depth of each node, and the number of events split at the node. Figure 5.17 shows the feature counts for an example event.
4. Condense the feature counts into a final feature ranking. One method is to assign weights to each node, with larger weights for nodes that are cut earlier, that are then summed for each feature. For example, a ‘geometric’ weight assignment could have a weight of 1 for the first cut, 0.5 for the second cut, 0.25 for the third cut, et cetera. Figure 5.18 shows the condensed feature importances using this geometric weight assignment, and Figure 5.19 shows the location of the outlier in the space of the two most important features. The number of events split at each node can also be incorporated into the summarization of the feature importances.

This procedure yields local feature importance, which is specific to a single event. Global feature importances can be obtained by averaging the feature counts over portions of the dataset,

↓ P, → T	γ-X	SS
γ-X	**95**	12.5
SS	5	**87.5**

Table 5.3: Confusion matrix from the IF identifying (in percentages) simulated γ-X events, with the predicted classes in the rows and the true classes in the columns.

to inspect certain populations, or over the entirety of the dataset, to assess certain features to focus on or drop.

5.2.2 Case studies

The IF was deployed on simulated and measured data in several contexts, including targeting specific anomalies such as γ-X events, aiding the scanning of waveforms by human analyzers, and diagnosing data quality issues. These cases are discussed below.

γ-X events

γ-X events were targeted with the IF using the features identified in Chapter 4. A dedicated simulation carried out during MDC3 to produce γ-X events was tested with the IF. With an appropriate selection in S1c–log S2c space, the IF correctly identified 95% of the γ-X events. The data, with IF predictions, are shown in Figure 5.20. A confusion matrix summarizing this result is given in Table 5.3.

The choice of data selection in S1–log(S2c) space was driven by two competing effects. If too much data were included, it is more likely for irrelevant anomalies to be identified, and on the other hand too little data is not conducive in building a robust and replicable IF. Preliminary studies indicate that the larger the fraction of γ-X events present in the data, the less accurate the IF becomes. More work is required to test and to determine the optimal cuts that would maximize the γ-X identification and single scatter acceptance rates. While the IF, being an unsupervised

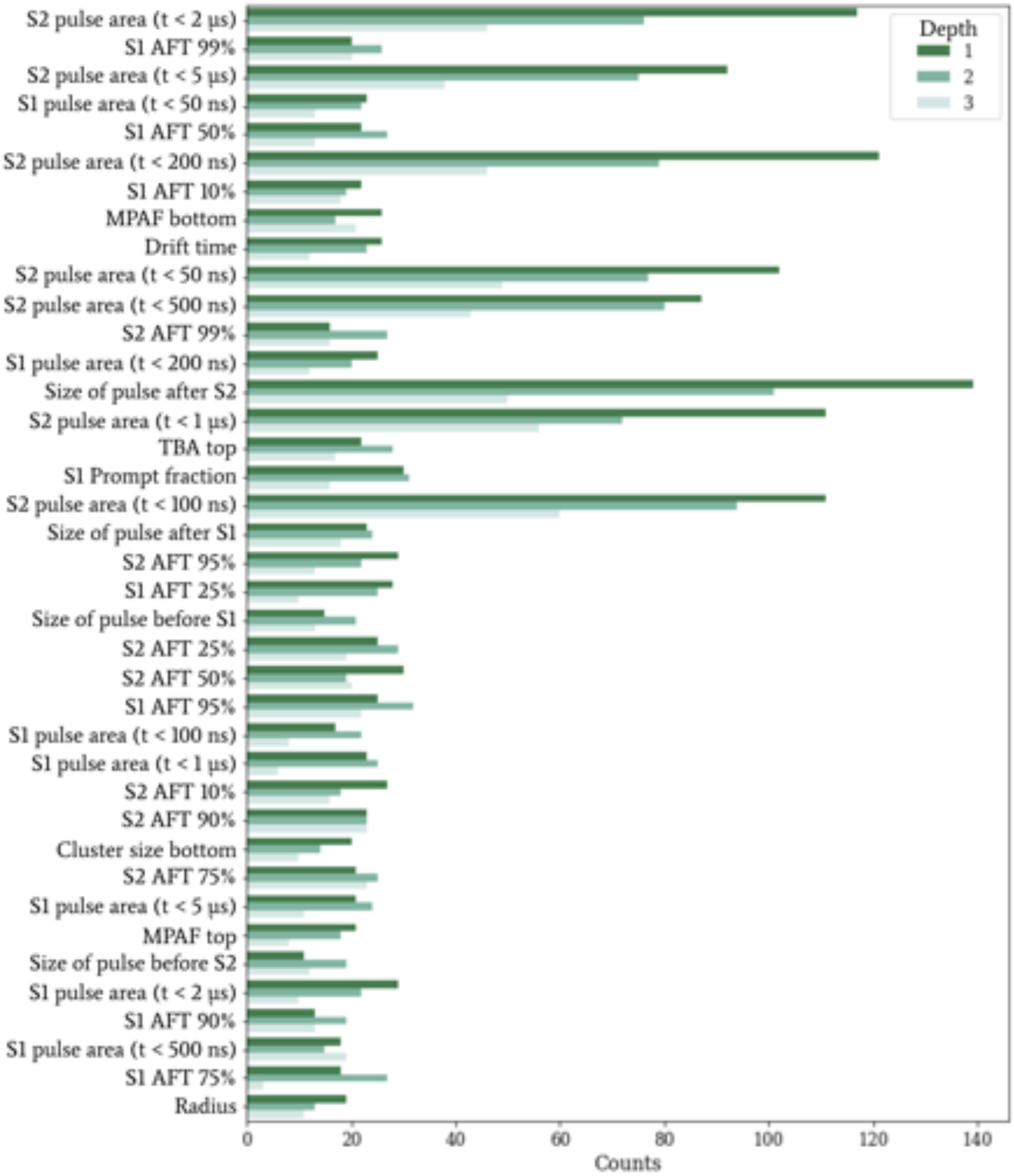


Figure 5.17: Feature counts for decision paths with three or fewer cuts for an example event. The darkest color corresponds to the first cuts in IF trees, the next darkest color to the seconds cuts, and the lightest color to the third cuts.

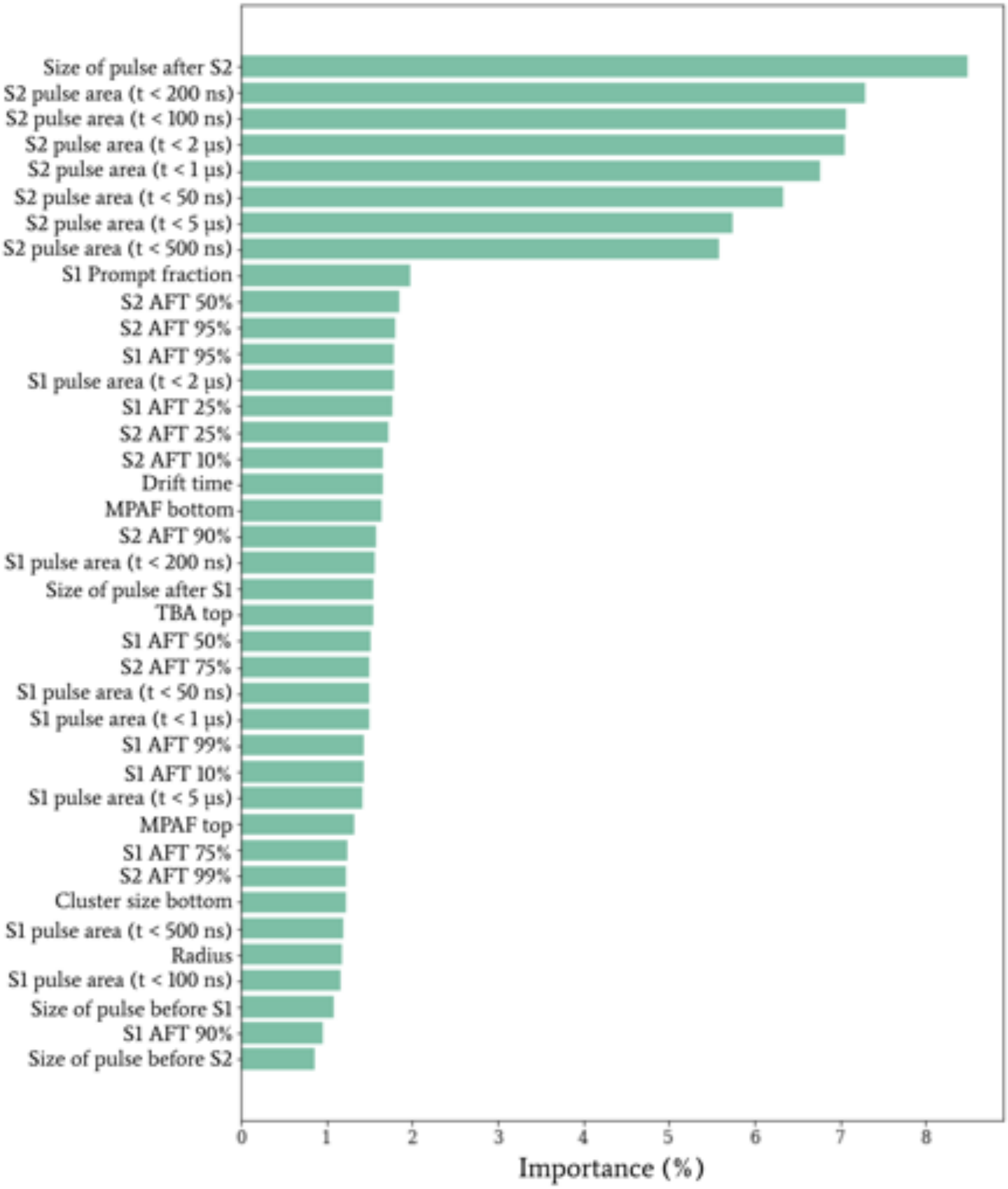


Figure 5.18: Feature importances for the example event in Figure 5.17, calculated using weights of 1, 0.5, and 0.25 for the first, second, and third cuts, respectively.

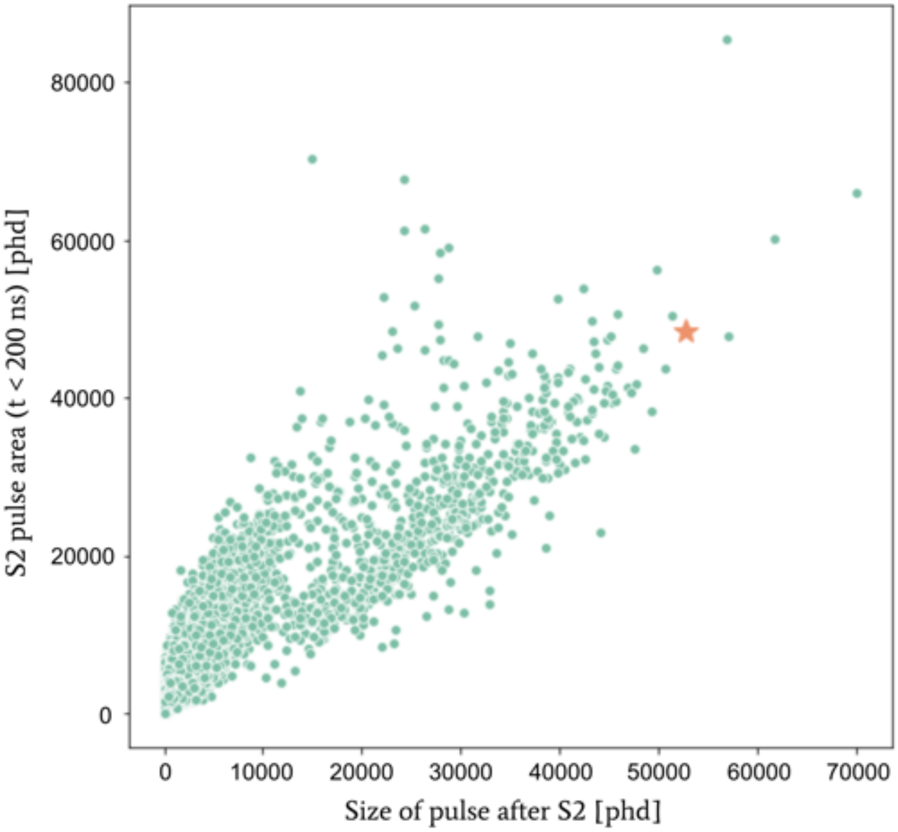


Figure 5.19: Distribution of events in the space of the two most important features. The anomaly (orange star) corresponds to the event in Figures 5.17 and 5.18.

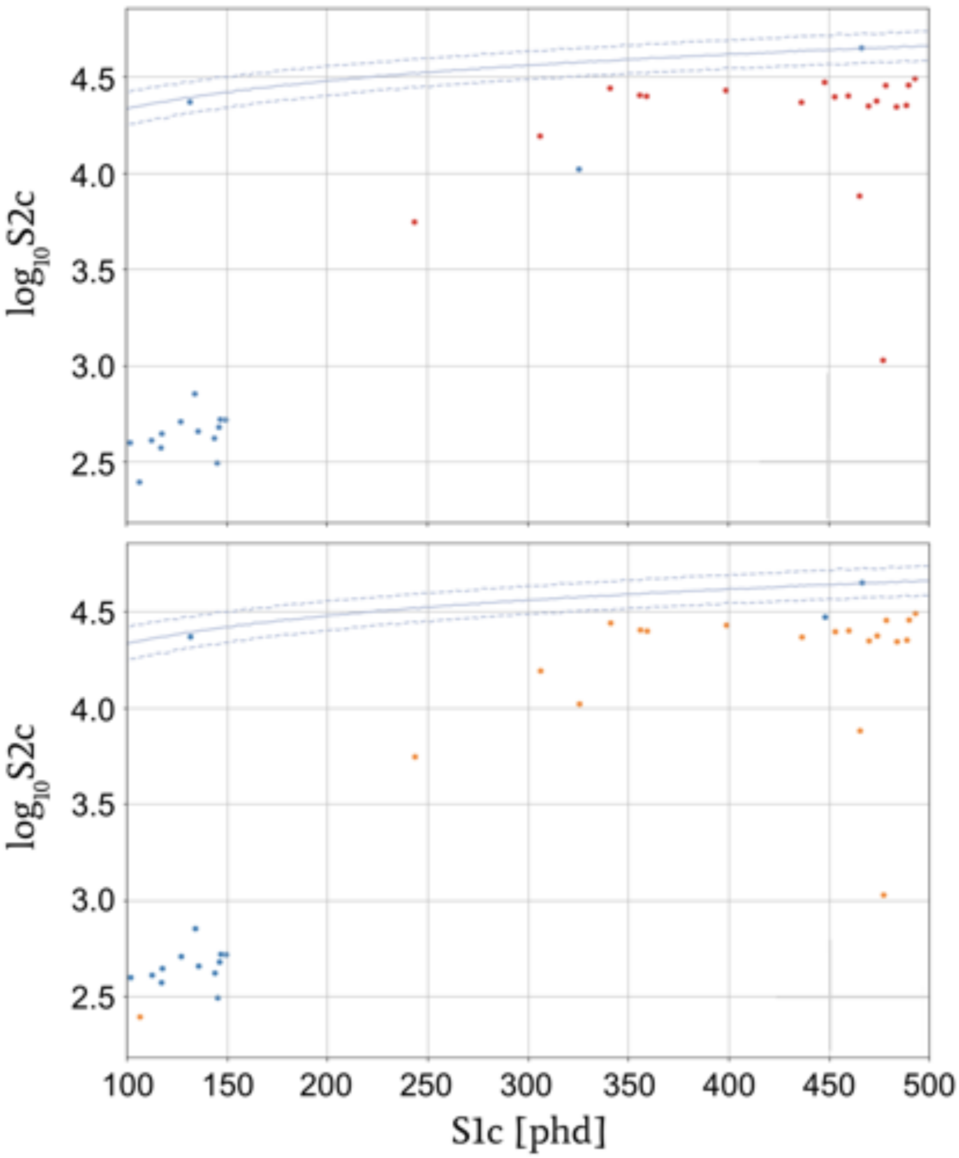


Figure 5.20: Top: Simulated data in S1c–log S2c space showing true γ-X events in red, and single scatters in blue. Bottom: Events predicted as γ-X events by the IF are shown in orange, and single scatters in blue.

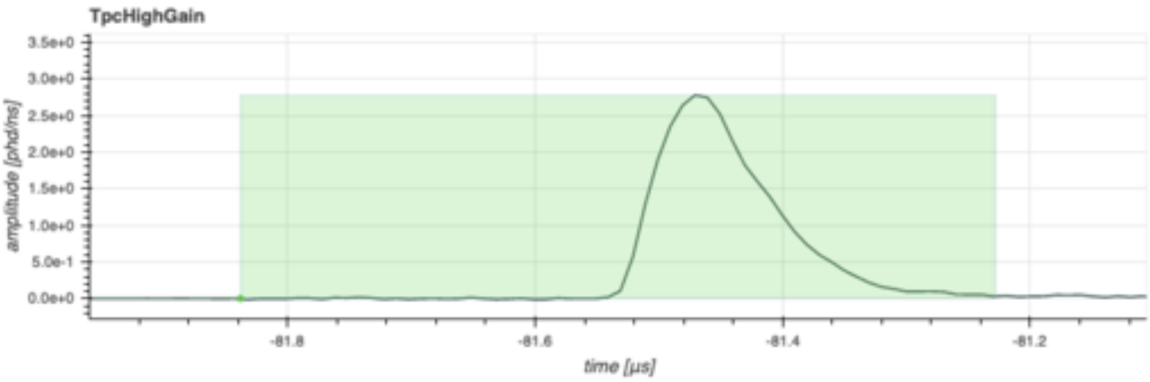


Figure 5.21: Example pulse boundary (green box) with the associated S1 pulse (black line) from the MDC3 simulations with a large gap between the start of the boundary and the actual pulse.

technique, is not suitable for actually removing γ-X events from data, it may serve as a useful cross-check on future applications of the γ-X BDT discussed in Chapter 4.

Pulse boundaries

When running over datasets with features describing pulse shape such as the AFT family, the IF found pulses with ill-defined boundaries. These pulse anomalies, whose boundaries either started early or ended late, were found both in MDC3 simulations and in commissioning data. An example from the MDC3 simulation is shown in Figure 5.21, and one from the commissioning data is shown in Figure 5.22. This issue has subsequently been corrected in the LZ pulse finder algorithm. The IF illustrates that waveform anomalies can be detected by using a relatively small set of quantities that summarize the shape of the pulse.

Periods of high pulse rate in the TPC

Right after the LZ TPC was filled with liquid xenon, there were periods in which the rate of small pulses were frequently elevated. The IF was used to diagnose the cause of these emissions, and noticed abnormally shaped S2 pulses during periods of high pulse rate. The number of pulses in each event and the anomaly score of the S2 is shown in Figure 5.23 for roughly 90 s of data

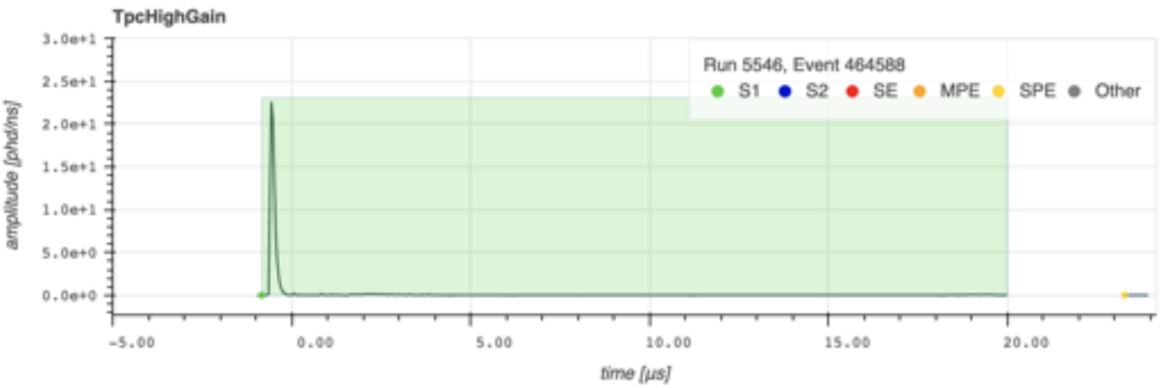


Figure 5.22: Example pulse boundary (green box) with the associated S1 pulse (black line) from commissioning with a large gap between the end of the pulse and the bounding box.

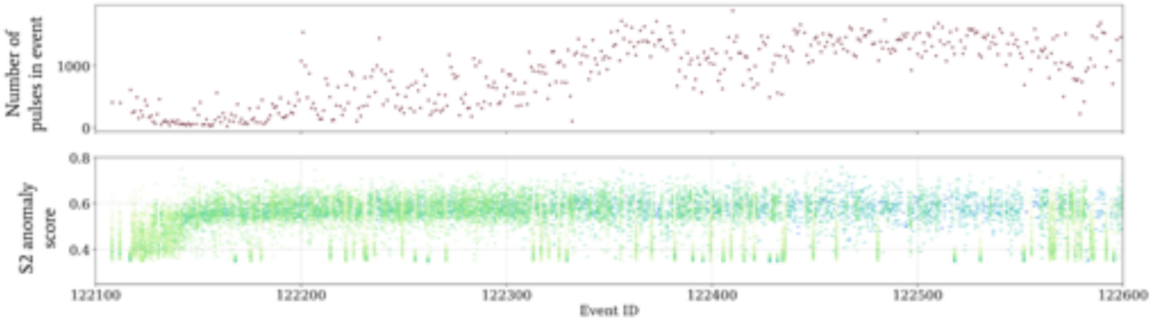


Figure 5.23: Number of pulses in each event (top) and the IF anomaly score for the S2 pulses in each event (bottom) shown for a roughly 90 s interval of data taken during commissioning. The anomaly scores increase before the number of pulses increase, indicating that the pulses classified as S2s deform as a result of being made up of several individual pulses, which then space out with time and start being classified as individual pulses.

beginning at the onset of a high rate period. The anomaly score indicates that the pulse shape starts deforming before the rise in pulse rate.

The anomaly score marked the high rate period before the number of pulses started to increase, because in the intense onset of these emissions, multiple small pulses were classified as single pulses due to their proximity to each other. This was facilitated by the IF being sensitive to pulse shape parameters. Examples of waveforms before and after the number of pulses were seen to increase are shown in Figures 5.24 and 5.25, respectively.

These emissions were later identified as originating from the electric field grids. Figure 5.26

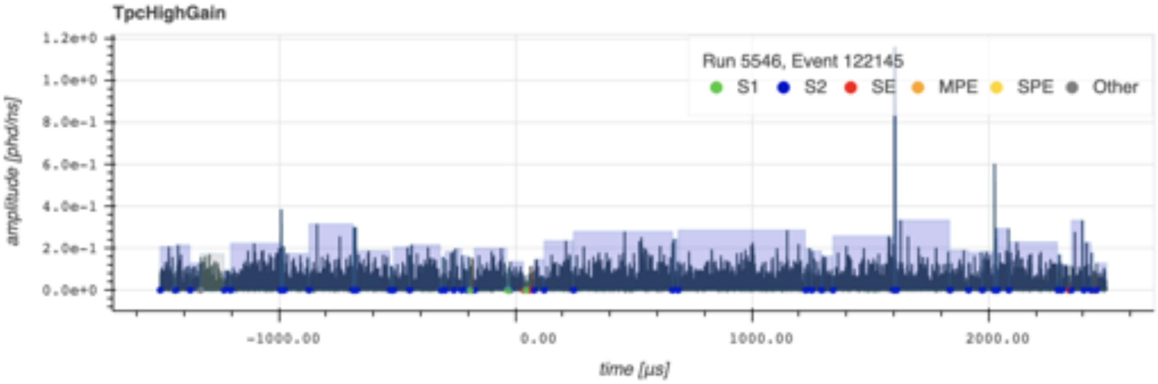


Figure 5.24: Example TPC waveform before the number of pulses increases as in Figure 5.23. Around 100 pulses are classified here, but the true number is far greater. Several single and multiple electron pulses are considered to be a single S2 pulse by the LZ pulse processing algorithms. These S2 pulses are considered highly anomalous by the IF.

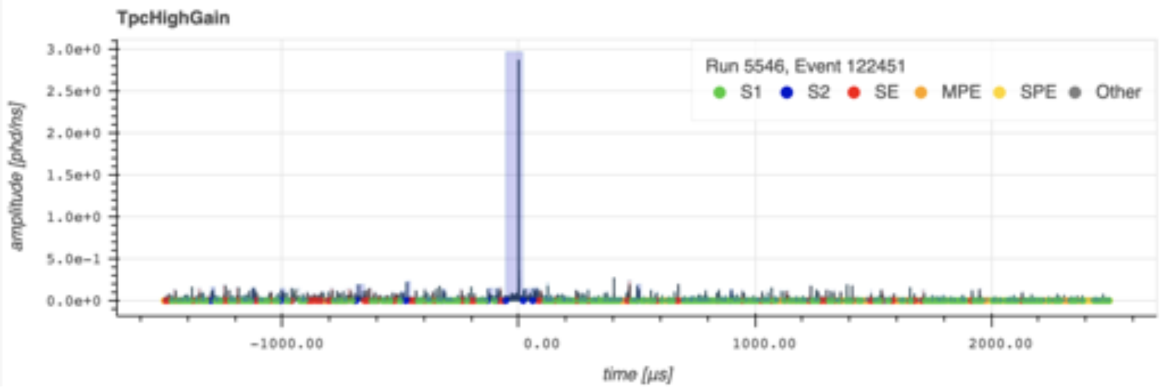


Figure 5.25: Example waveform after the number of pulses starts to increase. Around 1,000 pulses are classified here. The rate of single and multiple electrons has fallen, and thus more of the corresponding pulses are identified individually by the pulse finding algorithm.

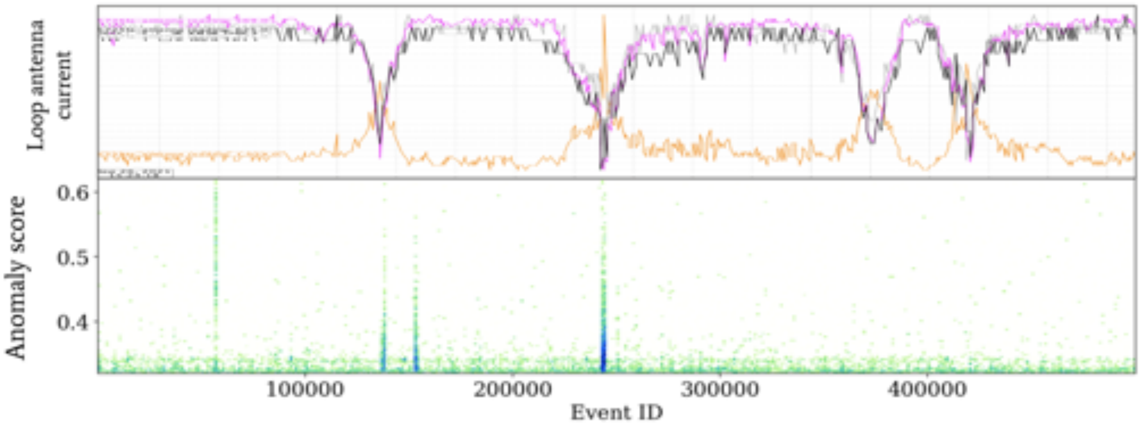


Figure 5.26: Loop antenna current (top), coincident with periods of anomalous S2 pulses (bottom), over a 5 hour duration. Only two periods of coincidence are observed.

shows some coincidence between the anomaly scores and the activity recorded on the loop antennas, which are instruments around the TPC that are sensitive to changes in the electromagnetic environment. There are periods with anomalous pulses that are not coincident with any quantity measured by LZ slow control, and conversely, there is sometimes electromagnetic activity without a coincident pulse anomaly. While further study of these coincidences is needed, the IF provided supporting evidence to the claim that the periods of high rate were caused by electron emissions from the grids.

Chapter 6

Conclusion

6.1 High energy nuclear recoil search

This thesis presents the results of a broadband search for galactic WIMPs with the LZ detector, a dual-phase TPC with an active mass of 5.5 tonnes of liquid xenon. The analysis targeted xenon nuclear recoils caused by energy deposits up to 270 keV_{nr}, encompassing the first result of LZ (which went to 50 keV_{nr}) while exceeding the energy windows of all previous experiments searching for nuclear recoils. The models tested in the enlarged energy window comprise of all possible WIMP-nucleon interactions derived from an Effective Field Theory (EFT), starting from the four non-relativistic degrees of freedom (DoFs) relevant at the energy scale of the scatter. These DoFs are the momentum transfer, relative WIMP velocity, WIMP spin, and nucleon spin bilinears, which are the building blocks of 14 unique interactions. Each interaction may be sensitive to the isospin properties of the target nuclei, so we tested for both isoscalar (where the WIMP interacts identically with protons and neutrons) and isovector (where the WIMP couples differently due to protons and neutrons due to isospin differences) scenarios. Further, we generalized the WIMP-nucleon interaction to include inelastic collisions where the WIMP scatters to a heavier, excited final state.

The event rates of many of these non-standard interactions peak at non-zero energies, as opposed to the canonical WIMP whose event rate exponentially falls with energy. The model spaces furnished by the EFT and inelastic dynamics are ideal to interpret the search for high energy nuclear recoils, since they cover the entire extended energy window. Conversely, the event rates accepted by the extended energy window comprise a significant fraction of the total event rates for all interactions tested (60% in the weakest case of $\mathcal{O}_6$ with a 4,000 GeV WIMP mass), whereas the 50 keV_{nr} cutoff of LZ's first WIMP search accepts virtually no events for momentum-suppressed interactions at high mass. The γ-X background classification tool developed by the author (based on simulations by Greg Rischbieter) was required to extend the energy window to 270 keV_{nr}, and demonstrated an excellent background rejection power.

No excess NR events were observed during the first 60 live–day run of LZ. A total of 56 models (14 interactions, 2 isospin states, elastic and inelastic) were tested and world-leading upper limits for their interaction strengths were placed. The upper bounds on interaction strengths were improved by several orders of magnitude for many interactions, and interactions with severe momentum suppression were the most tightly constrained with respect to previous analyses, by virtue of the large energy window of this analysis. The results of this search may be used in conjunction with results from experiments using target nuclei other than xenon to probe interactions that xenon is insensitive to due to its isospin. Further, the individual WIMP-nucleon couplings excluded by this analysis may inform phenomenological models of dark matter, whose interactions can be reduced down to the fourteen effective interactions considered here.

6.2 Anomaly finding

A general-purpose tool was developed and characterized to identify anomalies at all stages of the data processing chain. The performance on this tool was demonstrated on anomalies stemming from detector effects, rare event topologies, and simulation effects. Further, tools from deep learning were used to identify anomalies in the raw waveforms, which may be adapted for data quality monitoring and pulse classification.

6.3 Liquid xenon measurements with the MiX detector

In the following appendices, the author's work with the Michigan Xenon (MiX) detector is described. The MiX detector is a small dual-phase xenon TPC located at the University of Michigan, used for an experiment to measure the response of LXe to electronic and nuclear recoils. At time of writing, it is currently being prepared to obtain data to measure the ionization and excitation energies of LXe. A discussion of the LXe microphysics and a re-analysis of data from another experiment hinting at the result is presented in Appendix A. The author also outlined a procedure to measure the photon and electron yields of LXe in response to ultra-low energy (less than 0.3 $\mathrm{keV_{nr}}$) nuclear recoil events using neutron capture. The resulting paper, published in Phys. Rev. D [60], is reproduced in Appendix B.

6.4 Summary

Dark matter remains an enigma in particle physics and cosmology. The LXe TPC technology is highly viable for the direct detection of WIMP dark matter and has demonstrated leading sensitivity through two generations of experiments. LZ, currently the most sensitive WIMP detector, has not provided any evidence for standard WIMP scenarios where dark matter couples solely to the total charge and spin of atomic nuclei. This null result contributes to the erosion of theoretical priors in favor of the standard WIMP. This dissertation extends the WIMP search in energy and provides leading sensitivity for non-standard WIMP interactions to meet the widening scope of the dark matter community.

Appendices

Appendix A

Split W-Value Framework and Ongoing Measurements with the MiX Detector

The production of VUV photons and ionization electrons in the LXe medium is linked to the nature of the initial energy deposit and the atomic properties of xenon. For instance, detailed studies of LXe show differences in the S1/S2 ratio among different types of ER interactions such as between γ ray and β energy deposits, and in how an impinging particle distributes energy among the inner shell electrons of xenon atoms [84, 134]. A flexible framework to characterize LXe is tested here, which may lead to a more accurate energy reconstruction by the inclusion of recombination effects. Specifically, the simplifying assumption that treats the excitation and ionization thresholds together as the W-value, introduced in Chapter 2, is removed. Further, an experiment is designed and, at time of writing, is being conducted to measure the slight deviations in the regular energy reconstruction scale (Eq. 2.5) incurred by the difference in energy thresholds.

The ongoing experiment uses the Michigan Xenon (MiX) detector, a small LXe TPC at the University of Michigan designed for measurements characterizing the LXe medium [135]. A new measurement of the as-defined W-value is also desirable due to two recent independent results that measured $W = 11.5^{+0.2}_{-0.3}$ eV and $W = 11.5 \pm 0.5$ eV [136, 137]. These results contradict the consensus value of $W = 13.7 \pm 0.2$ eV reported in Ref. [88], adding to the variations seen in earlier measurements [89, 138]. Since the W-value is central to determine the g_1 and g_2 parameters in LXe TPCs employed in the dark matter search, and since there is potential for an improved energy reconstruction scale that accounts for recombination, the MiX measurement attempts to address the discrepancy.

A.1 Recombination effects in energy reconstruction

The common W-value framework for energy reconstruction, introduced in Ref. [88], results in a better energy resolution than either the S1 or S2 channels afford alone, while its independence on the drift field (via recombination) means that it is simple to use. The drawback is that the W-value, defined as the average energy required to create either an exciton or an ion, is not clearly defined on the atomic scale; it is not only agnostic to the different energetics between excitation and ionization, it also does not account for how the initial energy deposit is distributed across the atom, or group of atoms. A complete modeling of LXe microphysics would include the ionization and excitation of each atomic shell, to be used in conjunction with detailed distributions of the energy deposit. Since this is not feasible, here, only the gross excitation and ionization processes are assumed to have separate energy thresholds. The result is an explicit recombination dependence upon the energy reconstruction.

A.1.1 Excitation and ionization thresholds in LXe

Instead of starting with the W-value common to both excitation and ionization processes as in Eq. 2.7, $E = W[N_{\text{ex}} + N_i]$, each process now has its own energy threshold, such that

$$E = W_{\text{ex}} N_{\text{ex}} + W_i N_i, \tag{A.1}$$

where E is the energy deposited, $W_{\rm ex}$ is the excitation work function, and W_i is the ionization work function. The initial quanta are shuffled according to the recombination fraction r, and are written in terms of the final detectable quanta N^r such that

$$N_{\rm ex} = N^r_{\rm ex} - \left(\frac{r}{1-r}\right) N^r_i, \tag{A.2}$$

$$N_i = \frac{N^r_i}{1-r}. \tag{A.3}$$

Relating to observables and gain factors: $\mathrm{S1} = \mathrm{g}_1 N^r_{\rm ex}$ and $\mathrm{S2} = \mathrm{g}_2 N^r_i$, the following recombination dependent energy scale is constructed as

$$E = W_{\rm ex}\frac{\mathrm{S1}}{\mathrm{g}_1} + \left(\frac{W_i - rW_{\rm ex}}{1-r}\right)\frac{\mathrm{S2}}{\mathrm{g}_2}, \tag{A.4}$$

which can be rearranged in the form of the Doke plot, giving

$$\frac{\mathrm{S2}}{E} = -\frac{\mathrm{g}_2}{\mathrm{g}_1}\left(\frac{1-r}{G-r}\right)\frac{\mathrm{S1}}{E} + \frac{\mathrm{g}_2}{W_{\rm ex}}\left(\frac{1-r}{G-r}\right), \tag{A.5}$$

where $G = W_i/W_{\rm ex}$. The effect of $G \neq 1$ should in principle be measurable using data at different drift fields.

Photon production efficiency from recombination

Recombination may not produce scintillation light with perfect efficiency. Quantifying this efficiency with ε_r, we have the recombination process: $N^f_{\rm ex} = N^i_{\rm ex} + r\varepsilon_r N^i_i$, which in the split W-value framework, yields the most general Doke plot equation

$$\frac{\mathrm{S2}}{E} = -\frac{\mathrm{g}_2}{\mathrm{g}_1}\left(\frac{1-r}{G-r\varepsilon_r}\right)\frac{\mathrm{S1}}{E} + \frac{\mathrm{g}_2}{W_{\rm ex}}\left(\frac{1-r}{G-r\varepsilon_r}\right). \tag{A.6}$$

The EXO collaboration, using a single-phase xenon TPC, has measured ε_r to be 97%, rejecting a unity value at 3σ significance [137]. We do not pursue this measurement since we do not have

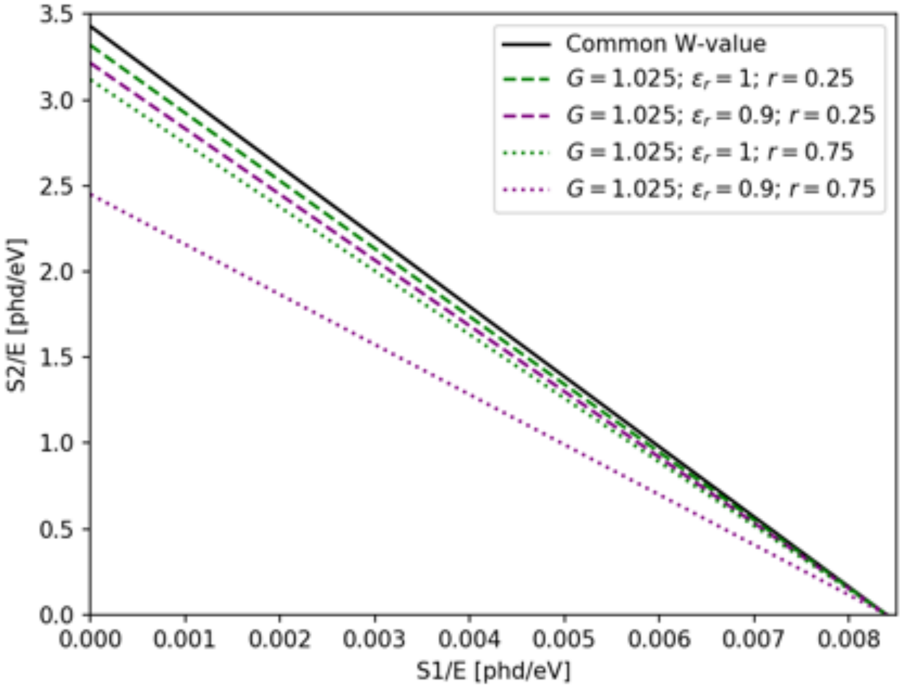


Figure A.1: Doke plot showing the expectations of photopeak trends, within the common W-value framework (black solid), and the split W-value framework with $G = 1.025$. Lines for recombination fraction $r = 0.25$ are shown for a perfect photon production efficiency (green dashed) and for $\varepsilon_r = 0.9$ (purple dashed). Lines for a higher recombination fraction $r = 0.75$, emulating a lower drift field, are shown for a perfect photon production efficiency (green dotted), and for $\varepsilon_r = 0.9$ (purple dotted).

an absolute light detection efficiency calibration in the MiX detector.

A.1.2 Analysis strategy

The goals of this experiment are to make a measurement of the W-value, defined as independent of recombination, in order to the address the discrepancy in published measurements, and to measure the deviation of the ratio $G = W_i/W_{\text{ex}}$ from unity. The W-value measurement procedure for a dual-phase LXe TPC such as the MiX detector has been described in Ref. [136], which is where one of the results with $W = 11.5$ eV was reported. It involves extracting the quantity g_2/W, the intercept of Eq 2.5, which physically corresponds to extrapolating the S1–S2 anticorrelation to the limit of infinite drift field, where S1$=0$. The extrapolation requires an absolute, i.e. independent of W, measurement of g_2.

In dual-phase xenon TPCs, an absolute measurement of g_2 is difficult due to the imperfect electron extraction across the liquid-gas interface; usually experiments assume a W-value and use it to calculate reported values of g_2 [5]. In a single-phase (liquid only) TPC, a calibrated charge pre-amplifier, with a measured efficiency (single-phase equivalent of g_2), may be used to read out the electrons [88, 137]. In a dual-phase TPC, such as Xurich II used in Ref. [136] which measured $W = 11.5$ eV, the extraction efficiency must be assumed to be 100% in the absence of an auxiliary measurement of g_2, implying that the reported W-value is an upper bound. An upper bound lower than the consensus value does not weaken the discrepancy, but making a measurement that is bound only by systematic uncertainties requires a high enough extraction field to justify assuming perfect extraction. On the other hand, it is unclear whether electron extraction is ever 100% even for the highest fields that have been measured; for instance, Ref. [95] defines the efficiency to be 100% for fields above 7.5 kV/cm, where it transitions from a rapid increase to rising with a small slope.

Measuring the deviation of G from unity requires data taken at various drift fields for various sources. A set of monoenergetic γ ray sources that span the 10 keV – 10 MeV range is preferable to sample the range of ionization and scintillation yields. The analysis amounts to determining the variations in the slope of Eq. A.6 at different drift fields. While the MiX detector is being upgraded, which is described in section A.2, public data from the Xurich II detector (Ref. [136]) is re-analyzed to demonstrate $G \neq 1$.

A.1.3 Xurich II re-analysis

Xurich II is a small dual-phase xenon TPC, similar to the MiX detector, which obtained data for ^{37}Ar (2.82 keV decay via electron capture) and $^{83\mathrm{m}}$Kr (9.41 keV and 32.15 keV two step decay) under several drift fields. The ^{37}Ar data were taken at different drift fields than the $^{83\mathrm{m}}$Kr, except

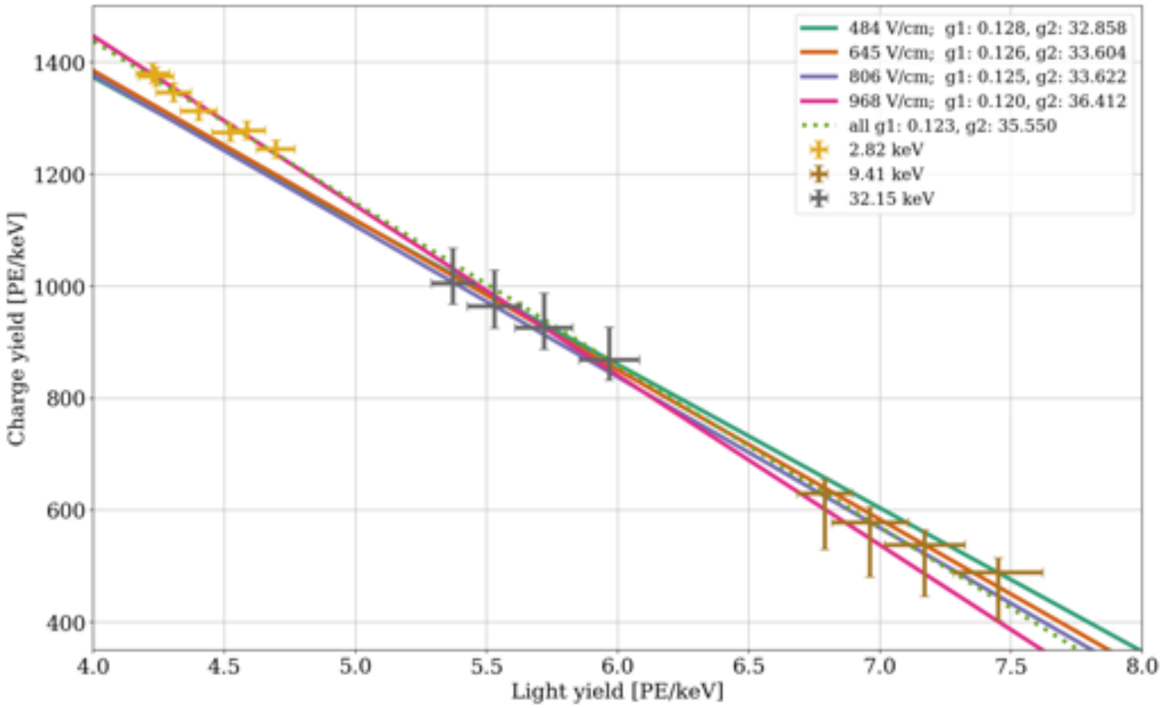


Figure A.2: The light and charge yields of the 2.82 keV ^{37}Ar peak (gold), the 9.41 keV $^{83\text{m}}$Kr peak (brown), and the 32.15 keV $^{83\text{m}}$Kr peak (gray). Extrapolating lines through each pair of $^{83\text{m}}$Kr points at each field are shown, along with the corresponding g_1 and g_2 values calculated assuming $W = 13.7 \pm 0.2$ eV. A line is fit to the three points at 968 V/cm. Further, a linear fit to all the points is shown in green (dotted), yielding a g_1/g_2 ratio that is consistent with the value reported in Ref. [136].

at 968 V/cm, which is the only drift field with all three energy peaks. At least two peaks per field are required to observe the anti-correlation of light and charge yields, so for the remaining fields (484, 645, 806 V/cm) only the two $^{83\text{m}}$Kr peaks are available. The data are shown in Figure A.2, which also shows straight lines through pairs of $^{83\text{m}}$Kr points at each field, in addition to a linear fit to the three energy peaks at 968 V/cm. The observed trend of the slope towards more negative values with increasing drift field (less recombination) is consistent with the predictions from Figure A.1. Table A.1 shows the different g_1 and g_2 values calculated at each electric field assuming $W = 13.7 \pm 0.2$ eV. The deviations observed in these drift-independent quantities may hint at the recombination dependence of the anti-correlation.

The scatter in the g_1 and g_2 values presented above can be recast as a deviation of G from

Drift field [V/cm]	Sources	Slope ($-g_2/g_1$)	Intercept (g_2/W)	g_1 [PE/keV]	g_2 [PE/keV]
All (Ref. [136])	^{37}Ar (7), $^{83\text{m}}$Kr (8)	$-289.5^{+11.3}_{-7.1}$	2596^{+52}_{-34}	0.122 ± 0.005	35.5 ± 0.9
All (fit)	^{37}Ar (7), $^{83\text{m}}$Kr (8)	-289.3 ± 5.4	2594 ± 25	0.123 ± 0.003	35.6 ± 0.3
484	$^{83\text{m}}$Kr (2)	-256	2398	0.128	32.9
645	$^{83\text{m}}$Kr (2)	-267	2453	0.126	33.6
806	$^{83\text{m}}$Kr (2)	-269	2454	0.125	33.6
968	$^{83\text{m}}$Kr (2)	-264	2426	0.126	33.2
968	^{37}Ar (1), $^{83\text{m}}$Kr (2)	-302.8 ± 17.3	2657 ± 78.7	0.120 ± 0.008	36.4 ± 1.1

Table A.1: Values for the slope, intercept, g_1, and g_2 for all data and for each value of the drift field for which more than two sources exist. The g_1 and g_2 parameters were calculated assuming $W = 13.7 \pm 0.2$. Values calculated using only two sources are shown without error bars; the rest are shown with fit errors. For the drift field of 968 V/cm, these values were calculated with and without the inclusion of the ^{37}Ar peak.

unity. In other words, the gain parameters can be held constant at all drift fields, which is physically the case, while the changes in the slope of the Doke plot are attributed to the factor $(1-r)/(G-r)$. To do this, $G = W_i/W_{\text{ex}}$ can be Taylor expanded such that $G \approx 1 + \delta_G$, where the deviation δ_G is positive since $W_i > W_{\text{ex}}$. Then Eq. A.4 becomes

$$\frac{\text{S2}}{E} = -\frac{\text{g}_2}{\text{g}_1}\left(1 - \frac{\delta_G}{1-r}\right)\frac{\text{S1}}{E} + \frac{\text{g}_2}{W_{\text{ex}}}\left(1 - \frac{\delta_G}{1-r}\right), \tag{A.7}$$

where the slope and intercept depend on the drift field via r. A field-independent quantity can be obtained by taking the ratio of the intercept to the slope, yielding $-g_1/W_{\text{ex}}$. The deviations of the slope and the field-independent quantity are shown in Table A.2 for the four drift fields.

From these quantities, the quantity δ_G can be extracted by assuming a recombination model. The recombination fraction r essentially translates the number of excitations and ionizations initially produced at the interaction site to the detected quantities. In the common W-value framework, it may be expressed as

$$r = 1 - \left(\frac{N_{\text{ex}}}{N_i} + 1\right)\left(\frac{\text{S2}}{\text{S2} + \frac{g_2}{g_1}\text{S1}}\right). \tag{A.8}$$

Using this definition of r in the measured $\delta_G/(1-r)$ allows a best fit value for δ_G over the four drift fields to be obtained, shown in Table A.3, indicating that G deviates from unity at the 2–3%

Drift field [V/cm]	Sources	$\delta_G/(1-r)$	$-g_1/W_{\text{ex}}$
All (Ref. [136])	^{37}Ar (7), $^{83\text{m}}$Kr (8)	–	–
All (fit)	^{37}Ar (7), $^{83\text{m}}$Kr (8)	10^{-5}	-9.0 ± 0.2
484	$^{83\text{m}}$Kr (2)	0.114	-8.97
645	$^{83\text{m}}$Kr (2)	0.077	-9.35
806	$^{83\text{m}}$Kr (2)	0.069	-9.18
968	$^{83\text{m}}$Kr (2)	0.086	-9.17
968	^{37}Ar (1), $^{83\text{m}}$Kr (2)	-0.047 ± 0.003	-8.78 ± 0.57

Table A.2: Values for the Doke slope deviation and field-independent slope-to-intercept ratio for each value of the drift field for which more than two sources exist. The slope deviations were calculated assuming $W = 13.7\pm0.2$. Values calculated using only two sources are shown without error bars; the rest are shown with propagated errors. For the drift field of 968 V/cm, these values were calculated with and without the inclusion of the ^{37}Ar peak.

N_{ex}/N_i	δ_G
0.06	0.0253 ± 0.0004
0.2	0.0283 ± 0.0006

Table A.3: Best fit values of δ_G for two values of N_{ex}/N_i presented in Ref. [84].

level. The ratio N_{ex}/N_i has been theoretically estimated to be 0.06, and measured to be as high as 0.2 [84], but δ_G is not very sensitive to this ratio as seen in Table A.3. Note that Eq. A.8 is only valid in the common W-value framework, so the results in Table A.3 are certainly biased; a proper treatment where a reparameterized recombination model is also profiled over to obtain the best fit δ_G is left to the analysis of the MiX detector's data.

Further scrutiny of the W-value requires more than two energy peaks for each drift field, ideally taken at more than four drift fields. In the next section, we describe the MiX detector, which at time of writing is being prepared to make a measurement of the common W-value and of δ_G in the split W-value framework.

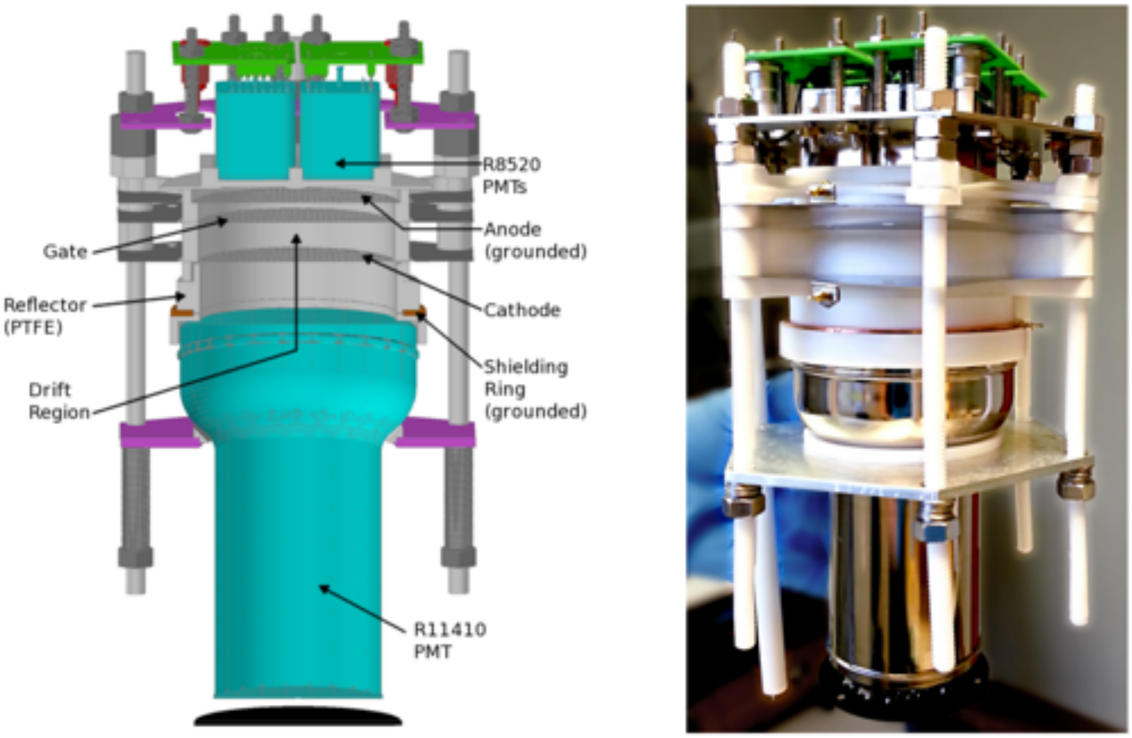


Figure A.3: Left: Schematic of the cross section of the MiX detector. A 3-inch PMT covers the entirety of the TPC cross section on the bottom, and four 1-inch PMTs cover the top. Three hexagonal meshes constitute the cathode, gate and anode grids. Right: Photograph of the assembled detector. Figure taken from Ref. [135].

A.2 The MiX detector

The MiX detector is a dual-phase xenon TPC with an active LXe mass of around 400 g, built by Scott Stephenson, a previous graduate student in our lab. Its small size allows a large fraction of the S1 light to be collected in its 5 PMTs; indeed, the g_1 parameter was originally measured (using $W = 13.7$ eV) to be 0.239 ± 0.012 PE/photon, which is twice that of a modern TPC used in dark matter searches [135]. A further distinction of the MiX detector is its high energy resolution, measured to be $1.03 \pm 0.11\%$ at 1.33 MeV, one of the best achieved in LXe. A full description of the MiX detector and measurements of its signal collection efficiencies and energy resolution is given in Refs. [135] and [139], and a schematic is shown in Figure A.3.

A high extraction field, preferably greater than 7.5 kV/cm in the LXe above the gate, is necessary to measure a W-value that is unaffected by incomplete charge extraction into the GXe [95]. The extraction field, which differs in the LXe and GXe by a factor of the LXe dielectric constant $\varepsilon_l = 1.85$, is set primarily by the gate electrode voltage and the height of LXe above the gate; the anode electrode is grounded. The expression for the LXe extraction field E_l is

$$E_l = \frac{\Delta V}{1.85d - 0.85h}, \tag{A.9}$$

where ΔV is the voltage across the gate and anode, which is just the gate voltage since the anode is grounded, d is the distance between the gate and anode, and h is the height of the LXe above the gate, where $h < d = 7$ mm. For a $E_l = 7.5$ kV/cm, with the liquid level halfway up to the anode, a gate voltage of about -7.5 kV is required, which the MiX electrodes are not designed for. To test drift fields in the range of 100–1,000 V/cm, the cathode has to hold an even larger voltage of around -8.5 kV.

At time of writing these voltages are unrealized, with electrical breakdown occurring just shy of the requirements. An instance of an electrical breakdown in air between the corners of the gate and anode is shown in Figure A.4. The grids and HV cables are currently being insulated with Kapton and PTFE to address this issue.

The MiX detector has been dormant for a few years since its construction, when it was first characterized in Ref. [135]. We have recommissioned the detector and demonstrated that the MiX detector is capable of taking data from multiple ER sources at various drift fields, in addition to improving its safety systems. Preliminary data from several sources are presented here.

A.2.1 Preliminary data

The sources ^{22}Na (511 keV positron annihilation peak), ^{137}Cs (662 keV γ ray), and ^{133}Ba (most prominently with a 365 keV γ ray) were used to take data with the MiX detector. These sources are available in the form of small pellets that can be attached to the side of the outer cryostat. Figure A.5 shows the (x, y) position reconstruction for a ^{22}Na source, using a center of gravity (CoG)

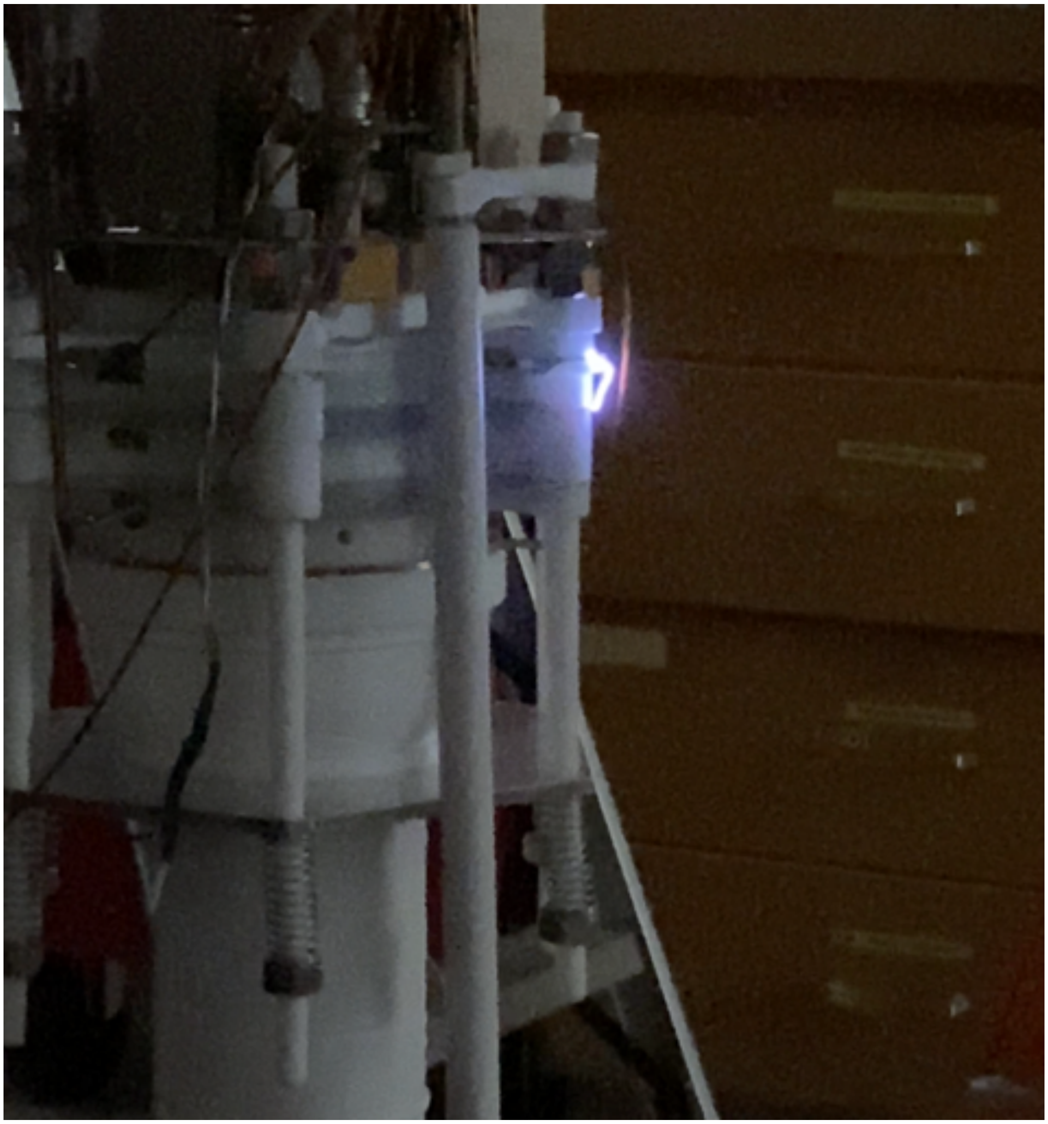


Figure A.4: Electrical breakdown in air between the negatively biased gate (-9.5 kV) and the grounded anode. The breakdown voltage in air is higher than that in GXe by a factor of 3 for a range of separation distances. Photograph by Samara Steinfeld.

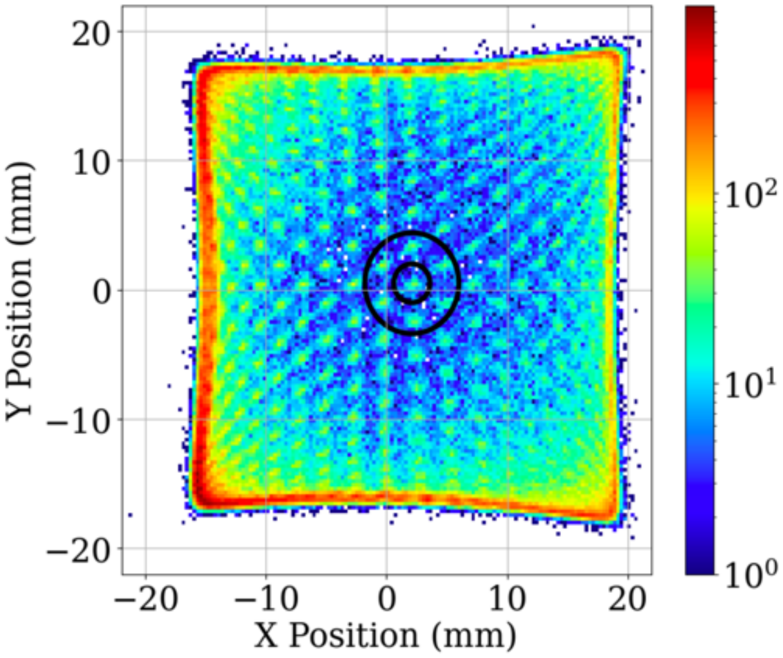


Figure A.5: Position reconstruction of events from an external a γ source using a center of gravity method. The concentric circles enclose the fiducial 'donut', chosen for an optimal energy resolution. The accuracy of the reconstruction is best at the center and degrades towards the walls, evidenced by the warping of the high density dots, which represent the holes in the hexagonal gate electrode.

method based on the amount of light observed by the top four PMTs, described in Ref. [135]. The high density spots observed in the (x, y) plane correspond to the holes in the hexagonal gate grid which the drifting electrons funnel through. The reconstruction asymptotes to a square in the large-radius limit due to the CoG method estimating the x and y positions separately using pairs of PMTs; an improved method is presented in Ref. [134], but CoG is deemed sufficient especially with the tight fiducialization planned for the MiX analysis.

S1 and S2 pulses are selected using cuts on the pulse shapes, including width and area, and a central fiducial volume is selected to be a donut bounded by concentric circles of radii 5 mm and 2 mm as shown in Figure A.5. A plot of events produced by exposure to 511 keV ^{22}Na γ rays are shown in bS1–tS2 space in Figure A.6, where bS1 corresponds to the S1 light observed by the bottom PMT and tS2 corresponds to the S2 light observed by the four top PMTs. Partial energy deposits due to Compton scatters are observed on the bS1 $\propto$ tS2 trendline, along with a prominent

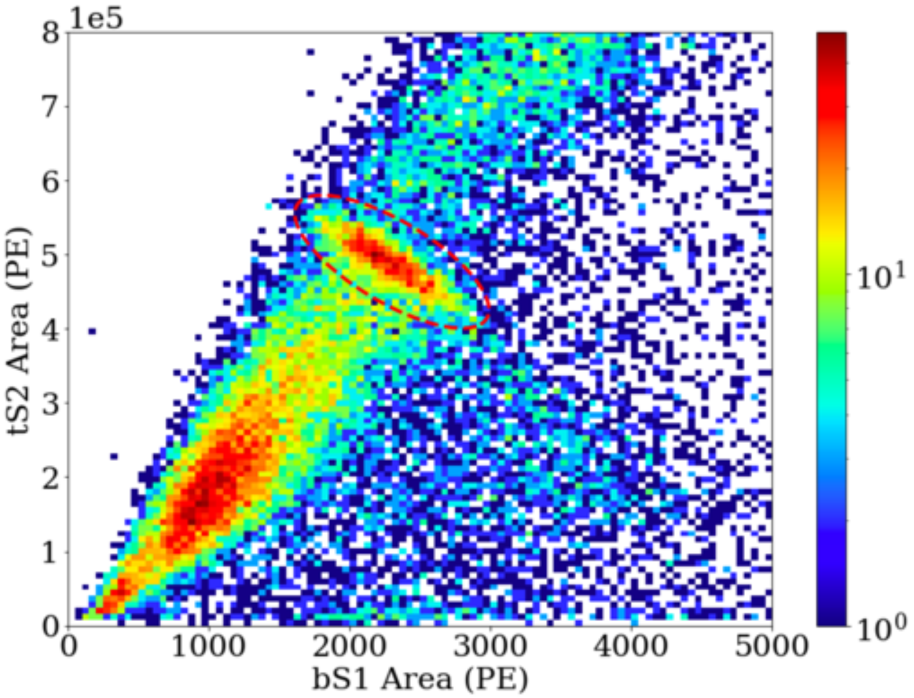


Figure A.6: ER events produced by a ^{22}Na in tS2–bS1 space, showing the spectrum of energy deposits from Compton scatters (below bS1$= 2,000$ PE), and the photoabsorption peak (red dashed) of the 511 keV annihilation γ ray at bS1$= 2,300$ PE. Higher energy events are due to the Compton scatters of the 1274.5 keV γ ray from the nuclear de-excitation of ^{22}Ne produced in the ^{22}Na decay. Events under the indicated ER populations are likely due to misreconstructions and accidental pairings of S1 and S2 pulses, and can be targeted with specialized cuts.

photoabsorption peak. The bS1–tS2 anticorrelation at the photoabsorption peak is primarily due to recombination fluctuations. The anticorrelation in a single peak can be used to determine the W-value if the g_1 and g_2 are absolutely known, although the anticorrelation observed in a collection of sources with different energies can be constrained to a higher precision.

Figure A.7 shows the photoabsorption peaks of three sources, fitted to 2D Gaussian functions in bS1–tS2 space. The anticorrelation is extrapolated to calculate the signal collection efficiencies (shown in the plots), assuming the consensus W-value of 13.7 ± 0.2 eV, yielding g_1 corresponding only to the bottom PMT, and g_2 corresponding only to the top PMTs. Typically, the gain parameters are obtained from a Doke plot with multiple sources, instead of using the anticorrelation within each peak as done here.

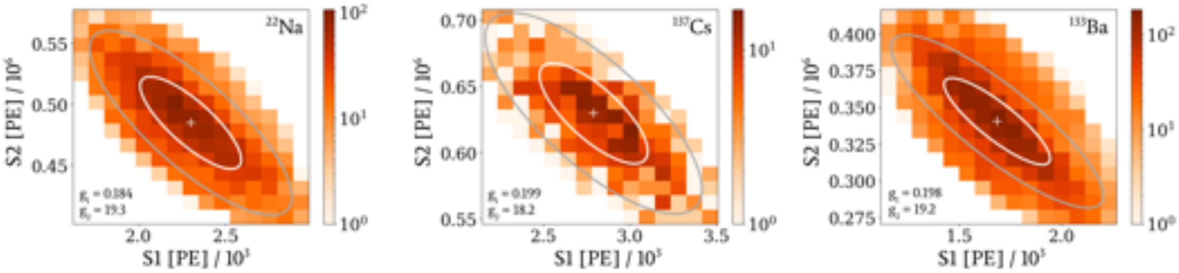


Figure A.7: Photoabsorption peak fit to 2D Gaussian functions, for ^{22}Na (left), ^{137}Cs (middle), and the 365 keV peak of ^{133}Ba (right) in the space of bS1 and tS2. The white cross is the center of the Gaussian distributions, while the 1σ and 2σ contours are shown in white and gray, respectively. The signal collection efficiencies (inset text) are calculated using the anticorrelation for each peak.

Work to maximize the extraction field is currently being carried out in our group, which includes adding electrical insulation and rebuilding cables and connections using parts rated for HV. Simultaneously, analyses are being done to characterize the position dependence of the light and charge collection efficiency, calibrate the PMT response to a single VUV photon and ensure the response is linear with PMT voltage, measure any time dependence of detected signals, and investigate sources of small electron emission.

Appendix B

Ultra-Low Energy Calibration Study Using Neutron Capture with the MiX Detector

Light WIMPs (see Figure 1.10) and the Coherent Elastic ν–Nucleus Scattering (CEνNS) of solar neutrinos may have significant event rates below the current energy thresholds of experiments sensitive to NRs. For instance, neutrinos from ^{8}B decay at the end of the solar *pp* chain may interact via CEνNS and produce up to 7 events above LZ's baseline energy threshold over the full 5,600 day exposure [50]. Lowering the energy threshold by calibrating the scintillation and ionization yield of LXe to sub-keV NR energy deposits will therefore improve the chances of a first measurement of the ^{8}B CEνNS process, in addition to raising LZ's sensitivity to light WIMPs. Figure B.1 shows the current state of measurements with three theoretical models representing the variety of predictions at low energies. This chapter contains an as-published paper (Ref. [60]) written by the author addressing the challenge of using neutron capture as a source of ultra-low energy deposits below 0.3 keV$_{\text{nr}}$ for calibrations in LXe.

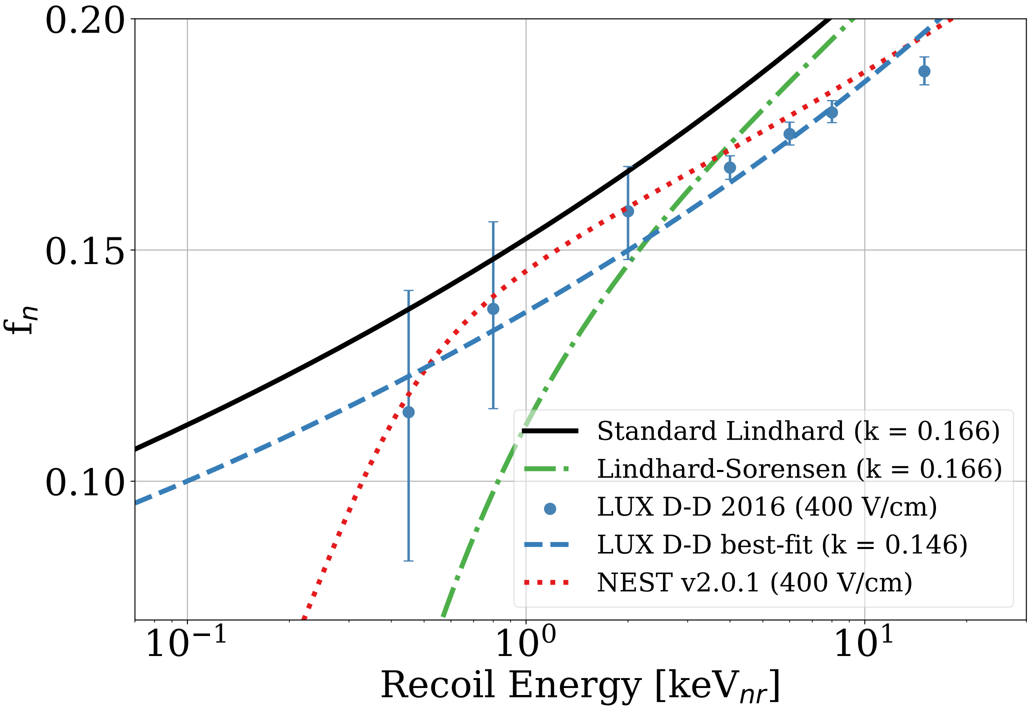


Figure B.1: Energy fraction f_n expended in producing photons and electrons (i.e. not lost to heat) as a function of NR energy. The lowest energy measurement (blue points) are from a neutron scattering analysis by the LUX collaboration [58]. The solid black line is the prediction from the Lindhard model of stopping power, and the dashed blue line depicts the same Lindhard model with a free parameter k fit to the LUX data [102]. The dotted-dashed green shows the Lindhard model modified to include an energy threshold for quanta production [140].

B.1 Abstract

The feasibility of an ultra-low energy nuclear-recoil measurement in liquid xenon using neutron capture is investigated for a small (sub-kilogram) liquid xenon detector that is optimized for a high scintillation gain, and a pulsed neutron source. The measurement uses the recoil energies imparted to xenon nuclei during the de-excitation process following neutron capture, where promptly emitted γ cascades can provide the nuclei with up to $0.3\,\mathrm{keV_{nr}}$ of recoil energy due to conservation of momentum. A successful calibration of scintillation photon and ionization electron yields below this energy will contribute to a greater sensitivity for liquid xenon experiments in searches for light WIMPs.

B.2 Introduction

Underground liquid xenon (LXe) time projection chambers (TPCs) have played an important role in constraining the parameter space available to dark matter in the form of Weakly Interacting Massive Particles (WIMPs) passing through Earth [141]. However, light ($< 10\,\mathrm{GeV}$) WIMPs are kinematically ill-matched with xenon ($A \approx 131$) and deposit less energy in the medium than their heavier counterparts. As a result, dark matter experiments that use LXe suffer a drastic drop in sensitivity for light WIMPs, where the expected signals approach the energy thresholds of the detectors [105]. Hints of light dark matter in several experiments that use other detector media, like CRESST-II [142], CDMS-II-Si [143], and CoGeNT [144], have therefore stoked interest in characterizing the response of LXe to sub-keV energy depositions.

Matter and radiation deposit energy in LXe by interacting with either atomic electrons, creating electronic recoil (ER) events, or with nuclei, creating nuclear recoil (NR) events [145]. WIMPs are predicted to scatter off nuclei, leaving behind NR signatures [31, 146]. Both ER and NR events create detectable scintillation photons (S1) and ionized electrons, with some energy being lost as heat [141]. In a dual-phase LXe TPC[1], the ionized electrons are drifted towards and extracted into a gaseous xenon space by an electric field, where a secondary larger flash of light (S2) is produced by electroluminescence. The ratio S2/S1 is smaller for an NR than for an ER, a feature of LXe that allows ER events to be rejected with high efficiency ($> 99\%$ at 50% NR acceptance) from potential WIMP-induced NR events [84, 106, 147].

For a particular experiment to infer the WIMP mass and interaction cross section in case of a discovery, a map from the space of observed $\{S1, S2\}$ signals to NR energy is required. The production of S1 and S2 signals in LXe due to NR events of known energies has been characterized in a series of measurements over the past two decades [148–152]. As a result, a detector-independent picture of how LXe produces photons and electrons in response to NR events has emerged. In recent years there has been a concerted effort to determine these quanta yields at lower energies, allowing experiments to be sensitive to lighter WIMPs [59, 153–156]. The current lowest energy measurements have found 1.1-1.3 ionized electrons per 0.3 $\mathrm{keV_{nr}}$ recoil [156, 157], and 1.3 scintillation photons per 0.45 $\mathrm{keV_{nr}}$ recoil [156]. This work presents an experimental concept to measure these yields below 0.3 $\mathrm{keV_{nr}}$.

Previous measurements of the photon and electron yields in LXe have used the elastic scattering of neutrons as a source of nuclear recoils. We propose to use xenon nuclei that have captured neutrons. The nuclear recoils of interest are generated by the asymmetric emission of de-excitation γ cascades that leave the TPC undetected, as suggested in Ref. [58]. The idea of using neutron capture to access low recoil energies was implemented for germanium in Ref. [158], and has been repeatedly studied in that material [159–161]. Here we introduce a technique to implement this idea in LXe.

The details of this work correspond to simulations carried out for the Michigan Xenon (MiX) detector, a small (400g active volume) dual-phase LXe TPC with an excellent light collection efficiency and energy resolution [135], although the principles apply to any small TPC. A pulsed neutron source and a neutron moderator surrounding the detector are assumed for the simula-

---

[1] The basic operating principle of a typical dual-phase LXe TPC is described in Section III. E of Ref. [141].

tion. These components are found to be crucial in creating a collection of neutron captures in each pulse that are unaccompanied in time by other sources of NR, in addition to reducing backgrounds from spurious electron emissions.

The paper is organized as follows. Section B.3 describes how the neutron capture induced nuclear recoils are selected, while in section B.4 we present details about the Monte Carlo simulation. In section B.5, we report how to optimize the neutron capture signal by varying aspects of the experimental setup. Background and pile-up events are discussed in section B.6, along with changes to the setup required to minimize them. Section B.7 describes the implications this measurement could have on the sensitivities for light WIMP searches. We conclude in section B.8.

B.3 General approach

After a xenon nucleus captures a neutron, the γ cascade leaves it with up to $0.3\,\mathrm{keV_{nr}}$ of kinetic energy that it dissipates among neighboring atoms, producing photons and electrons. In order for the associated S1 and S2 signals to be cleanly recorded by a data acquisition system, the acquisition window cannot be contaminated by other ER or NR events. Only acquisition windows free of ER events are chosen for the measurement, by selecting captures in which the de-excitation γ cascade escapes the active volume, and also by rejecting events with ER events that originate externally. A neutron capture event can be positively identified if a separate detector outside the TPC detects the γ cascade, providing a timestamp to tag the capture NR.

Using a pulsed neutron source and a thin moderator between the source and detector, a set of neutron capture events suitable for the measurement can be produced in each pulse. Since the neutron capture cross section is roughly proportional to the inverse speed of the incident neutron (except at resonances), capture events are mostly caused by slow neutrons in the TPC. The role of the moderator is to slow down monoenergetic neutrons from the source, while discouraging neutron capture in the moderator itself, as the resulting γ rays are a source of pile-up. Accordingly, the simulation shows that partial neutron moderation is ideal. The thin moderator allows fast neutrons into the TPC first, which are likely to scatter, followed by slower neutrons that are captured. In this arrangement NR events due to neutron capture can be isolated from scattering events with an appropriate time cut.

The observed S1 and S2 pulses have to be associated with the energy of the nuclear recoil that produced them. While many previous measurements have had precise knowledge of the recoil energies, for example by using the angle of the scattered neutrons [59, 154], this measurement relies on a model of energy deposition in LXe due to the neutron capture process. The distribution of NR energies simulated by this model will be used to calculate the sizes of the S1 and S2 signals according to parameterized estimates of the yields below 0.3 $\mathrm{keV_{nr}}$. These parameters can be adjusted to fit the calculated S1 and S2 sizes to the observed data, as performed in Ref. [58]. The energy deposition model and its uncertainty are presented in Section B.4 and discussed in detail in Appendix B.9.

B.3.1 Neutron interactions in liquid xenon

Upon capturing a neutron, most xenon isotopes promptly de-excite (within 1 ns) to their ground state by releasing a cascade of γ rays: ${}^{A}\mathrm{Xe} + \mathrm{n} \rightarrow {}^{A+1}\mathrm{Xe} + \sum\gamma$ [162]. In some cases, this process also releases internal conversion electrons. If the de-excitation transition is direct, that is if a single γ ray carries away all the excitation energy (or equivalently, if several γ rays are emitted in the same direction), a nucleus initially at rest is given the maximum recoil energy

$$E_{R,\mathrm{max}} = \frac{S_n^2}{2M_{\mathrm{Xe}}} \approx S_n^2 \left(\frac{4 \times 10^{-6}}{\mathrm{MeV}} \right), \tag{B.1}$$

where S_n is the neutron separation energy of the newly created xenon isotope ${}^{A+1}\mathrm{Xe}$, and M_{Xe} is its mass. Table B.1 shows neutron separation energies and the corresponding maximum recoil energies for each of the naturally occurring xenon isotopes.

Most de-excitations occur with the emission of several γ rays that exit the nucleus in different directions, leaving the nucleus with recoil energy $E_R \leq E_{R,\mathrm{max}}$. As a result, the recoil spectra

Target Isotope	Abundance (%)	E_1^* (keV)	Capture Cross Section (b)	Product Isotope	S_n (keV)	$E_{\mathrm{R,\,max}}$ (keV$_{\mathrm{nr}}$)
^{124}Xe	0.1	354.0	165 ± 20	^{125}Xe	7603	0.230
^{126}Xe	0.1	388.6	3.8 ± 0.5	^{127}Xe	7223	-
^{128}Xe	1.9	422.9	5.2 ± 1.3	^{129}Xe	6908	0.187
^{129}Xe	26.4	39.6	21 ± 5	^{130}Xe	9256	0.332
^{130}Xe	4.1	536.1	4.8 ± 1.2	^{131}Xe	6605	0.168
^{131}Xe	21.2	80.2	85 ± 10	^{132}Xe	8937	0.305
^{132}Xe	26.9	667.7	0.42 ± 0.05	^{133}Xe	6440	-
^{134}Xe	10.4	847.0	0.27 ± 0.02	^{135}Xe	8548	-
^{136}Xe	8.9	1313.0	0.26 ± 0.02	^{137}Xe	4025	0.060

Table B.1: Properties of xenon nuclei that are relevant to interactions with slow neutrons: natural abundances [163], energies of the first excited nuclear state E_1^* [164], thermal neutron capture cross sections, neutron separation energies S_n of the product nuclei, and the maximum recoil energy $E_{\mathrm{R,\,max}}$ imparted to the product nuclei by the γ cascades following capture [163]. Of primary interest to the proposed measurement are ^{129}Xe and ^{131}Xe due to their large natural abundances, large thermal neutron capture cross sections, and the prompt γ cascades of their capture products. The isotopes with missing data in the last column produce activated products upon neutron capture that do not decay promptly.

of each isotope will be a distribution bounded from above by $E_{R,\max}$, assuming the momentum transferred to the nuclei from the collision with the neutrons is negligible. Otherwise one has to add to this bound the energy transferred to the nucleus from the collision of approximately $E_{\mathrm{n}}/131$, where E_{n} is the kinetic energy of the neutron when it was captured. The NR events selected for this measurement are produced from the capture of neutrons with an average energy of 20 eV, which results in a negligible 0.15 eV$_{\mathrm{nr}}$ contribution to the recoil energy. In contrast to studies performed with germanium detectors, where monoenergetic recoils of 0.245 keV$_{\mathrm{nr}}$ were tagged using a γ ray from a low energy excited state of ^{73}Ge [159], the entire distribution of capture-induced recoils in xenon will be used.

Metastable states with lifetimes many orders of magnitude greater than the capture states can be populated by neutron capture or by the inelastic scattering of neutrons by xenon. The most abundantly created metastable states are $^{129\mathrm{m}}$Xe and $^{131\mathrm{m}}$Xe, which produce prominent 236 keV and 164 keV γ lines, respectively [165]. These γ rays also recoil xenon nuclei, but the resulting events do not contribute to the NR calibration for two reasons. Most importantly, the magnitude of nuclear recoils caused by the emission of these γ rays is $O(0.1\,\mathrm{eV_{nr}})$, and will not be sufficient to produce quanta. Second, the half lives of the metastable states are too long (seconds to days) for them to be selected in time along with the neutron captures in each pulse.

Elastic scattering is an inefficient process to transfer energy from slow neutrons to xenon nuclei due to the large difference in their masses and because no energy goes into altering nuclear states [73]. Due to this inefficiency, a neutron has to scatter numerous times before it is captured, resulting in a high rate of elastic scatters immediately after the neutron pulse. The simulations show that the thermalization time of neutrons in LXe is $O(10\,\mu\text{s})$, after which they are readily captured. The average time it takes for a neutron to be captured in the TPC after being emitted is $O(100\,\mu\text{s})$. This is long enough for the time cut to be effective in isolating a collection of NR events produced only by captures, with an acceptance of around 80% for the experimental configuration discussed in Section B.4.2.

B.3.2 Signal selection

Although all capture events result in a recoiling nucleus, the signal events are defined to be neutron captures that did not deposit more than 10 eV_{nr} before capture in the TPC, and where the entire γ cascade escapes the TPC without depositing energy in it. If internal conversion electrons and subsequent atomic emissions (X-rays and Auger electrons) are produced in an event, it is discarded. This ensures that signal events have a pure NR signature. An example of an NR spectrum due to the neutron capture de-excitation process, and the subset of signal events, is shown in Fig. B.2 for a detector geometry that is discussed in Section B.4. Also shown are the low energy nuclear recoil events due to elastic and inelastic neutron scattering. At energies below $0.3\,\text{keV}_{\text{nr}}$, neutron capture events contribute to the majority of the NR spectrum, and signal events make up around 15% of captures in the TPC. The capture-induced NR events above around $0.3\,\text{keV}_{\text{nr}}$ are due to collisions with faster moving neutrons, and can be removed with a time cut.

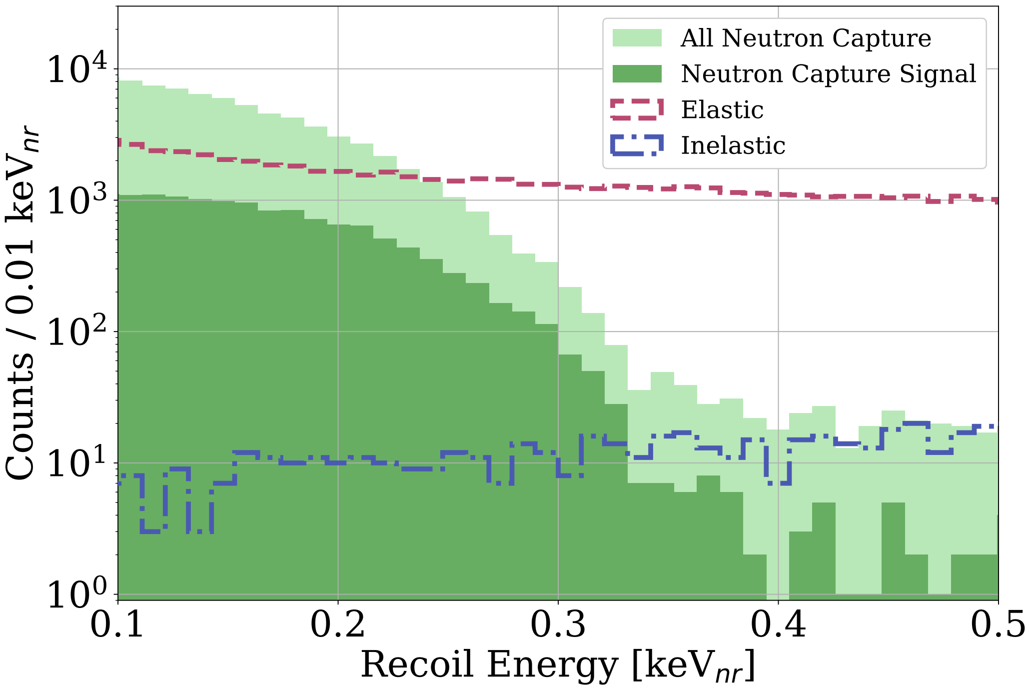


Figure B.2: Nuclear recoil spectrum due to neutron interactions simulated in the MiX detector. The shaded light green histogram (140,000 counts) shows all recoil events due to neutron captures, while the shaded dark green portion (20,000 counts) only retains those where all of the γ-rays from the nuclear de-excitation process escape the active volume. The concentration of signals below about 0.3 keV$_{\mathrm{nr}}$ provides an opportunity for a measurement of quanta in this energy region. Also shown are the recoil events due to elastic (dashed magenta) and inelastic (dashed-dot blue) neutron scatters. All inelastic scatters are shown, regardless of whether their de-excitation γ-products escape the TPC.

B.3.3 Tagging signal events using the LXe skin

Signal events can be positively identified if their γ cascades are detected outside the TPC. A natural location to detect the interactions of these γ rays is the detector skin, the volume of LXe immediately outside the TPC. In the MiX detector, this volume is ideal for tagging signal events due to its large size. Simulations show that 70% of signal events can be tagged using an instrumented skin with a $100\,\mathrm{keV_{ee}}$ energy threshold. In other words, less than 30% of signal events emit γ cascades that escape not only the TPC, but the surrounding skin region as well. No significant bias on the NR energy spectrum is observed when taggable signals are selected. Tagging offers a major reduction in the single-electron background commonly observed in LXe TPCs that may otherwise dominate the number of events from neutron capture that also produce single electrons [99].

B.4 Simulation

The NR energy distribution of neutron capture events depends on the detector's neutron environment and the nuclear properties of xenon. The passage of neutrons emitted from an external source through the MiX detector, and the energy deposits of neutron capture events were studied using a Monte Carlo simulation built with the GEANT4-based application BACCARAT, a detector independent framework developed by the LUX and LZ collaborations [166, 167]. The MiX detector geometry was tessellated from existing CAD drawings and imported into this framework using the CADMesh package [168].

B.4.1 Neutron model

The low energy neutron transport processes are modeled using the QGSP_BIC_HP physics list in GEANT4, and the de-excitation process following neutron capture is simulated with the

GEANT4 photon evaporation model. The photon evaporation model simulates discrete and continuous γ cascades using the Evaluated Nuclear Structure Data File (ENSDF), and also simulates internal conversion electrons [169]. The photon evaporation algorithm conserves energy and momentum, and appears to handle the dynamics of cascade production sufficiently well, although it has not been experimentally validated. Validation requires measurements of the γ spectra for each multiplicity[2] that have so far only been made for the target isotope ^{136}Xe [170]. Since the experimental concept relies on a comparison with simulations, a custom algorithm was implemented to generate nuclear recoils from neutron capture and used to calculate the uncertainty of the NR energy spectrum. This uncertainty, shown in Fig. B.3, incorporates discrepancies in the γ spectra between ENSDF and the evaluated gamma ray activation file (EGAF), which is an experimental database of multiplicity-independent neutron capture γ energies [171]. The uncertainty calculation is discussed in Appendix B.9.

^{136}Xe is not important to the proposed measurement due its low neutron capture cross section, comprising only 0.1% of neutron capture events in natural xenon, and because its largest recoil energy is $60\,\mathrm{eV_{nr}}$, which is too small to produce a signal. However it is the only isotope for which data exists to make a comparison with GEANT4 that properly takes into account γ spectra at each multiplicity. Using the custom algorithm that generates nuclear recoil events from neutron capture, the prediction of the NR spectrum from those data was compared with GEANT4 and a weighted average difference of 39.7% was found in the $0-60$ $\mathrm{eV_{nr}}$ range. In Appendix B.9 this calculation is presented and it is argued why such large discrepancies are not expected for the other isotopes of xenon if measurements of their γ spectra are eventually made.

B.4.2 Description of the setup

The MiX detector is a small dual-phase TPC at the University of Michigan that is ideally suited to study properties of LXe. A cross section of the detector is shown in Fig. B.4. The MiX

---

[2] Multiplicity refers to the number of γ rays emitted in a de-excitation.

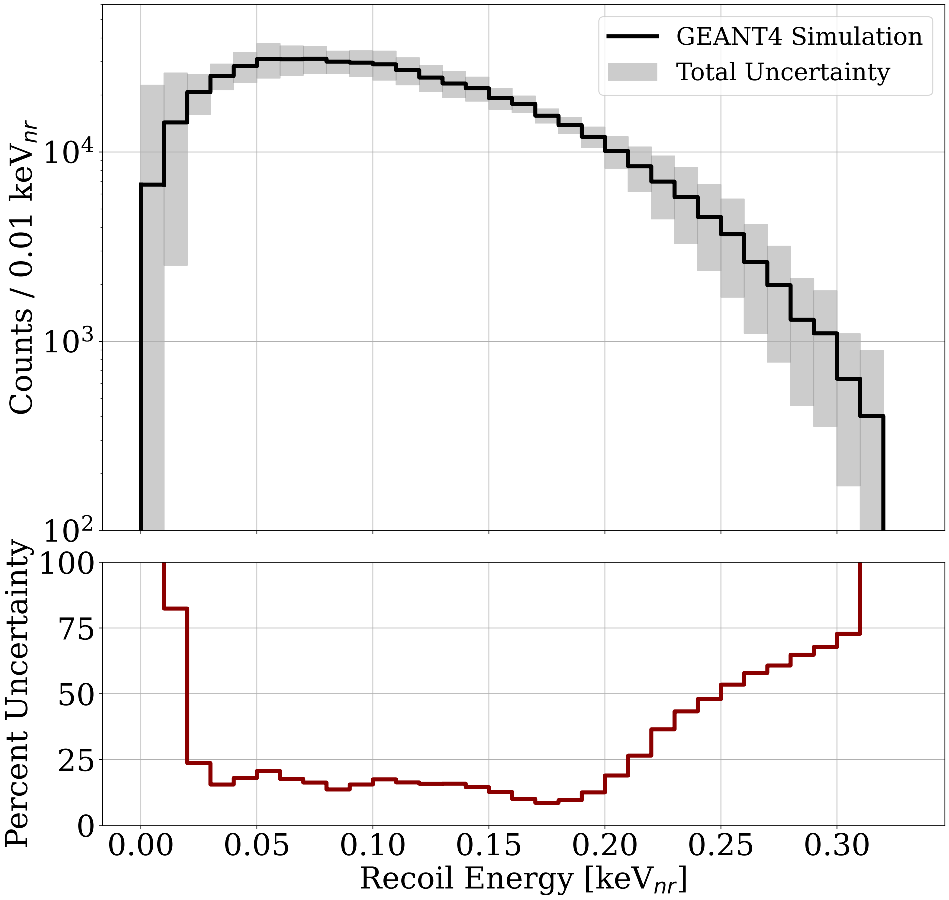


Figure B.3: *Top:* NR spectrum due to thermal neutron capture in LXe simulated using GEANT4. The gray uncertainty band represents the total uncertainty, which incorporates discrepancies in the γ spectra between the ENSDF and the EGAF files. *Bottom:* The error band in the top panel is presented as percent uncertainty for clarity.

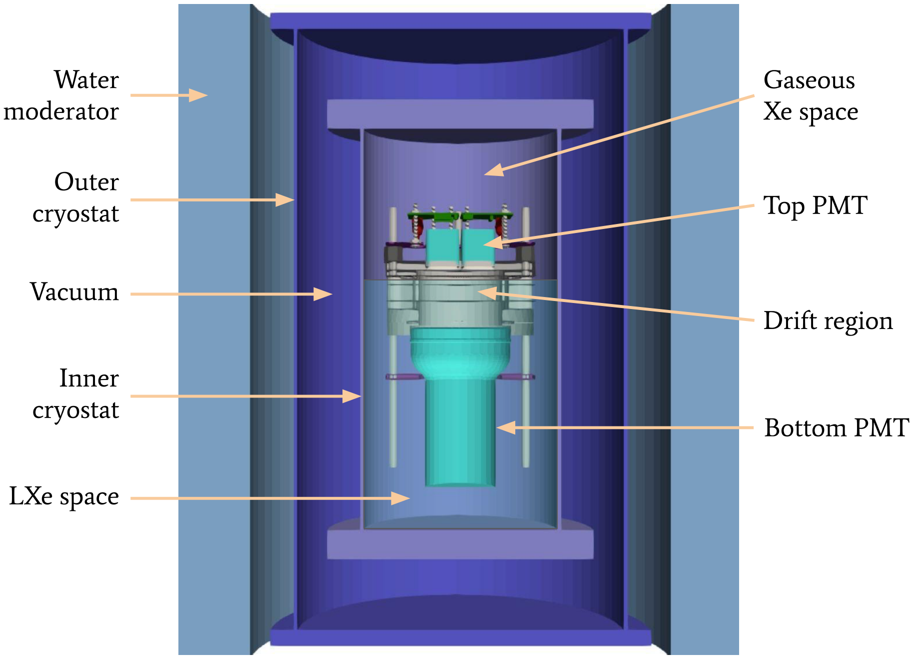


Figure B.4: 3D model of the MiX detector. The inner cryostat encloses the LXe space that partially submerges the TPC assembly, and thus the TPC contains only a small fraction of the LXe in the system. The thickness of the water tank shown here is 5 cm.

detector has a drift chamber with a diameter of 62.5 mm and a height of 12 mm. It was designed and built to have good signal gains, with scintillation and ionization gains of (0.239 ± 0.012) photoelectrons/photon and (16.1 ± 0.6) photoelectrons/electron, respectively [135]. The high scintillation gain, which is crucial to measure the LXe response to low energy interactions [141], is more than a factor of 2 larger than that of typical $O(100\,\text{kg})$ scale detectors. This makes the MiX detector a suitable candidate to perform an ultra-low energy NR calibration in LXe.

The feasibility study assumes a 2.45 MeV monoenergetic neutron source, modeled after an upgraded Adelphi Technologies DD109 Deuterium-Deuterium (D-D) neutron generator. The source has the ability to create pulses as short as 20 μs at an instantaneous rate of 10^9 n/s [58]. A point source of neutrons that originates one meter[3] away from the center of the TPC is simulated. The solid angle of the neutrons that intercept the setup ranges from 0.1 to 0.35 steradians, depending on the size of the water tank. The ability to produce short pulses of neutrons is essential

---

[3] The conclusions of this study do not strongly depend on this distance.

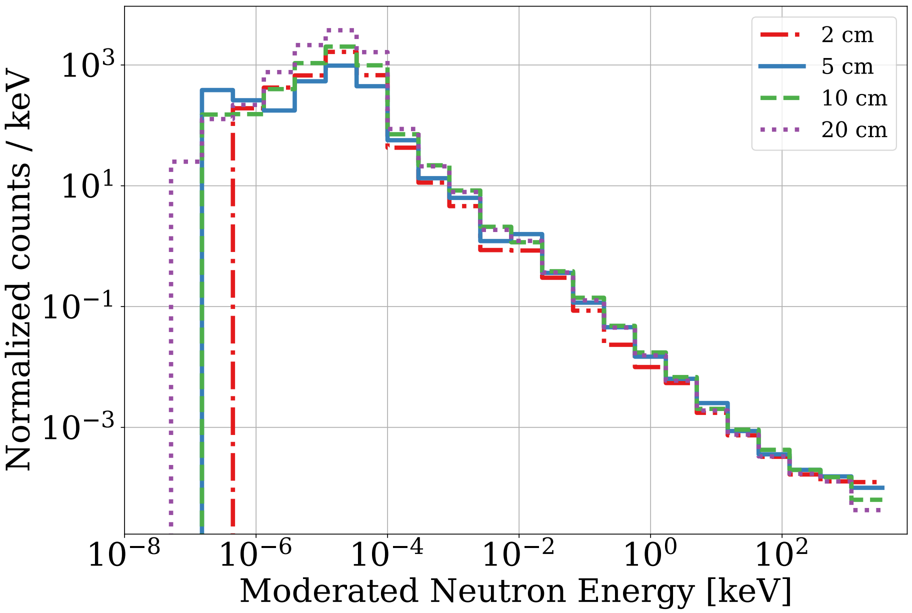


Figure B.5: Kinetic energy distributions of neutrons as they enter the TPC after being moderated by the water tank, shown for various thicknesses of the tank.

to isolate neutron capture events and mitigate single-electron background events, as discussed in Sections B.5 and B.6, respectively. It is required that neutron interactions following a pulse completely die off before the next pulse starts. This ensures that the timing effects in each cycle can be treated independently.

A cylindrical water tank surrounds the detector to moderate the D-D neutrons for capture. Neutron kinetic energies are shown in Fig. B.5 as they enter the TPC for various tank thicknesses. The neutrons are further moderated by xenon in the TPC. The NR energy distributions for neutrons prior to capture are shown in Fig. B.6 for various tank thicknesses. For signal events surviving the time cut, discussed in Section B.5.2, the energy dissipated in the TPC before the neutrons are captured is insufficient to produce quanta. The simulation shows that 90% of those signals events are due to neutrons that deposit less than 6×10^{-5} keV_{nr} in the TPC by scattering.

The radial profiles of neutron capture interactions in the active volume of the MiX detector are shown in Fig. B.7. Neutron capture events are concentrated on the edge of the TPC closest to the neutron source. The signal population is also largely near that edge, because there is a geometric advantage for γ cascades escaping the TPC near a wall. Although the fiducial volume

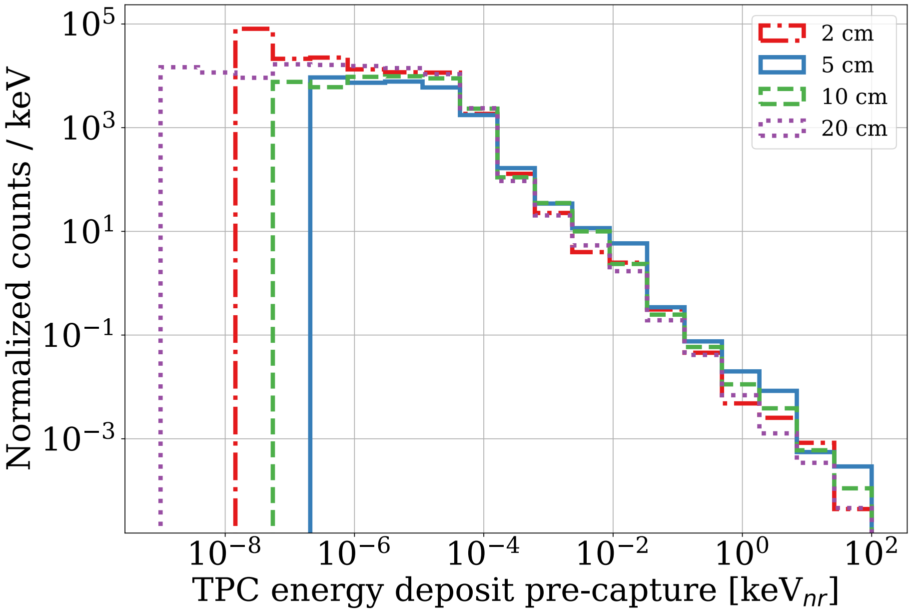


Figure B.6: Distributions of the NR energy transferred to xenon in the TPC by neutron scattering before capture, shown for various thicknesses of the water tank.

in the MiX detector is only well defined within a radius of 29 mm, 80% of the signal events are retained [135].

B.5 Signal optimization

There are two factors that contribute to an optimal signal: a high yield of signal events, and the separability of these events from neutron scattering events. The simulation shows that the presence of a water tank to moderate the D-D neutrons boosts the fraction of signal events, and that its thickness can be tuned to gain a favorable separation of neutron capture events. The analysis of the time distributions of TPC neutron capture events provides suitable values for the neutron generator pulse width w_n and pulsing frequency f. However, w_n and f are more strongly constrained by the rates of background and pile-up events. As discussed in Sections B.5.2 and B.6.3, the optimal parameters for this experiment are a water tank of thickness 5 cm, neutron pulse width $w_n = 30\,\mu$s, and pulsing frequency $f = 60$ Hz.

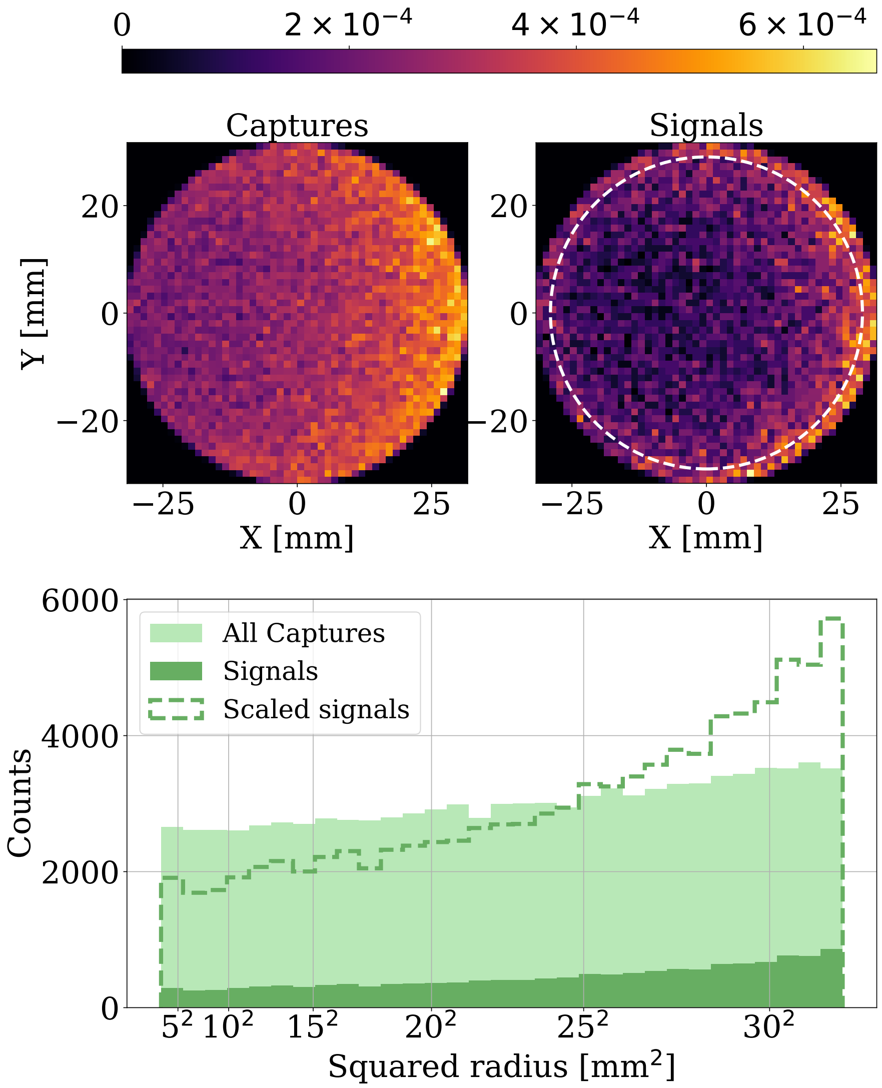


Figure B.7: *Top:* Neutron capture locations in the TPC (left), and signal event locations (right), each normalized to unity. The neutrons enter the water tank from the right, which causes the higher concentration of captures on the right edge. The white circle on the right plot indicates the fiducial radius defined in the MiX detector [135]. Only 20% of signal events fall outside its radius. *Bottom:* Radial positions of capture (light green) and signal (dark green) events. The signal population scaled to the total counts of captures is also shown (dashed) to demonstrate the higher concentration of signal events near the walls of the TPC.

B.5.1 Signal and target energy estimates

Following the optimal configuration presented in Section B.5, estimates for the signal event rate and target energy are discussed. With an instantaneous rate of 10^9 n/s emitted isotropically, 330 neutrons enter the water tank in each pulse. Of these, roughly 0.1 neutrons (0.03%) are captured in the TPC, but only 0.015 events (15% of neutron capture events) end up as signal events. After applying position cuts that only keep events within the MiX fiducial volume, i.e. within a 29 mm radius of the active region, 0.004 signal events per pulse survive. This results in a signal event every 250 pulses, or roughly 1 signal every 4 seconds at a pulsing frequency of 60 Hz. Using an instrumented LXe skin with an estimated capture tagging efficiency of 70%, a final rate of 0.2 usable signal events per second is expected.

The lowest NR energy for this experimental configuration depends on the detector's intrinsic and neutron-induced backgrounds, the exposure, and the scintillation and ionization yields. Even a basic estimate of this target NR energy requires an assumption of the yields below 0.3 keV_{nr} where there is currently no data, in addition to assumptions about the yet unmeasured background levels in the MiX detector. The Noble Element Simulation Technique (NEST) v2.0.1 NR yield model (which was modified to remove the sharp cutoff in the yields at 0.2 keV_{nr}) and the photon evaporation model were used to simulate quanta produced by the neutron capture events. The simulation predicts a drop in quanta production at 0.13 keV_{nr}, where an average of 0.2 ionized electrons are expected. The drop in quanta production was confirmed by weighting the yields directly from NEST with the NR spectrum. The quanta simulated for two months of runtime (10^6 usable signal events) were compared with the weighted NEST yields using a χ^2 test, scanned over various energy thresholds. The threshold energy at which the goodness of the fit stopped improving is consistent with the 0.13 keV_{nr} target energy. The NEST extrapolation predicts an average of 0.2 ionized electrons at 0.13 keV_{nr}. Thus this energy threshold is within reach of a two-month run.

B.5.2 Timing of neutron interactions

Most neutrons are captured between 10 μs and 1 ms after they are emitted by the source as shown in Fig. B.8 for a 5 cm water tank and a 30 μs pulse width. This is due to the joint effect of the neutrons spending most of this time losing energy in the moderator tank and the fact that neutron capture cross sections scale inversely to the incident neutron speed [162]. The *signal window* corresponding to a pulse is defined as the period of time that starts when all neutron scattering has died off, and ends when 99% of neutron capture signal events have been produced. The typical size of a signal window for a 5 cm thick water tank is 0.55 ms.

Less than 2% of the captured neutrons are not slowed down significantly and are captured early in the TPC, creating extra recoil energy due to the collision. This population can be seen in the top left quadrant of Fig. B.8 with recoil energies greater than 0.3 $\mathrm{keV_{nr}}$, which is the maximum recoil energy expected due to γ emissions from stationary xenon nuclei. These events extend up to the width of the pulse, and can be easily removed with a time cut.

An advantage of using a water tank moderator together with a pulsed neutron source is that the former's thermalization effect separates the scattering from capture events in time, creating a pure collection of neutron capture events over several pulses. The signal separability $\mathrm{T_{NR}}$ is quantified (averaged over numerous cycles) as

$$\mathrm{T_{NR}}(E_{\downarrow}, w_n) = \frac{\text{Number of capture signals after the last scatter}}{\text{Number of scattering events}}, \tag{B.2}$$

where only events with deposited energy below $E_{\downarrow}$ are kept, and where the neutron pulse width is w_n. The time at which all the scattering interactions have died off is defined as the last scatter time. Therefore, the numerator represents the signal events in a cycle that are desirable for the measurement since they will not be accompanied by recoil events due to scattering. Figure B.9 shows these populations for $E_{\downarrow} = 1$ keV, a 5 cm water tank, and a pulse width $w_n = 30\,\mu$s. The area of the hatched portion represents the numerator of Eq. B.2 while the area of the elastic

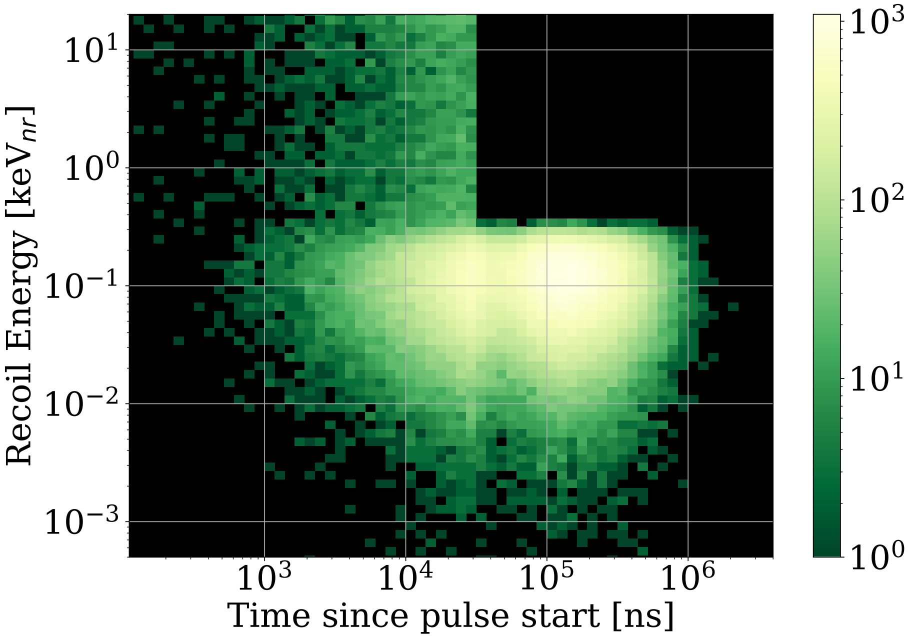


Figure B.8: Recoil energy of the xenon atoms at the time the neutrons were captured for 10^5 neutron captures. The simulation corresponds to a 5 cm water tank and a 30 μs pulse width. A small fraction of neutrons, shown in the top left quadrant, reaches the TPC early with enough energy to cause collisional recoil energies noticeably greater than the 0.3 keV$_{\mathrm{nr}}$ possible by the γ cascades alone. The time of flight of these events is $O(100\,\mathrm{ns})$, so they abruptly cease shortly after the pulse ends at 30 μs.

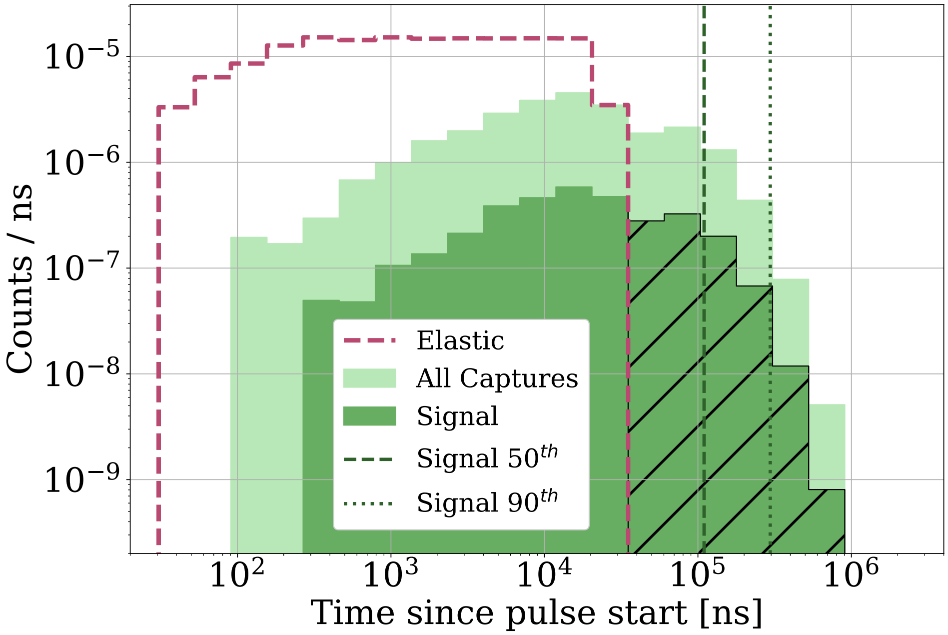


Figure B.9: Time distribution of the neutrons that interact with the active LXe volume in the TPC, from a simulation done for a 5 cm water tank and $w_n = 30\,\mu$s. The total counts due to neutron capture (light green) and elastic scattering (dashed magenta) are normalized to unity. Inelastic scattering events are omitted from this plot for clarity as their rate is a hundred-fold less than the elastic rate. All events shown here deposit less than 1 keV_{nr}. The dark green histogram shows all signal events, and the hatched portion shows the signal events that occur after the last scattering time. Visual checkpoints for when 50% and 90% of all signal events occur are shown with the vertical dashed and dotted lines, respectively.

scatter portion (under the dashed magenta line) represents the denominator. Although a neutron pulse width of $w_n = 30\,\mu$s is used for Fig. B.9, the timing of neutron capture events is not very sensitive to w_n. Rather, it is the timing of the neutron scattering processes, which take place in less than $50\,\mu$s, that more keenly depend on w_n. This fact can be used to maximize the number of signal events that occur after the last scatter time. The effects of varying the pulse width are discussed in Section B.5.3.

Although the time distribution of neutron capture events is relatively unaffected by the pulse width, the time of last scatter and therefore the number of capture events that occur in the signal window is sensitive to w_n. $\text{N}_{\text{signal}}(E_{\downarrow}, w_n)$ is defined as a measure of the fraction of usable signal

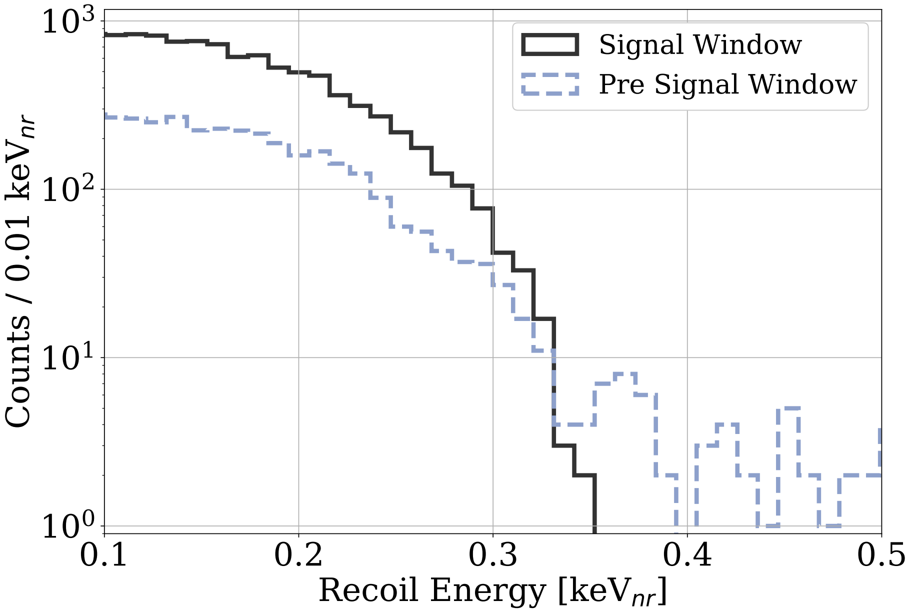


Figure B.10: Recoil energy distributions of the signal events inside the signal window (solid) and before the signal window (dashed) for $w_n = 30\mu$s and a 5 cm water tank moderator. Waiting until the last scatter occurs ensures that the capture of fast neutrons, which are associated with larger recoil energies, are not included in the analysis.

events, such that

$$\mathrm{N}_{\mathrm{signal}}(E_{\downarrow}, w_n) = \frac{\text{Number of capture signals after the last scatter}}{\text{Total number of capture signals}}, \tag{B.3}$$

where as in Eq. B.2, only NR deposits with energy less than $E_{\downarrow}$ are kept. Figure B.10 shows the recoil energy distributions of signal events before and inside the signal window for a 5 cm water tank and $w_n = 30\mu$s. Excluding early signal events with a time cut has the benefit of removing events with extra recoil energies attributed to the faster neutron collisions. This time cut retains a majority (80%) of the signal events.

The metrics T_{NR} and $\mathrm{N}_{\mathrm{signal}}$ summarize the general features of the neutron interaction time structure. These are evaluated for different thicknesses of the water tank and shown in Table B.2 for $w_n = 30\mu$s. The rate of neutron capture in the TPC drops for both small and large tanks. For small tanks, the rate drops due to insufficient neutron moderation. For large tanks, it is due to fewer neutrons making their way into the TPC. However, larger tanks offer a greater degree of

Thickness [cm]	Captures [%]	Signals [%]	T_{NR}	N_{signal}
2	0.007	0.0010	0.02	0.61
5	0.032	0.0047	0.15	0.81
10	0.045	0.0068	0.36	0.89
15	0.033	0.0051	0.57	0.92
20	0.020	0.0030	0.71	0.94
25	0.010	0.0016	0.98	0.95

Table B.2: Properties that influence the choice of water tank moderator thickness, including the neutron capture and signal percentage of neutrons entering the water tank, the signal separability metric $T_{NR}(1\ \text{keV}_{nr}, 30\mu s)$, and $N_{signal}(1\ \text{keV}_{nr}, 30\mu s)$.

scatter-capture separation.

B.5.3 Neutron pulse width

At first glance, it may appear advantageous to have a large pulse width by considering the proportional increase in neutrons emitted per pulse. However, signal events are selected using the signal window time cut, which has an efficiency that depends on the pulse width. Since the signal window is defined to be between when the neutron scattering and neutron capture events end, the timing of the capture and scatter processes are analyzed as a function of pulse width.

The time structure of the neutron capture population does not have a significant dependence on w_n. The thermalization process in the water tank sets a characteristic time scale of $O(100\,\mu\text{s})$ for the neutron capture distribution (see Fig. B.9 for the 5 cm tank). As long as this time scale is greater than w_n, the time structure of the captures is insensitive to changes in w_n. By the same argument it is noted that the neutron scattering population is more responsive to changes in w_n because scatters occur much earlier than the bulk of neutron captures. As an example, the characteristic scattering time set by a 5 cm water tank is around $10\,\mu$s, and thus the resulting time structure is affected by values of w_n larger than $10\,\mu$s (in Fig. B.9 the last scatter time is prolonged to $30\,\mu$s). While the signal window shrinks as its beginning is postponed with increasing w_n, the reduction is negligible until w_n approaches the time scale of neutron captures.

The results are summarized in the top two panels of Fig. B.11. The first panel shows the number of signal events in the signal window as a function of neutron pulse width for various water tank thicknesses. The proportional increase tapers off when w_n approaches the neutron thermalization time set by the water tank, because the signal window begins at later and later times as the scattering events are prolonged. The second panel confirms that the scattering and capture distributions, which set the beginning and end of the signal window, are insensitive to changes in the pulse width until it is on the order of the thermalization time in each tank. These considerations suggest that using an arbitrarily large neutron pulse width would be beneficial, if not for the degradation of the separability metrics $\mathrm{N}_{\mathrm{signal}}$ and T_{NR}, defined in Eqs. B.2 and B.3. These quantities are shown as a function of w_n in the bottom two panels of Fig. B.11. The fall of these metrics at longer pulse widths is caused by the extension of the last scatter time.

As a result the neutron pulse width is constrained to be no more than $O(100\,\mu\mathrm{s})$. The generalization of this constraint can be obtained by comparing the characteristic timing of scattering and capture processes in a detector. A stronger constraint on the pulse width arises when considering the mitigation of background and puile-up events produced directly by neutron captures that could pollute the signal window. This is further discussed in Section B.6.

B.6 Expected backgrounds

We now consider three types of non-NR events that could reduce usable signal counts: i) the low energy ER background from the γ cascades of activated and capture products, and from radiation in the environment, ii) the single electron (SE) background, and iii) the high energy ER events in the TPC. The first two produce small S1 and S2 signals that may overlap the faint signature of neutron capture events. In contrast, the third produces large S1 and S2 signals that may coincide with the signal events in time, temporarily blinding the detector.

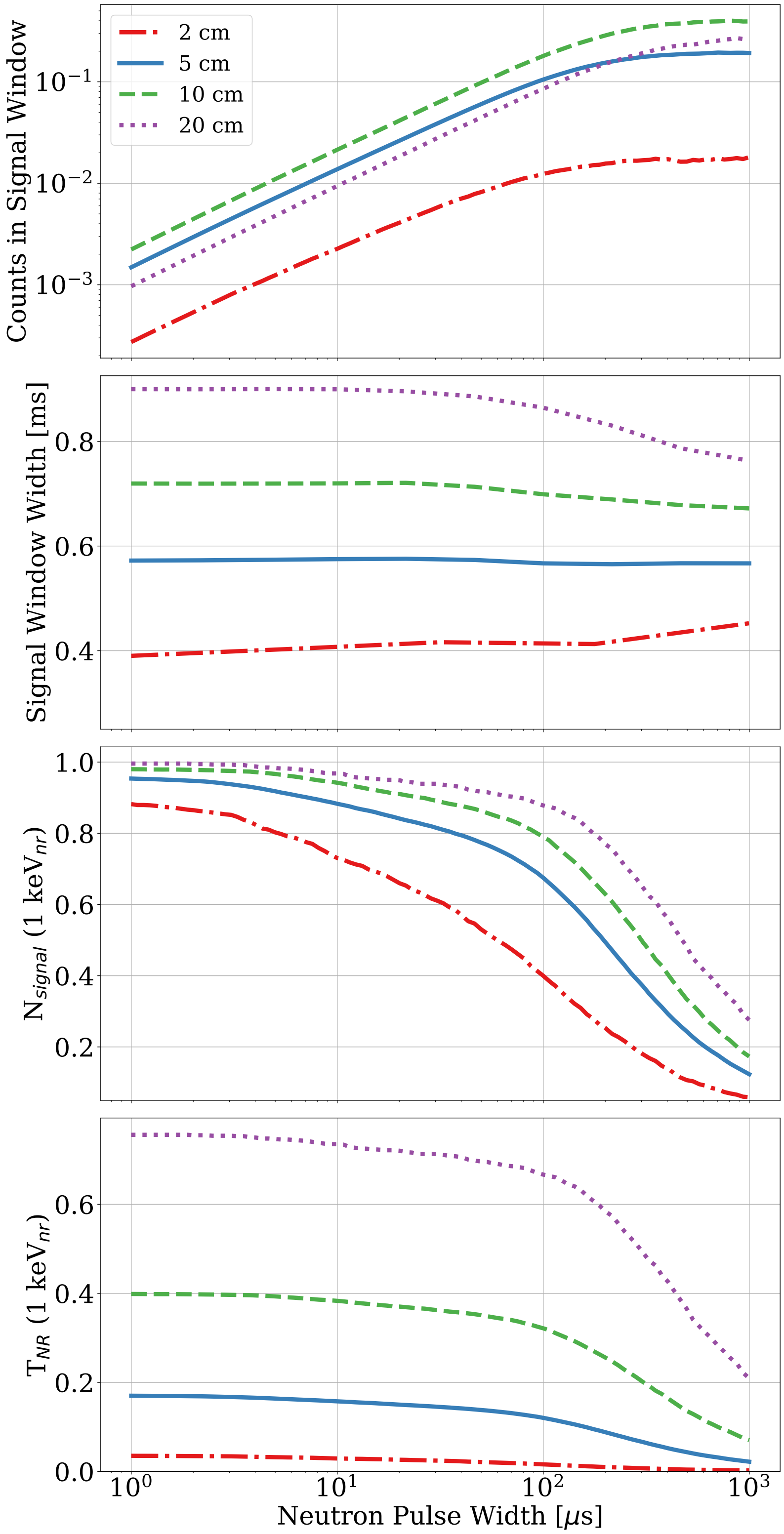


Figure B.11: Simulated metrics as a function of neutron pulse width for 10^9 n/s and various water tank thicknesses. *Top:* Number of signal events falling inside a signal window. *Center top:* Width of the signal window, which begins after the last scattering event and ends when 99% of signal events have been produced after the last scatter. *Center bottom:* N_{signal} for events that deposit less than 1 keV_{nr} in the TPC. *Bottom:* T_{NR} for events that deposit less than 1 keV_{nr} in the TPC.

B.6.1 Low energy ER background

Neutrons in the vicinity of the detector are an indirect source of ER events in the TPC due to the de-excitation cascades of nuclei that undergo neutron interactions. γ-producing neutron interactions (capture or inelastic scatter) can happen both inside and outside the TPC. Most of the ER events in the MiX detector originate from outside the TPC, where there are large amounts of LXe and water (see Fig. B.4). These events are called external ER events, as opposed to internal ER events that are accompanied by a small NR signature. Table B.3 shows neutron capture events partitioned according to where the capture and subsequent ER energy deposit occur.

		ER Deposit	
		Inside TPC	Outside TPC
Neutron Capture	Inside TPC	Internal Bkgd.	Signal
	Outside TPC	External Bkgd.	Undetected

Table B.3: Classification of events based on where the γ-producing neutron interaction and subsequent ER energy deposit took place.

The low energy component of both the internal and external ER events are found to be small compared to the number of neutron capture signal events for the tank thicknesses considered here. Figure B.12 shows the internal and external components of the ER background for a 5 cm water moderator without clustering applied to the energy deposition sites. This represents an upper bound of the ER counts, amounting to less than 0.1% of the number of neutron capture signals below 0.5 keV. As expected for a large volume of LXe outside the TPC, most of the background is external, and a large majority (95%) originate from neutron capture γ cascades.

B.6.2 Single electron Background

Small electron backgrounds are one of the biggest obstacles to the low energy sensitivity of LXe TPCs. Their high rate poses challenges to searches where the expected ionization signal is

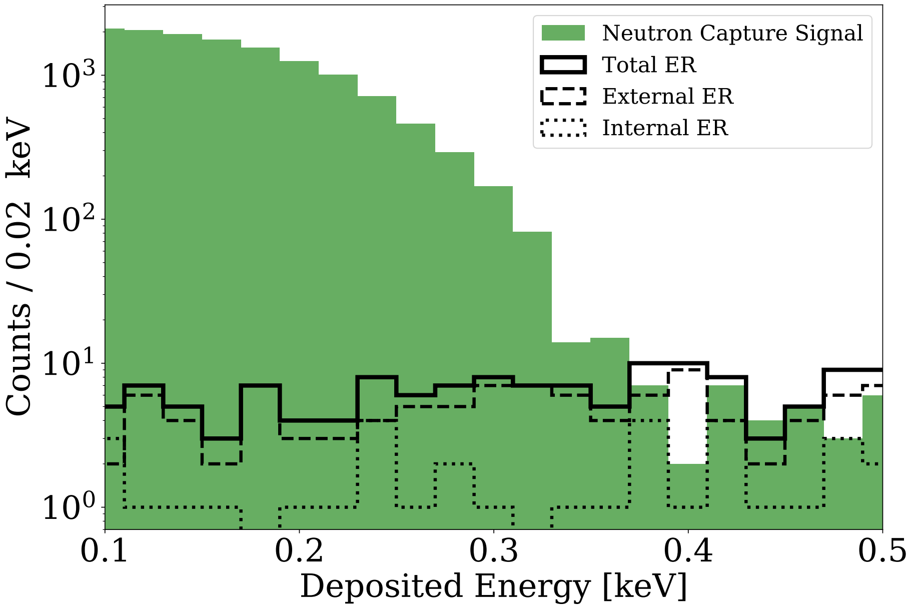


Figure B.12: Deposited energy spectrum due to the internal (dotted) and external (dashed) ER background below 0.5 keV without clustering applied, for a 5 cm water moderator. Also shown is the corresponding recoil spectrum due to the neutron capture signals. This simulation assumes a 1.2 day exposure with a 30 μs pulse width and 60 Hz pulsing frequency, resulting in 20,000 neutron capture signal events. The number of ER counts below 0.5 keV is less than 0.1% of the number of NR signal counts.

only a few electrons, as for ionization-only analyses, or in searches for the coherent scattering of solar neutrinos [172–176]. Single electron (SE) backgrounds are particularly challenging for low energy yield measurements because a significant fraction of NR events below 0.3 keV_{nr} produce only one electron [157]. According to the NEST model, of all neutron capture signals that produce an ionization signal, 80% produce a single electron.

Although the origin of the SE background is not known with certainty, it has been observed that it is almost always preceded by large ionization signals. Background SE events have been observed to persist much longer than the maximum drift time after the initial interaction [99, 177]. Further, this time behavior has been found to depend on runtime parameters like the purity of the LXe and the magnitude of the electric field [178]. Despite the dedicated studies that have been performed using data from multiple detectors, an accurate simulation of this background is still out of reach [99, 179].

The rate of the SE background is expected to be higher in the signal window, due to the capture-induced γ cascades adding on to γ radiation from the environment. The background SE events are indistinguishable from electrons produced by capture-induced recoils of signal events. They will have to be subtracted following a measurement of the SE rate in the signal window, in a manner similar to the background subtraction in Ref. [59]. Due to difficulties in modeling the SE background, it can only be properly addressed after an explicit measurement. After such a measurement there are solutions for mitigation at the hardware [99, 180] and analysis [181] levels.

The background SE rate must be taken into account when deciding the neutron pulsing frequency. Given that all neutron interactions following a $30\,\mu$s pulse die off 1 ms after the start of the pulse (see Fig. B.9), a strict upper bound on the pulsing frequency can be set at 1 kHz. Above this frequency, neutron scatters will start to overlap the isolated capture population. In practice it is likely that high background SE rates in the signal window will disfavor the maximum pulsing frequency, and that some time is needed after a pulse for the background SE rate to decay away. A trade off will have to be made on the pulsing frequency to optimize the number of background-subtracted single electrons produced by signal events. For this study, a pulsing frequency of 60 Hz is assumed based on an investigation into the decay rate of the SE background performed by the LUX experiment, where the intensity of the SE rate was observed to drop tenfold in 16 ms [99]. The pulsing frequency will have to be tuned following a measurement of the SE rate and its decay constant in the MiX detector.

B.6.3 Pileup from high energy ER events

High energy ER events can coincide with the neutron capture signal and contribute to pileup, decreasing the number of clean acquisition windows that contain only the S1 and S2 pulses of the signal event. The source of these ER events can be both internal and external, although in the MiX detector the internal component is negligible. A distribution of ER energy deposits in time, summed over many neutron pulses, is shown in Fig. B.13 for a 5 cm water tank. The two main contributors to this background are the hydrogen in the water and the LXe outside the TPC. In the following, the ER pileup is minimized and constraints for the size of the water tank and length of the neutron pulse are obtained.

Within a given signal window, both the number of high energy ER events originating from neutron capture and the number of neutron capture signals are modeled according to Poisson distributions. This is a valid approach as long as neutron-induced interactions from previous pulses do not leak into the current signal window. As mentioned in Section B.6.2, this leakage would only occur for pulsing frequencies larger than 1 kHz. Each pulse can be treated as independent. The quantity of interest is the probability P that a given signal window has no large ER deposits, while containing a signal event, such that

$$P = \mathrm{Pois}(0, \mathrm{ER}_{\mathrm{external}}) \times \mathrm{Pois}(1, \mathrm{NR}_{\mathrm{signal}}), \tag{B.4}$$

where $\mathrm{ER}_{\mathrm{external}}$ is the average number of external ER events in a signal window, and $\mathrm{NR}_{\mathrm{signal}}$ is the average number of signal events in a signal window. $\mathrm{NR}_{\mathrm{signal}}$ automatically excludes

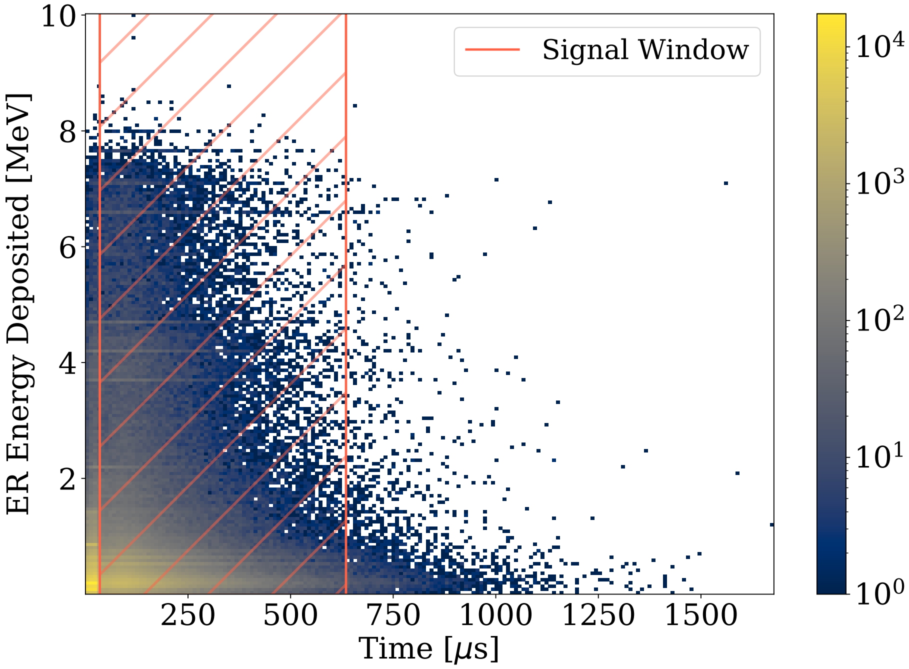


Figure B.13: Distribution of high energy ER deposits in the TPC as a function of time elapsed since the beginning of a neutron pulse of width 30 μs. Simulated for a 5 cm water tank, ER events resulting from 3,000 pulses are shown, corresponding to about 12 signal events. Shown in orange is the signal window for this configuration, presented in the center top panel of Fig. B.11.

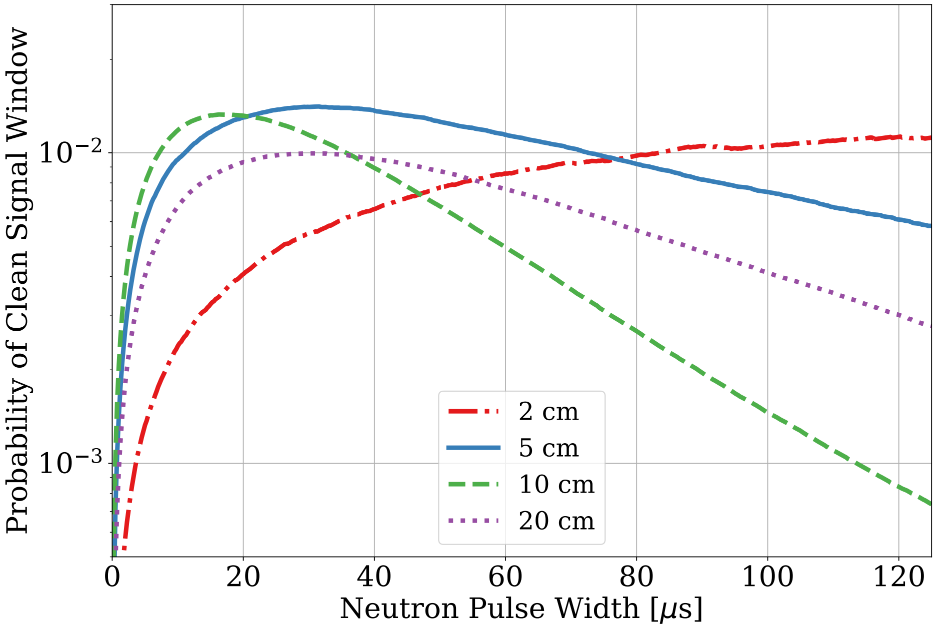


Figure B.14: Probability P of obtaining a clean signal window where no signal event is accompanied by an ER deposit, as a function of neutron pulse width w_n. Curves are shown for a representative set of water tank thicknesses.

internal ER contributions since signal events are defined as captures inside the TPC that are not accompanied by ER deposits. Figure B.14 shows the probability of a given signal window having one signal event and no ER deposits as a function of w_n for several thicknesses of the water tank moderator. While the number of signal events proportionally increases with w_n, there is a corresponding rise in the rate of external ER background due to the capture products of material outside the TPC. This sets a strong constraint on the optimal value for the neutron pulse width, specific to each water tank thickness. A water tank with thickness 5 cm and $w_n = 30\,\mu$s are identified as optimal. Note that the ER pileup can be further mitigated if the capture signal is tagged by its γ cascade using the LXe skin. This allows to precisely determine the time when the capture occurred.

B.7 Implications for dark matter searches

Measurements of the NR quanta yields in LXe below 0.3 $\mathrm{keV_{nr}}$ would provide an absolute detector-independent calibration for LXe experiments. A lower energy threshold allows xenon interactions with slower WIMPs to be detected, increasing the number of observable WIMP events. The increase in counts becomes more significant for lighter WIMPs, where the cutoff velocity for particles to produce detectable NR events is in the tail of the Maxwell-Boltzmann velocity distribution. Figure B.15 shows the gain in sensitivity for light WIMPs in an idealized LXe detector, assuming the yield models in NEST v2.0.1 (which was modified to remove the sharp cutoff in the yields at 0.2 $\mathrm{keV_{nr}}$) are experimentally realized down to 0.1 $\mathrm{keV_{nr}}$ [182]. These sensitivity curves assume a two extracted-electron threshold, a 0% WIMP acceptance for recoil energies below various energy thresholds, and no PMT coincidence required in NEST. In addition to greater sensitivity to light WIMPs, lower energy thresholds offer the following benefits. First, if light (below 10 GeV) WIMPs are discovered, the interaction cross section can be reconstructed with higher precision [183]. The interaction cross section for light WIMPs suffers a degeneracy because only the high-v tail of the velocity distribution is probed. Furthermore, performing an ultra-low energy NR calibration will allow the routine projections down to 0.1 $\mathrm{keV_{nr}}$ found in the direct detection literature to be either corroborated or refuted [184, 185].

B.8 Conclusion

This study focuses on an experimental concept to measure the LXe response to low energy NR interactions using the recoils of neutron capture products. It is emphasized that using a small detector with high light and charge collection efficiencies enhances the chances of a successful measurement.

To establish feasibility, simulations were performed for the MiX detector, a small dual phase TPC designed to maximize light collection, and a pulsed D-D generator as a neutron source. The

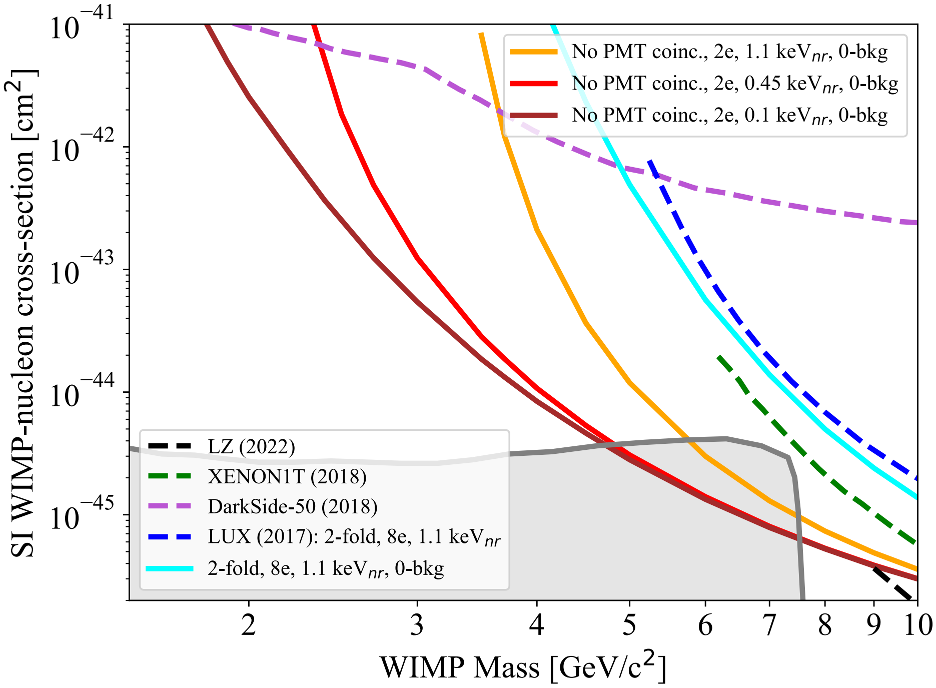


Figure B.15: The projected 90% sensitivities for a generic background-free LXe experiment with a full LUX-like exposure are shown for different energy thresholds in solid maroon, red, and orange curves. The limits are generated using the NEST 2.0.1 default yield models [182]. The searches use both ionization and scintillation channels with no PMT coincidence requirement and a two extracted-electron threshold. A 0% signal acceptance is enforced for recoil energies below the indicated values. The solid and dashed blue curves verify that the LUX result is fairly reproduced [186]. Also shown are limits from LZ (dashed black) [187], XENON1T (dashed green) [188], and DarkSide-50 [189] (dashed purple)

.

small size of the active LXe volume allows about 15% of the γ cascades resulting from neutron capture to escape the TPC, leaving behind a pure NR signature. A pulsed neutron source induces a time structure for the neutron interactions that allows a large fraction of the neutron capture events (80%) to be isolated. These signals can be positively identified using an instrumented LXe volume outside the TPC that can record the γ cascades. The isolated fraction is found to depend on the thickness of the water tank moderator that surrounds the detector, and a trade off has to be made between the higher statistics allowed by smaller tanks, and the higher separability of signals allowed by larger tanks.

The parameters of the neutron generator are found to affect the numbers of signal, background, and pile-up events. A trade off also has to be made between a large neutron pulse width, which increases the number of neutron capture events, and a small pulse width, which decreases the rate of ER pile-up originating from neutron capture outside the TPC. A similar but independent compromise is struck for the pulsing frequency, which is constrained from above to mitigate single electron backgrounds, but needs to be sufficiently high for a proper background subtraction.

The neutron capture population identified in this study constitutes an ideal set of events for probing the scintillation and ionization yields down to $0.13\,\mathrm{keV_{nr}}$, with the recoil events at $0.3\,\mathrm{keV_{nr}}$ serving as a cross-reference to the current lowest measured ionization yield [58]. Whether or not the fundamental limits of NR quanta production are accessible, the results of such a measurement will extend the present knowledge of low energy physics in LXe, and increase the power of direct detection experiments that use it.

B.9 Appendix: Neutron capture model uncertainty

The GEANT4 photon evaporation model simulates γ cascades by sampling the evaluated experimental nuclear structure data (ENSDF) maintained by the national nuclear data center at Brookhaven National Laboratory [190, 191]. The model then uses the γ energy spectra and multiplicity distributions as inputs to generate the neutron capture recoil events. To the best of our knowledge, the γ energy spectra for each multiplicity and the multiplicity distributions have not been measured for any isotope of xenon except ^{136}Xe [170]. A custom model to generate recoil events was written using the γ spectra and multiplicities as input parameters to study the effect on the NR spectrum. The sources of uncertainty arising from those parameters and the sampling method in the custom model were combined to obtain an uncertainty of the GEANT4 NR spectrum. A direct comparison of GEANT4 and data from ^{136}Xe is also made.

B.9.1 NR uncertainty calculation

The evaluated gamma ray activation file (EGAF) is a database of neutron capture γ ray energies and cross sections prepared by the International Atomic Energy Agency [171]. This database was formed by merging elemental γ spectrum measurements taken in 2002 at the Budapest Research Reactor using a high-purity germanium detector with nuclear structure data [192]. By comparing the GEANT4 spectra with the EGAF database, the uncertainty in the NR spectrum was estimated.

Since the EGAF database does not include multiplicity information, it is not possible to adjust the photon energies in GEANT4 and repeat the simulation. Instead, a model is constructed to produce recoil events from a single, multiplicity-independent photon energy spectrum for each isotope. The recoil events generated by the model can be combined according to isotope abundances and an estimate for the multiplicity distributions to produce a final NR spectrum. The algorithm is as follows: Given a normalized photon energy spectrum, neutron separation energy S_n [163], and desired multiplicity κ, take κ-1 random samples from the spectrum and calculate the sum of their energies $E_{\kappa-1} = \sum_{i=1}^{k-1} E_i$. Let E_{min} and E_{max} be the respective lowest and highest photon energies in the sampled spectrum. Then if S_n - $E_{max} < E_{\kappa-1} <$ S_n - E_{min}, set the final γ energy to S_n - $E_{\kappa-1}$ to conserve energy. If $E_{\kappa-1}$ is not in the acceptable range, reject the event and re-sample a new set of $\kappa - 1$ energies. Once a complete set of κ photons has been generated, choose random directions in 4π for each γ and calculate the nuclear recoil using momentum conservation.

To compare the GEANT4 results with the EGAF database, the uncertainty associated with the sampling method must be accounted for. The simulation provides multiplicity-independent γ energy spectra for each isotope. The average multiplicity κ_{avg} is calculated by dividing S_n by the average γ energy. Although it is possible to directly extract multiplicities from GEANT4, a Gaussian distribution about κ_{avg} is assumed so that the same process may be applied to the EGAF data, which does not contain multiplicity information. The effective cross section is calculated for each isotope as the product of the natural abundance and the thermal neutron capture cross section. The effective cross sections are used to weigh the fraction of samples taken from each isotope's energy spectrum. The effective cross sections indicate that ^{129}Xe, ^{130}Xe, and ^{131}Xe make up more than 99% of neutron capture events. Thus the analysis is restricted to these three isotopes. With a total of 500,000 events, the sampling method is used to generate recoils, and the result is compared with recoils simulated by GEANT4. As shown in Fig. B.16, the recoil spectra match well above 0.13 keV$_{\mathrm{nr}}$, validating the sampling method and the multiplicity assumption.

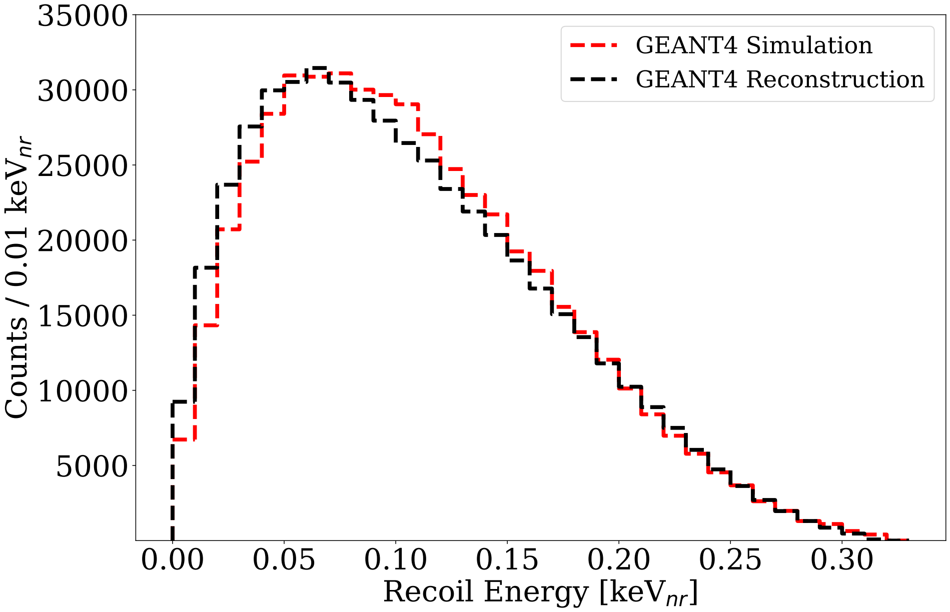


Figure B.16: Nuclear recoil spectra for 500,000 neutron capture events produced by the GEANT4 simulation and reconstruction for ^{129}Xe, ^{130}Xe, and ^{131}Xe. The reconstructed recoil energies skew slightly lower than the GEANT4 simulation, but the spectra match well above 0.13 keV$_{\mathrm{nr}}$.

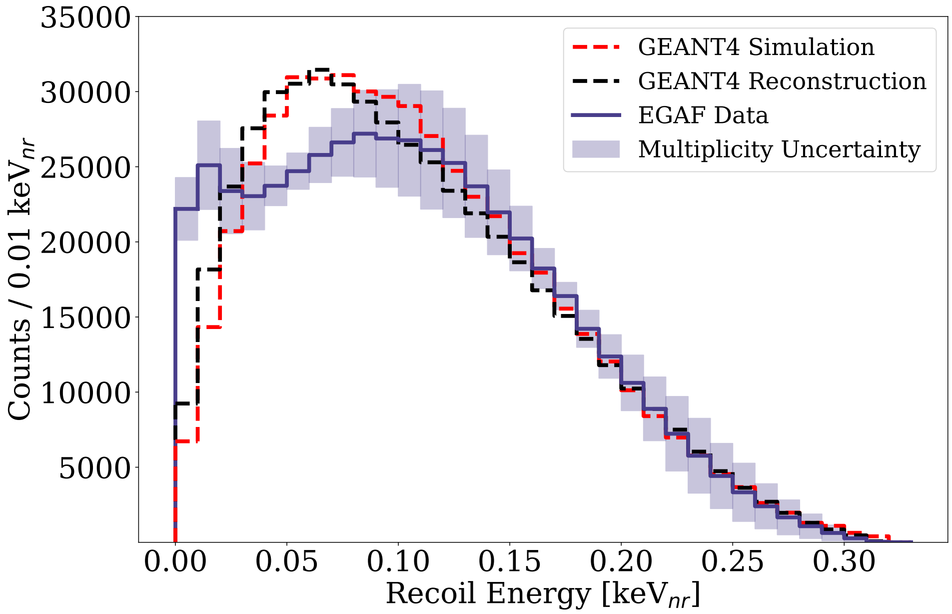


Figure B.17: Nuclear recoil spectra for 500,000 neutron capture events produced by GEANT4 simulation, GEANT4 reconstruction, and reconstruction using the EGAF database. The uncertainty in the EGAF NR spectrum (purple band) is calculated by varying the width of the Gaussian γ multiplicity distribution for each isotope. The EGAF reconstruction matches the GEANT4 simulation closely for energies greater than 0.13 keV$_{\mathrm{nr}}$.

Next, the same sampling procedure is carried out using γ energy spectra from the EGAF database to compare with the NR spectrum obtained from GEANT4. A uniform standard deviation of 1.5 was chosen for the Gaussian multiplicity distributions for the γ spectra obtained from GEANT4, because that value produced the recoil spectrum that most closely matched the GEANT4 simulation. For EGAF, however, there exist no multiplicity distributions. For each isotope, the EGAF data are sampled using Gaussian distributed multiplicities with standard deviations between 1 and 5. The absolute minimum and maximum counts of NR events are calculated in each energy bin. Those counts become the lower and upper bounds for the EGAF NR spectrum, shown in Fig. B.17. The EGAF spectrum matches the GEANT4 simulation well for energies above 0.13 keV$_{\mathrm{nr}}$. The discrepancy at low energies is attributed to disagreements between the GEANT4 and EGAF γ energy spectra for ^{130}Xe.

The three sources of uncertainty, which include i) the uncertainty associated with simplifying assumptions made by the sampling method, ii) the discrepancy between the γ spectra obtained from the EGAF database and the GEANT4 simulation, and iii) the variability in the EGAF NR spectrum calculated from varying the widths of the Gaussian multiplicity distributions, are added in quadrature to produce the final uncertainty band on the NR spectrum, shown in Fig. B.3. Ultimately, the uncertainty in the NR spectrum will be propagated to the yield models following the neutron capture calibration in the MiX detector.

B.9.2 Comparison with ^{136}Xe data

A similar analysis was performed for ^{136}Xe using measurements of neutron capture γ cascades taken at the Detector for Advanced Neutron Capture Experiments (DANCE) at the Los Alamos Neutron Science Center in 2016 [170]. Unlike EGAF, these data include γ energy spectra for each multiplicity as well as the overall multiplicity distribution. Since the maximum recoil energy of ^{136}Xe is $60\,\mathrm{eV_{nr}}$, which is well below the target energy threshold, it will not contribute to the neutron capture calibration. However, it is the only isotope that allows a comparison between the NR spectrum simulated in GEANT4 and a NR spectrum calculated from measured γ spectra with multiplicity information.

^{136}Xe has a relatively small capture cross section and natural abundance (see Table B.1), and contributes only 0.1% of the neutron capture events in natural xenon. Therefore, GEANT4 simulations were run using isotopically pure ^{136}Xe rather than natural xenon to extract both the multiplicity distribution and the γ energy spectra. NR spectra were produced for each multiplicity, then combined according to the weights specified by the multiplicity distribution. The same analysis was performed using DANCE data, and the count-weighted relative difference was calculated in the resulting NR spectra. Taking into account the multiplicity weights from the GEANT4 simulation and the DANCE data eliminates the systematic uncertainties associated with sampling the same γ spectrum irrespective of multiplicity, and the assumption of Gaussian multiplicity distributions. This allows to more accurately quantify the impact of variability of the

γ energies on the resulting NR spectra.

The discrepancy between the ^{136}Xe NR spectra produced using γ energies from GEANT4 and DANCE is represented by a weighted average difference of 40%, taken over the full 0 - 60 eV$_{\text{nr}}$ energy range. Above a 30 eV$_{\text{nr}}$ threshold, the weighted average difference drops to 27%. These differences indicate a disparity between the GEANT4 photon evaporation model and experimental data, and that it is particularly pronounced at low NR energies. However, note that ^{136}Xe is not representative of other xenon isotopes because the GEANT4 γ spectra for ^{136}Xe are sparse (15 lines total) compared to those of more abundant isotopes (more 500 lines each for ^{129}Xe and ^{131}Xe). The small number of γ energies makes the resulting ^{136}Xe NR spectrum sensitive to discrepancies between the γ spectra from GEANT4 and DANCE. The larger number of lines in the other isotopes are expected to lead to smaller differences in the NR spectra due to discrepancies in γ energy distributions.

Bibliography

[1] Lam Hui, Jeremiah P Ostriker, Scott Tremaine, and Edward Witten. Ultralight scalars as cosmological dark matter. *Physical Review D*, 95(4):043541, 2017.

[2] Daniel JH Chung, Patrick Crotty, Edward W Kolb, and Antonio Riotto. Gravitational production of superheavy dark matter. *Physical Review D*, 64(4):043503, 2001.

[3] Bernard Carr, Florian Kühnel, and Marit Sandstad. Primordial black holes as dark matter. *Physical Review D*, 94(8):083504, 2016.

[4] Giorgio Arcadi, Maíra Dutra, Pradipta Ghosh, Manfred Lindner, Yann Mambrini, Mathias Pierre, Stefano Profumo, and Farinaldo S Queiroz. The waning of the wimp? a review of models, searches, and constraints. *The European Physical Journal C*, 78:1–57, 2018.

[5] Jelle Aalbers, DS Akerib, CW Akerlof, AK Al Musalhi, F Alder, A Alqahtani, SK Alsum, CS Amarasinghe, A Ames, TJ Anderson, et al. First dark matter search results from the lux-zeplin (lz) experiment. *arXiv preprint arXiv:2207.03764*, 2022.

[6] Phillip James Edwin Peebles. *Cosmology's century: an inside history of our modern understanding of the universe*. Princeton University Press, 2020.

[7] Friedrich Wilhelm Bessel. On the variations of the proper motions of procyon and sirius. *Monthly Notices of the Royal Astronomical Society*, 6:136–141, 1844.

[8] Fritz Zwicky. Die rotverschiebung von extragalaktischen nebeln. *Helvetica Physica Acta, Vol. 6, p. 110-127*, 6:110–127, 1933.

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

[10] Sinclair Smith. The mass of the virgo cluster. *The Astrophysical Journal*, 83:23, 1936.

[11] Horace W Babcock. The rotation of the andromeda nebula. *Lick observatory bulletin*, 19:41–51, 1939.

[12] HC Van de Hulst, E Raimond, and H Van Woerden. Rotation and density distribution of the andromeda nebula derived from observations of the 21-cm line. *Bulletin of the Astronomical Institutes of the Netherlands, Vol. 14, p. 1*, 14:1, 1957.

[13] 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.

[14] Yoshiaki Sofue. Rotation and mass in the milky way and spiral galaxies. *Publications of the Astronomical Society of Japan*, 69(1), 2017.

[15] Yoshiaki Sofue, Y Tutui, M Honma, A Tomita, T Takamiya, J Koda, and Y Takeda. Central rotation curves of spiral galaxies. *The Astrophysical Journal*, 523(1):136, 1999.

[16] Frank Hohl. Numerical experiments with a disk of stars. *The Astrophysical Journal*, 168:343, 1971.

[17] Jeremiah Paul Ostriker and Philip JE Peebles. A numerical study of the stability of flattened galaxies: or, can cold galaxies survive? *The Astrophysical Journal*, 186:467–480, 1973.

[18] Albert Einstein. Lens-like action of a star by the deviation of light in the gravitational field. *Science*, 84(2188):506–507, 1936.

[19] Douglas Clowe, Anthony Gonzalez, and Maxim Markevitch. Weak-lensing mass reconstruction of the interacting cluster 1e 0657–558: Direct evidence for the existence of dark matter. *The Astrophysical Journal*, 604(2):596, 2004.

[20] Imagesof 1E 0657-56. Accessed: 2023-04-10.

[21] Scott Dodelson and Fabian Schmidt. *Modern cosmology*. Academic press, 2020.

[22] Nabila Aghanim, Yashar Akrami, Mark Ashdown, J Aumont, C Baccigalupi, M Ballardini, AJ Banday, RB Barreiro, N Bartolo, S Basak, et al. Planck 2018 results-vi. cosmological parameters. *Astronomy & Astrophysics*, 641:A6, 2020.

[23] DJ Fixsen. The temperature of the cosmic microwave background. *The Astrophysical Journal*, 707(2):916, 2009.

[24] DJ Fixsen, ES Cheng, JM Gales, John C Mather, RA Shafer, and EL Wright. The cosmic microwave background spectrum from the full cobe* firas data set. *The Astrophysical Journal*, 473(2):576, 1996.

[25] Temperature anisotropy map. Accessed: 2023-04-10.

[26] Planck power spectrum. Accessed: 2023-04-10.

[27] Mark Vogelsberger, Shy Genel, Volker Springel, Paul Torrey, Debora Sijacki, Dandan Xu, G Snyder, Simeon Bird, Dylan Nelson, and Lars Hernquist. Properties of galaxies reproduced by a hydrodynamic simulation. *Nature*, 509(7499):177–182, 2014.

[28] Anadi Canepa. Searches for supersymmetry at the large hadron collider. *Reviews in Physics*, 4:100033, 2019.

[29] Rebecca K Leane, Tracy R Slatyer, John F Beacom, and Kenny CY Ng. Gev-scale thermal wimps: Not even slightly ruled out. *Physical Review D*, 98(2):023016, 2018.

[30] Pablo Fernández de Salas, Khyati Malhan, Katherine Freese, K Hattori, and M Valluri. On the estimation of the local dark matter density using the rotation curve of the milky way. *Journal of Cosmology and Astroparticle Physics*, 2019(10):037, 2019.

[31] Gerard Jungman, Marc Kamionkowski, and Kim Griest. Supersymmetric dark matter. *Physics Reports*, 267(5-6):195–373, 1996.

[32] Katherine Freese, Joshua Frieman, and Andrew Gould. Signal modulation in cold-dark-matter detection. *Physical Review D*, 37(12):3388, 1988.

[33] J Engel. Nuclear form factors for the scattering of weakly interacting massive particles. *Physics Letters B*, 264(1-2):114–119, 1991.

[34] Pekka Pirinen, Jenni Kotila, and Jouni Suhonen. Spin-dependent wimp-nucleus scattering off 125te, 129xe, and 131xe in the microscopic interacting boson-fermion model. *Nuclear Physics A*, 992:121624, 2019.

[35] Martin Hoferichter, Javier Menéndez, and Achim Schwenk. Coherent elastic neutrino-nucleus scattering: Eft analysis and nuclear responses. *Physical Review D*, 102(7):074018, 2020.

[36] BS Hu, J Padua-Argüelles, S Leutheusser, T Miyagi, SR Stroberg, and JD Holt. Ab initio structure factors for spin-dependent dark matter direct detection. *Physical Review Letters*, 128(7):072502, 2022.

[37] Daniel Baxter, IM Bloch, E Bodnia, X Chen, Jan Conrad, P Di Gangi, JEY Dobson, D Durnford, SJ Haselschwardt, A Kaboth, et al. Recommended conventions for reporting results from direct dark matter searches. *The European Physical Journal C*, 81:1–19, 2021.

[38] N Wyn Evans, Ciaran AJ O'Hare, and Christopher McCabe. Refinement of the standard halo model for dark matter searches in light of the gaia sausage. *Physical Review D*, 99(2):023012, 2019.

[39] Jatan Buch, JiJi Fan, and John Shing Chau Leung. Implications of the gaia sausage for dark matter nuclear interactions. *Physical Review D*, 101(6):063026, 2020.

[40] P Adhikari, R Ajaj, DJ Auty, CE Bina, W Bonivento, MG Boulay, M Cadeddu, B Cai, M Cárdenas-Montes, STEFANO Cavuoti, et al. Constraints on dark matter-nucleon effective couplings in the presence of kinematically distinct halo substructures using the deap-3600 detector. *Physical Review D*, 102(8):082001, 2020.

[41] P Cushman, C Galbiati, DN McKinsey, H Robertson, TMP Tait, D Bauer, A Borgland, B Cabrera, F Calaprice, J Cooley, et al. Snowmass cf1 summary: Wimp dark matter direct detection. *arXiv preprint arXiv:1310.8327*, 2013.

[42] Jodi Cooley, Tongyan Lin, W Hugh Lippincott, Tracy R Slatyer, Tien-Tien Yu, Daniel S Akerib, Tsuguo Aramaki, Daniel Baxter, Torsten Bringmann, Ray Bunker, et al. Report of the topical group on particle dark matter for snowmass 2021. *arXiv preprint arXiv:2209.07426*, 2022.

[43] Rouven Essig, Graham K Giovanetti, Noah Kurinsky, Dan McKinsey, Karthik Ramanathan, Kelly Stifter, and Tien-Tien Yu. Snowmass2021 cosmic frontier: The landscape of low-threshold dark matter direct detection in the next decade. *arXiv preprint arXiv:2203.08297*, 2022.

[44] David E Kaplan, Markus A Luty, and Kathryn M Zurek. Asymmetric dark matter. *Physical Review D*, 79(11):115016, 2009.

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

[46] Valentin V Khoze and Gunnar Ro. Dark matter monopoles, vectors and photons. *Journal of High Energy Physics*, 2014(10):1–27, 2014.

[47] Marco Battaglieri, Alberto Belloni, Aaron Chou, Priscilla Cushman, Bertrand Echenard, Rouven Essig, Juan Estrada, Jonathan L Feng, Brenna Flaugher, Patrick J Fox, et al. Us cosmic visions: new ideas in dark matter 2017: community report. *arXiv preprint arXiv:1707.04591*, 2017.

[48] DS Akerib, S Alsum, HM Araújo, X Bai, AJ Bailey, J Balajthy, P Beltrame, EP Bernard, A Bernstein, TP Biesiadzinski, et al. Results from a search for dark matter in the complete lux exposure. *Physical review letters*, 118(2):021303, 2017.

[49] Raymond Davis. A review of the homestake solar neutrino experiment. *Progress in Particle and Nuclear Physics*, 32:13–32, 1994.

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

[51] Babak Abi, Roberto Acciarri, Mario A Acero, Giorge Adamov, David Adams, Marco Adinolfi, Zubayer Ahmad, Jhanzeb Ahmed, Tyler Alion, S Alonso Monsalve, et al. Deep underground neutrino experiment (dune), far detector technical design report, volume ii: Dune physics. *arXiv preprint arXiv:2002.03005*, 2020.

[52] Searching for dark matter. Accessed: 04/20/2023.

[53] Akira Hitachi. Properties of liquid xenon scintillation for dark matter searches. *Astroparticle Physics*, 24(3):247–256, 2005.

[54] Yue Meng, Zhou Wang, Yi Tao, Abdusalam Abdukerim, Zihao Bo, Wei Chen, Xun Chen, Yunhua Chen, Chen Cheng, Yunshan Cheng, et al. Dark matter search results from the pandax-4t commissioning run. *Physical Review Letters*, 127(26):261802, 2021.

[55] E Aprile, K Abe, F Agostini, S Ahmed Maouloud, L Althueser, B Andrieu, E Angelino, JR Angevaare, VC Antochi, D Antón Martin, et al. First dark matter search with nuclear recoils from the xenonnt experiment. *arXiv preprint arXiv:2303.14729*, 2023.

[56] J Aalbers, SS AbdusSalam, K Abe, V Aerne, F Agostini, S Ahmed Maouloud, DS Akerib, DY Akimov, J Akshat, AK Al Musalhi, et al. A next-generation liquid xenon observatory for dark matter and neutrino physics. *Journal of Physics G: Nuclear and Particle Physics*, 50(1):013001, 2022.

[57] DS Akerib, CW Akerlof, D Yu Akimov, A Alquahtani, SK Alsum, TJ Anderson, N Angelides, HM Araújo, A Arbuckle, JE Armstrong, et al. The lux-zeplin (lz) radioactivity and cleanliness control programs. *The European Physical Journal C*, 80:1–52, 2020.

[58] D. Q. Huang. Ultra-Low Energy Calibration of the LUX and LZ Dark Matter Detectors. PhD thesis, Brown University, 2020.

[59] Brian Lenardo et al. Measurement of the ionization yield from nuclear recoils in liquid xenon between 0.3-6 keV with single-ionization-electron sensitivity, 2019.

[60] CS Amarasinghe, R Coronel, DQ Huang, Y Liu, M Arthurs, S Steinfeld, W Lorenzon, and R Gaitskell. Feasibility study to use neutron capture for an ultralow energy nuclear-recoil calibration in liquid xenon. *Physical Review D*, 106(3):032007, 2022.

[61] DS Akerib, CW Akerlof, D Yu Akimov, SK Alsum, HM Araújo, X Bai, AJ Bailey, J Balajthy, S Balashov, MJ Barry, et al. Lux-zeplin (lz) conceptual design report. *arXiv preprint arXiv:1509.02910*, 2015.

[62] DS Akerib, CW Akerlof, D Yu Akimov, A Alquahtani, SK Alsum, TJ Anderson, N Angelides, HM Araújo, A Arbuckle, JE Armstrong, et al. The lux-zeplin (lz) experiment. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 953:163047, 2020.

[63] R Linehan, RL Mannino, A Fan, CM Ignarra, S Luitz, K Skarpaas, TA Shutt, DS Akerib, SK Alsum, TJ Anderson, et al. Design and production of the high voltage electrode grids and electron extraction region for the lz dual-phase xenon time projection chamber. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 1031:165955, 2022.

[64] DS Akerib, CW Akerlof, D Yu Akimov, SK Alsum, HM Araújo, Isaac J Arnquist, M Arthurs, Xinhua Bai, AJ Bailey, J Balajthy, et al. Identification of radiopure titanium

for the lz dark matter experiment and future rare event searches. *Astroparticle Physics*, 96:1–10, 2017.

[65] SJ Haselschwardt, S Shaw, HN Nelson, MS Witherell, Minfang Yeh, KT Lesko, A Cole, S Kyre, and DT White. 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.

[66] EH Miller, J Busenitz, TK Edberg, C Ghag, C Hall, R Leonard, K Lesko, X Liu, Yue Meng, A Piepke, et al. Constraining radon backgrounds in lz. In *AIP Conference Proceedings*, volume 1921, page 050003. AIP Publishing LLC, 2018.

[67] DS Akerib, CW Akerlof, SK Alsum, N Angelides, HM Araújo, JE Armstrong, M Arthurs, X Bai, J Balajthy, S Balashov, et al. Measurement of the gamma ray background in the davis cavern at the sanford underground research facility. *Astroparticle physics*, 116:102391, 2020.

[68] W Turner, A Baxter, HJ Birch, B Boxer, S Burdin, E Fraser, A Greenall, S Powell, and P Sutcliffe. Optical calibration system for the lux-zeplin (lz) outer detector. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 1010:165551, 2021.

[69] DS Akerib, X Bai, E Bernard, A Bernstein, A Bradley, D Byram, SB Cahn, MC Carmona-Benitez, D Carr, JJ Chapman, et al. An ultra-low background pmt for liquid xenon detectors. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 703:1–6, 2013.

[70] B López Paredes, HM Araújo, F Froborg, N Marangou, I Olcina, TJ Sumner, R Taylor, A Tomás, and A Vacheret. Response of photomultiplier tubes to xenon scintillation light. *Astroparticle Physics*, 102:56–66, 2018.

[71] Roberto Acciarri, C Adams, R An, A Aparicio, S Aponte, J Asaadi, M Auger, N Ayoub, L Bagby, B Baller, et al. Design and construction of the microboone detector. *Journal of Instrumentation*, 12(02):P02017, 2017.

[72] AC Fonseca, R Meleiro, V Chepel, A Pereira, V Solovov, and MI Lopes. Study of secondary scintillation in xenon vapour. In *IEEE Symposium Conference Record Nuclear Science 2004.*, volume 1, pages 572–576. IEEE, 2004.

[73] Glenn F Knoll. *Radiation detection and measurement*. John Wiley & Sons, 2010.

[74] Matthew Szydagis, Grant A Block, Collin Farquhar, Alexander J Flesher, Ekaterina S Kozlova, Cecilia Levy, Emily A Mangus, Michael Mooney, Justin Mueller, Gregory RC Rischbieter, et al. A review of basic energy reconstruction techniques in liquid xenon and argon detectors for dark matter and neutrino physics using nest. *Instruments*, 5(1):13, 2021.

[75] AI Bolozdynya. Two-phase emission detectors and their applications. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 422(1-3):314–320, 1999.

[76] VN Solovov, VA Belov, D Yu Akimov, HM Araujo, EJ Barnes, AA Burenkov, V Chepel, A Currie, L DeViveiros, B Edwards, et al. Position reconstruction in a dual phase xenon scintillation detector. *IEEE Transactions on Nuclear Science*, 59(6):3286–3293, 2012.

[77] n_TOF Collaboration, M Mastromarco, A Manna, O Aberle, J Andrzejewski, L Audouin, M Bacak, J Balibrea, M Barbagallo, F Bečvář, et al. Cross section measurements of 155,157 gd (n, $\gamma\gamma$) induced by thermal and epithermal neutrons. *The European Physical Journal A*, 55:1–20, 2019.

[78] M Apollonio, A Baldini, C Bemporad, E Caffau, F Cei, Y Declais, H de Kerret, B Dieterle, A Etenko, L Foresti, et al. Search for neutrino oscillations on a long base-line at the chooz

nuclear power station. *The European Physical Journal C-Particles and Fields*, 27:331–374, 2003.

[79] F Boehm, J Busenitz, B Cook, G Gratta, H Henrikson, J Kornis, D Lawrence, KB Lee, K McKinny, L Miller, et al. Search for neutrino oscillations at the palo verde nuclear reactors. *Nuclear Physics B-Proceedings Supplements*, 91(1-3):91–98, 2001.

[80] FP An, Q An, JZ Bai, AB Balantekin, HR Band, W Beriguete, M Bishai, S Blyth, RL Brown, GF Cao, et al. A side-by-side comparison of daya bay antineutrino detectors. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 685:78–97, 2012.

[81] J Aalbers, DS Akerib, AK Al Musalhi, F Alder, SK Alsum, CS Amarasinghe, A Ames, TJ Anderson, N Angelides, HM Araújo, et al. Background determination for the lux-zeplin (lz) dark matter experiment. *arXiv preprint arXiv:2211.17120*, 2022.

[82] Maris Arthurs. *Radon reduction and the first science results of the LZ experiment*. PhD thesis, University of Michigan, 2022.

[83] Xiao Hua-Lin, Deng Jing-Shan, and Wang Nai-Yan. Oxygen quenching in a lab based liquid scintillator and the nitrogen bubbling model. *Chinese Physics C*, 34(5):571, 2010.

[84] Dmitry Yu Akimov, Alexander I Bolozdynya, Alexey F Buzulutskov, and Vitaly I Chepel. *Two-phase Emission Detectors*. World Scientific, 2021.

[85] Nikolaus Schwentner, Ernst-Eckhard Koch, and Joshua Jortner. *Electronic excitations in condensed rare gases*. Springer, 1985.

[86] James F Ziegler, Matthias D Ziegler, and Jochen P Biersack. Srim–the stopping and range of ions in matter (2010). *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, 268(11-12):1818–1823, 2010.

[87] M Szydagis, N Barry, K Kazkaz, J Mock, D Stolp, M Sweany, M Tripathi, S Uvarov, N Walsh, and M Woods. Nest: a comprehensive model for scintillation yield in liquid xenon. *Journal of Instrumentation*, 6(10):P10002, 2011.

[88] 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.

[89] Tadayoshi Doke, Akira Hitachi, Jun Kikuchi, Kimiaki Masuda, Hiroyuki Okada, and Eido Shibamura. Absolute scintillation yields in liquid argon and xenon for various particles. *Japanese journal of applied physics*, 41(3R):1538, 2002.

[90] Akira Hitachi, Tan Takahashi, Nobutaka Funayama, Kimiaki Masuda, Jun Kikuchi, and Tadayoshi Doke. Effect of ionization density on the time dependence of luminescence from liquid argon and xenon. *Physical Review B*, 27(9):5279, 1983.

[91] GJ Boyle, RP McEachran, DG Cocks, MJ Brunger, SJ Buckman, Sasa Dujko, and RD White. Ab initio electron scattering cross-sections and transport in liquid xenon. *Journal of Physics D: applied physics*, 49(35):355201, 2016.

[92] Oumarou Njoya, T Tsang, M Tarka, W Fairbank, KS Kumar, T Rao, Trey Wager, Soud Al Kharusi, G Anton, Isaac J Arnquist, et al. Measurements of electron transport in liquid and gas xenon using a laser-driven photocathode. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 972:163965, 2020.

[93] JB Albert, PS Barbeau, D Beck, V Belov, M Breidenbach, T Brunner, A Burenkov, GF Cao, WR Cen, C Chambers, et al. Measurement of the drift velocity and transverse diffusion of electrons in liquid xenon with the exo-200 detector. *Physical Review C*, 95(2):025502, 2017.

[94] Ralph L Amey and Robert H Cole. Dielectric constants of liquefied noble gases and methane. *The Journal of Chemical Physics*, 40(1):146–148, 1964.

[95] Jingke Xu, Sergey Pereverzev, Brian Lenardo, James Kingston, Daniel Naim, Adam Bernstein, Kareem Kazkaz, and Mani Tripathi. Electron extraction efficiency study for dual-phase xenon dark matter experiments. *Physical Review D*, 99(10):103024, 2019.

[96] John David Jackson. Classical electrodynamics, 1999.

[97] Chen Cheng, Pengwei Xie, Abdusalam Abdukerim, Wei Chen, Xun Chen, Yunhua Chen, Xiangyi Cui, Yingjie Fan, Deqing Fang, Changbo Fu, et al. Search for light dark matter–electron scattering in the pandax-ii experiment. *Physical Review Letters*, 126(21):211803, 2021.

[98] Elena Aprile, K Abe, F Agostini, S Ahmed Maouloud, M Alfonsi, L Althueser, E Angelino, JR Angevaare, Vasile C Antochi, D Antón Martin, et al. Emission of single and few electrons in xenon1t and limits on light dark matter. *Physical Review D*, 106(2):022001, 2022.

[99] D. S. Akerib et al. Investigation of background electron emission in the LUX detector. *Phys. Rev. D*, 102(9):092004, 2020.

[100] JA Wilkerson, J McLaren, and U Raut. Measurement of the near infrared xenon scintillation yield near atmospheric pressure. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 500(1-3):345–350, 2003.

[101] CAO Henriques, P Amedo, JMR Teixeira, D González-Díaz, CDR Azevedo, A Para, J Martín-Albo, A Saa Hernandez, JJ Gomez-Cadenas, DR Nygren, et al. Neutral bremsstrahlung emission in xenon unveiled. *Physical Review X*, 12(2):021005, 2022.

[102] Jens Lindhard, V Nielsen, M Scharff, and PV Thomsen. Integral equations governing radiation effects. *Mat. Fys. Medd. Dan. Vid. Selsk*, 33(10):1–42, 1963.

[103] Jens Lindhard, Morten Scharff, and Hans E Schiøtt. *Range concepts and heavy ion ranges*, volume 33. Munksgaard Copenhagen, 1963.

[104] J Linhard. Approximation method in classical scattering by screened coulomb fields. *Kgl. Danske Vidensk. Selsk. mat.-fys. Medd.*, 36(10):3–32, 1968.

[105] JD Lewin and PF Smith. Review of mathematics, numerical factors, and corrections for dark matter experiments based on elastic nuclear recoil. *Astroparticle Physics*, 6(1):87–112, 1996.

[106] DS Akerib et al. Discrimination of electronic recoils from nuclear recoils in two-phase xenon time projection chambers. *Phys. Rev. D*, 102(11):112002, 2020.

[107] Spencer Chang, Aaron Pierce, and Neal Weiner. Momentum dependent dark matter scattering. *Journal of Cosmology and Astroparticle Physics*, 2010(01):006, 2010.

[108] A Liam Fitzpatrick, Wick Haxton, Emanuel Katz, Nicholas Lubbers, and Yiming Xu. The effective field theory of dark matter direct detection. *Journal of Cosmology and Astroparticle Physics*, 2013(02):004, 2013.

[109] Nikhil Anand, A Liam Fitzpatrick, and WC Haxton. Model-independent analyses of dark-matter particle interactions. *Physics Procedia*, 61:97–106, 2015.

[110] Andrew Cheek, Darren D Price, and Ellen M Sandford. Isospin-violating dark matter at liquid noble detectors: new constraints, future projections, and an exploration of target complementarity. *arXiv preprint arXiv:2302.05458*, 2023.

[111] Michael Peskin. *An introduction to quantum field theory*. CRC press, 2018.

[112] Injun Jeong, Sunghyun Kang, Stefano Scopel, and Gaurav Tomar. Wimpydd: An object–oriented python code for the calculation of wimp direct detection signals. *Computer Physics Communications*, 276:108342, 2022.

[113] Teal Pershing, Daniel Naim, BG Lenardo, Jingke Xu, James Kingston, Eli Mizrachi, Vladimir Mozin, Phillip Kerr, Sergey Pereverzev, Adam Bernstein, et al. Calibrating the scintillation and ionization responses of xenon recoils for high-energy dark matter searches. *Physical Review D*, 106(5):052013, 2022.

[114] Observation of two-neutrino double electron capture in 124xe with xenon1t. *Nature*, 568(7753):532–535, 2019.

[115] J Aalbers, DS Akerib, AK Al Musalhi, F Alder, SK Alsum, CS Amarasinghe, A Ames, TJ Anderson, N Angelides, HM Araújo, et al. Cosmogenic production of ar 37 in the context of the lux-zeplin experiment. *Physical Review D*, 105(8):082004, 2022.

[116] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. Asymptotic formulae for likelihood-based tests of new physics. *The European Physical Journal C*, 71:1–19, 2011.

[117] Elena Aprile, J Aalbers, F Agostini, M Alfonsi, FD Amaro, M Anthony, Francesco Arneodo, P Barrow, L Baudis, Boris Bauermeister, et al. Effective field theory search for high-energy nuclear recoils using the xenon100 dark matter detector. *Physical review D*, 96(4):042004, 2017.

[118] Jingkai Xia, Abdusalam Abdukerim, Wei Chen, Xun Chen, Yunhua Chen, Xiangyi Cui, Deqing Fang, Changbo Fu, Karl Giboni, Franco Giuliani, et al. Pandax-ii constraints on spin-dependent wimp-nucleon effective interactions. *Physics Letters B*, 792:193–198, 2019.

[119] DS Akerib, S Alsum, HM Araújo, X Bai, J Balajthy, J Bang, A Baxter, EP Bernard, A Bernstein, TP Biesiadzinski, et al. Constraints on effective field theory couplings using 311.2 days of lux data. *Physical Review D*, 104(6):062005, 2021.

[120] W. Haxton and K. McElvain. Ground state to ground state one body density matrices. `https://github.com/Berkeley-Electroweak-Physics/Elastic`. Accessed: 2022-12-10.

[121] David Smith and Neal Weiner. Inelastic dark matter. *Physical Review D*, 64(4):043502, 2001.

[122] Tianqi Chen and Carlos Guestrin. XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, pages 785–794, New York, NY, USA, 2016. ACM.

[123] Greg Rischbieter. Signal Yields and Detector Modeling in Xenon Time Projection Chambers, and Results of an Effective Field Theory Dark Matter Search Using LUX Data. PhD thesis, State University of New York at Albany, 2022.

[124] Logan G Wright, Tatsuhiro Onodera, Martin M Stein, Tianyu Wang, Darren T Schachter, Zoey Hu, and Peter L McMahon. Deep physical neural networks trained with backpropagation. *Nature*, 601(7894):549–555, 2022.

[125] Mirco Hünnefeld. Combining maximum-likelihood with deep learning for event reconstruction in icecube. *arXiv preprint arXiv:2107.12110*, 2021.

[126] Benjamin Nachman. A guide for deploying deep learning in lhc searches: How to achieve optimality and account for uncertainty. *SciPost Physics*, 8(6):090, 2020.

[127] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016. `http://www.deeplearningbook.org`.

[128] Geoffrey E Hinton and Sam Roweis. Stochastic neighbor embedding. *Advances in neural information processing systems*, 15, 2002.

[129] Martin Ester, Hans-Peter Kriegel, Jörg Sander, Xiaowei Xu, et al. A density-based algorithm for discovering clusters in large spatial databases with noise. In *kdd*, volume 96, pages 226–231, 1996.

[130] Markus M Breunig, Hans-Peter Kriegel, Raymond T Ng, and Jörg Sander. Lof: identifying density-based local outliers. In *Proceedings of the 2000 ACM SIGMOD international conference on Management of data*, pages 93–104, 2000.

[131] Fei Tony Liu, Kai Ming Ting, and Zhi-Hua Zhou. Isolation forest. In *2008 eighth ieee international conference on data mining*, pages 413–422. IEEE, 2008.

[132] Maurice G Kendall. A new measure of rank correlation. *Biometrika*, 30(1/2):81–93, 1938.

[133] Scott M Lundberg and Su-In Lee. A unified approach to interpreting model predictions. *Advances in neural information processing systems*, 30, 2017.

[134] Dylan J Temples, Jacob McLaughlin, Jack Bargemann, Daniel Baxter, Amy Cottle, C Eric Dahl, W Hugh Lippincott, Alissa Monte, and Jason Phelan. Measurement of charge and light yields for xe 127 l-shell electron captures in liquid xenon. *Physical Review D*, 104(11):112001, 2021.

[135] S Stephenson et al. MiX: a position sensitive dual-phase liquid xenon detector. *Journal of Instrumentation*, 10(10):P10040, 2015.

[136] Laura Baudis, Patricia Sanchez-Lucas, and Kevin Thieme. A measurement of the mean electronic excitation energy of liquid xenon. *The European Physical Journal C*, 81:1–12, 2021.

[137] G Anton, I Badhrees, PS Barbeau, D Beck, V Belov, T Bhatta, M Breidenbach, T Brunner, GF Cao, WR Cen, et al. Measurement of the scintillation and ionization response of liquid xenon at mev energies in the exo-200 experiment. *Physical Review C*, 101(6):065501, 2020.

[138] E Aprile, J Angle, F Arneodo, L Baudis, A Bernstein, A Bolozdynya, P Brusov, LCC Coelho, CE Dahl, L DeViveiros, et al. Design and performance of the xenon10 dark matter experiment. *Astroparticle Physics*, 34(9):679–698, 2011.

[139] Scott Stephenson. *Probing Spin-Independent WIMP-Nucleon Interactions with the PandaX-I Detector.* PhD thesis, 2015.

[140] Peter Sorensen. Atomic limits in the search for galactic dark matter. *Phys. Rev. D*, 91(8):083509, 2015.

[141] E Aprile and T Doke. Liquid xenon detectors for particle physics and astrophysics. *Rev. Mod. Phys.*, 82(3):2053, 2010.

[142] Godehard Angloher et al. Results from 730 kg days of the CRESST-II dark matter search. *Eur. Phys. J. C*, 72(4):1–22, 2012.

[143] R Agnese et al. Silicon detector dark matter results from the final exposure of CDMS II. *Phys. Rev. Lett.*, 111(25):251301, 2013.

[144] CE Aalseth et al. Results from a search for light-mass dark matter with a p-type point contact germanium detector. *Phys. Rev. Lett.*, 106(13):131301, 2011.

[145] GJ Davies, JD Davies, JD Lewin, PF Smith, and WG Jones. Liquid xenon as a dark matter detector. Prospects for nuclear recoil discrimination by photon timing. *Phys. Lett. B*, 320(3-4):395–399, 1994.

[146] Mark W Goodman and Edward Witten. Detectability of certain dark-matter candidates. *Phys. Rev. D*, 31(12):3059, 1985.

[147] E Aprile et al. Dark matter search results from a one ton-year exposure of XENON1T. *Phys. Rev. Lett.*, 121(11):111302, 2018.

[148] Elena Aprile, CE Dahl, Luiz de Viveiros, RJ Gaitskell, Karl-Ludwig Giboni, J Kwong, P Majewski, Kaixuan Ni, T Shutt, and M Yamashita. Simultaneous measurement of ionization and scintillation from nuclear recoils in liquid xenon for a dark matter experiment. *Phys. Rev. Lett.*, 97(8):081302, 2006.

[149] P Sorensen et al. The scintillation and ionization yield of liquid xenon for nuclear recoils. *Nucl. Instrum. Methods Phys. Res.*, 601(3):339–346, 2009.

[150] A Manzur, A Curioni, L Kastens, DN McKinsey, K Ni, and T Wongjirad. Scintillation efficiency and ionization yield of liquid xenon for monoenergetic nuclear recoils down to 4 keV. *Phys. Rev. C*, 81(2):025808, 2010.

[151] M Horn et al. Nuclear recoil scintillation and ionisation yields in liquid xenon from ZEPLIN-III data. *Phys. Lett. B*, 705(5):471–476, 2011.

[152] E Aprile et al. Response of the XENON100 dark matter detector to nuclear recoils. *Phys. Rev. D*, 88(1):012006, 2013.

[153] Qing Lin, Jialing Fei, Fei Gao, Jie Hu, Yuehuan Wei, Xiang Xiao, Hongwei Wang, and Kaixuan Ni. Scintillation and ionization responses of liquid xenon to low energy electronic and nuclear recoils at drift fields from 236 V/cm to 3.93 kV/cm. *Phys. Rev. D*, 92(3):032005, 2015.

[154] D. S. Akerib et al. Low-energy (0.7-74 keV) nuclear recoil calibration of the LUX dark matter experiment using D-D neutron scattering kinematics, 2016.

[155] Elena Aprile, Matthew Anthony, Qing Lin, Zach Greene, Patrick de Perio, Fei Gao, Joseph Howlett, Guillaume Plante, Yun Zhang, and Tianyu Zhu. Simultaneous measurement of the light and charge response of liquid xenon to low-energy nuclear recoils at multiple electric fields. *Phys. Rev. D*, 98(11):112003, 2018.

[156] D. S. Akerib et al. Improved dark matter search sensitivity resulting from LUX low-energy nuclear recoil calibration. 2022.

[157] BG Lenardo et al. Low-energy physics reach of xenon detectors for nuclear-recoil-based dark matter and neutrino experiments. *Phys. Rev. Lett.*, 123(23):231106, 2019.

[158] KW Jones and HW Kraner. Energy lost to ionization by 254-eV ^{73}Ge atoms stopping in Ge. *Phys. Rev. A*, 11(4):1347, 1975.

[159] JI Collar, ARL Kavner, and CM Lewis. Germanium response to sub-keV nuclear recoils: a multipronged experimental characterization. *Phys. Rev. D*, 103(12):122003, 2021.

[160] R Agnese et al. New results from the search for low-mass weakly interacting massive particles with the CDMS low ionization threshold experiment. *Phys. Rev. Lett.*, 116(7):071301, 2016.

[161] L Thulliez et al. Calibration of nuclear recoils at the 100 eV scale using neutron capture. *Journal of Instrumentation*, 16(07):P07032, 2021.

[162] Gabor Molnar. Handbook of prompt gamma activation analysis, volume 1. Springer Science & Business Media, 2004.

[163] H.D. Choi et al. *Database of Prompt Gamma Rays from Slow Neutron Capture for Elemental Analysis*. International Atomic Energy Agency, Vienna, Italy, 2007.

[164] Sergey I Sukhoruchkin and ZN Soroko. *Excited nuclear states*. Springer, 2012.

[165] K Ni, R Hasty, TM Wongjirad, L Kastens, A Manzur, and DN McKinsey. Preparation of neutron-activated xenon for liquid xenon detector calibration. *Nucl. Instrum. Methods Phys. Res.*, 582(2):569–574, 2007.

[166] D. S. Akerib et al. LUXSim: A component-centric approach to low-background simulations. *Nucl. Instrum. Methods Phys. Res.*, 675:63–77, 2012.

[167] D. S. Akerib et al. Simulations of events for the LUX-ZEPLIN (LZ) dark matter experiment. *Astroparticle Physics*, 125:102480, 2021.

[168] C Poole, Iwan Cornelius, Jamie Trapp, and Christian Langton. A CAD interface for GEANT4. *Australasian Physical & Engineering Sciences in Medicine*, 35:329–34, 2012.

[169] J Allison et al. Recent developments in GEANT4. *Nucl. Instrum. Methods Phys. Res.*, 835:186–225, 2016.

[170] Joshua B Albert, Sean J Daugherty, Tessa N Johnson, Thomasina O'Conner, Lisa J Kaufman, Aaron Couture, John L Ullmann, and Milan Krtička. Measurement of neutron capture on ^{136}Xe. *Phys. Rev. C*, 94(3):034617, 2016.

[171] International Atomic Energy Agency. Prompt gamma-ray neutron activation analysis, March 2014.

[172] J Angle et al. Search for light dark matter in XENON10 data. *Phys. Rev. Lett.*, 107(5):051301, 2011.

[173] E Aprile et al. Low-mass dark matter search using ionization signals in XENON100. *Phys. Rev. D*, 94(9):092001, 2016.

[174] E Aprile et al. Light dark matter search with ionization signals in XENON1T. *Phys. Rev. Lett.*, 123(25):251801, 2019.

[175] D.S. Akerib et al. Results of a search for sub-GeV dark matter using 2013 LUX data. *Phys. Rev. Lett.*, 122(13):131301, 2019.

[176] Elena Aprile et al. Search for coherent elastic scattering of solar ^{8}B neutrinos in the XENON1T dark matter experiment. *Phys. Rev. Lett.*, 126(9):091301, 2021.

[177] Abigail Kopec, Amanda L Baxter, Michael Clark, Rafael F Lang, Shengchao Li, Juehang Qin, and Riya Singh. Correlated single-and few-electron backgrounds milliseconds after interactions in dual-phase liquid xenon time projection chambers. *arXiv preprint arXiv:2103.05077*, 2021.

[178] P Sorensen and K Kamdin. Two distinct components of the delayed single electron noise in liquid xenon emission detectors. *Journal of Instrumentation*, 13(02):P02032, 2018.

[179] E Aprile et al. Observation and applications of single-electron charge signals in the XENON100 experiment. *Journal of Physics G: Nuclear and Particle Physics*, 41(3):035201, 2014.

[180] Peter Sorensen. Electron train backgrounds in liquid xenon dark matter search detectors are indeed due to thermalization and trapping, 2017.

[181] DS Akerib et al. Improving sensitivity to low-mass dark matter in LUX using a novel electrode background mitigation technique. *Phys. Rev. D*, 104(1):012011, 2021.

[182] M. Szydagis et al. New, flexible LXe NR yields and resolution model + G4 improvements + linear noise + much more. *Zenodo,*, 2019.

[183] Annika HG Peter, Vera Gluscevic, Anne M Green, Bradley J Kavanagh, and Samuel K Lee. WIMP physics with ensembles of direct-detection experiments. *Physics of the Dark Universe*, 5:45–74, 2014.

[184] D. S. Akerib et al. Projected WIMP sensitivity of the LUX-ZEPLIN dark matter experiment. *Phys. Rev. D*, 101(5):052002, 2020.

[185] Graciela B Gelmini, Volodymyr Takhistov, and Samuel J Witte. Casting a wide signal net with future direct dark matter detection experiments. *Journal of Cosmology and Astroparticle Physics*, 2018(07):009, 2018.

[186] D. S. Akerib et al. Results from a search for dark matter in the complete LUX exposure. *Phys. Rev. Lett.*, 118(2):021303, 2017.

[187] Jelle Aalbers et al. First dark matter search results from the LUX-ZEPLIN (LZ) experiment, 2022.

[188] E. Aprile et al. Dark matter search results from a one ton-year exposure of XENON1T. *Phys. Rev. Lett.*, 121(11):111302, 2018.

[189] AH Abdelhameed, G Angloher, P Bauer, A Bento, E Bertoldo, C Bucci, L Canonica, Antonio D'Addabbo, X Defay, S Di Lorenzo, et al. First results from the CRESST-III low-mass dark matter program. *Phys. Rev. D*, 100(10):102002, 2019.

[190] Geant4 Collaboration, Geneva, Switzerland. *Geant4 Physics Reference Manual*, 10.7 edition, 2020.

[191] E. Mendoza, D. Cano-Ott, D. Jordan, J.L. Taín, and A. Algora. A new photon evaporation model for Geant4, April 2019.

[192] R.B. Firestone, G.L. Molnar, T. Belgya, D.P. McNabb, and B.W. Sleaford. The evaluated gamma-ray activation file (EGAF). In *AIP Conference Proceedings*, volume 769, pages 219–224, 2005.