# Evaluation of Low-Energy Multi-scattered Neutron Calibrations for LUX-ZEPLIN

A thesis presented for the degree of
Bachelor of Science


**Anna Zuckerman**
**Advisor: Dr. Richard Gaitskell**

Department of Physics
Brown University
April 18, 2021

## Abstract

Over the past several decades, evidence for the existence of dark matter in the universe has become irrefutable. However, the nature of dark matter remains elusive. A prevailing theory is that dark matter takes the form of a Weakly Interacting Massive Particle (WIMP), which could theoretically be detected through various types of experiments. LUX-ZEPLIN (LZ) is one such experiment which seeks to detect WIMPS through the light and charge produced in their collisions with Xenon nuclei in a large underground detector. In order to understand and interpret the results of the LZ experiment, it is important to calibrate the expected light and charge signals for events which deposited known energies into the detector. Thus, various types of calibrations have been developed in the past. However, calibrating for low energies is made difficult by low photon yields, and past calibration techniques have not fully exploited the additional statistics provided by neutrons which scatter multiple times within the detector. This work develops and compares two methods for calibration of the LZ detector at low energies using multi-scattered neutrons.

## Acknowledgements

I would like to thank my advisor Dr. Richard Gaitskell for his guidance and insight over the course of this project. I will take the skills I have learned in this project with me as I continue my academic career. I am also grateful for the help of Casey Rhyne, whose expertise and patience was invaluable to me in starting this project with no prior experience in the field. Many others have also directly or indirectly supported this work, for instance in developing the models I relied on or in devising past calibration techniques. I specifically would like to thank Xin Xiang for his always prompt and helpful responses to queries about NEST. And of course, I want to thank my family and friends for putting up with me while I completed this project.

# Contents

# List of Figures

# Chapter 1

# Introduction

In this work, two methods for calibration of the LZ dark matter detector at low energies are presented and evaluated. Before describing the LZ experiment and the need for calibration, it is important to first understand the motivation for a dark matter search and the principles behind direct detection experiments.

## 1.1 Dark Matter

### 1.1.1 Evidence for the Existence of Dark Matter

The night sky seems full of light from stars and galaxies, and for hundreds of years astronomers used telescopes, and their own eyes, to learn about these bodies which helpfully signal their presence and properties by emitting electromagnetic radiation. However, we now know that over 80% of the matter in the universe does not interact electromagnetically, and so cannot be detected by these otherwise useful pieces of equipment. This matter is termed "dark" because, unlike regular matter which is found mostly in stars, it does not emit light.

Many lines of evidence confirm the existence of dark matter. The idea that unseen matter might exist in the universe was first suggested seriously in the early $20^{th}$ century. In 1904, Lord Kelvin, and later Henri Poincaré, noted that the dispersion of velocities $\langle v^2 \rangle$ of stars within the Milky Way should be related to the total mass M of the galaxy in a predictable way [6]. If the galaxy is assumed to be a stable, gravitationally bound system of stars which has reached equilibrium, the virial theorem states that the (time averaged) potential energy of the system will be equal to twice its (time averaged) kinetic energy. Equating these quantities gives

$$M = \frac{\langle v^2 \rangle r_h}{\alpha G} \tag{1.1}$$

where $r_h$ is the radius containing half of the mass of the galaxy and $\alpha$ is a constant observed to be near 0.45 for most clusters. The measurements available to Poincaré were unable to demonstrate a discrepancy between the observed Milky Way mass and the mass implied by the velocity dispersion, but in 1933 Fritz Zwicky applied a similar theory to the gravitationally bound galaxies within the Coma cluster and found that the velocity dispersion of 80 $km/s$ predicted by equation 1.1 from the observed luminous matter was in stark disagreement to the observed velocity dispersion of 1000 $km/s$ [33]. The finding was the first evidence that significant non-luminous

mass was required for the cluster to remain gravitationally bound despite the high dispersion of velocities of the galaxies it contains.

Another important line of evidence was developed from the rotation curves of galaxies. Measurements of the orbital velocity of stars in the galaxy M31 using the Doppler shifts of stellar emission lines showed no fall off in rotation rate with radius R from the center of the galaxy [25]. This observation could not be reconciled with the expected $1/\sqrt{R}$ decrease that is predicted from simple orbital mechanics if the mass of the galaxy is assumed to follow the disk brightness as a function of radius. The presence of additional dark matter was required in order to produce a mass distribution such that the mass contained within a certain radius continues to increase beyond the radius at which the mass in stars remains constant.

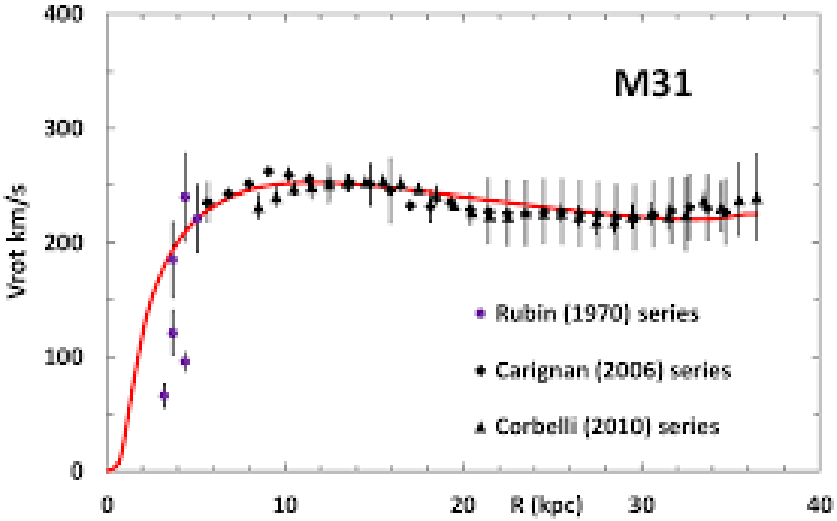



Figure 1.1: Original data from Rubin and Ford, and extended data from more recent measurements, indicate a flat rotation curve for M31 at large radii (red curve). This is fit poorly by the rotation curve expected if all matter was in the luminous disk, requiring additional matter in the form of a dark matter halo. Figure from [20].

Other observations also support the existence of a significant component of non-luminous matter in the universe. For instance, the extent of gravitational lensing of background galaxies by closer galaxies is too great to be produced only by the observed luminous mass of the foreground galaxies. Thus the presence of additional matter in galaxies is required [26].

Cosmological considerations also provide evidence for the existence of dark matter. For instance, the scale of fluctuations in the Cosmic Microwave Background indicate initial overdensities in baryonic matter which are too small to produce the observed formation of structure in the universe. Dark matter is necessary to provide the initial overdensities in the early universe from which structure grew [23]. These structure formation observations support a cold, rather than hot, dark matter component, because hot dark matter has the tendency to erase any local inhomogeneities in matter distribution on scales smaller than its free-streaming length (which can be thought of as the particle speed multiplied by the age of the universe). Because dark matter does not interact electromagnetically, its early distribution was not recorded in the observed CMB fluctuations.

In addition, models of the nucleosynthesis which occured during the Big Bang

provide strong contraints on the amount of baryonic matter in the universe, and require a total baryonic mass which is too small to explain the observed lack of curvature in the universe on its own. Measurements of the D/H ratio give $\Omega_b h^2 = 0.022 \pm 0.02$ [15], which alone could not produce a flat universe. Though some theories suggest baryonic forms of dark matter, this can clearly not contribute most of the mass of dark matter in the universe. Thus, most theories predict that this additional matter is non-baryonic.

### 1.1.2 Dark Matter Candidates

The above lines of reasoning all point to the existence of additional invisible matter in the universe, but say little about its identity. Evidence from both particle physics and cosmology can help provide candidates and constrain their properties. Any dark matter candidate must satisfy three basic characteristics: it must not interact with electromagnetic radiation (and thus be 'dark'), must have a nonzero mass (in order to interact gravitationally with other matter), and must be stable on the timescale of the age of the universe. Several potential identities for dark matter have been suggested, including MAssive Compact Halo Objects (MACHOs), Axions, sterile neutrinos, and Weakly Interacting Massive Particles (WIMPS). The LUX-ZEPLIN experiment seeks to detect WIMP dark matter, and so a few other potential candidates are described here only briefly, in order to provide context for the subsequent discussion of direct detection techniques.

MACHOS are perhaps the most intuitive initial guess as a dark matter candidate. It seems reasonable that the observed excess of mass could stem from a large quantity of regular matter that is simply non-luminous, such as black holes, neutron stars, brown dwarfs, or even exoplanets. These MACHOS could be lighter than one solar mass, or up to several hundreds of solar masses. If these bodies constitute a large fraction of the mass of galaxies, this could account for the inability of mass estimations which assume most matter is luminous in the form of stars to correctly predict galactic rotation curves or velocity dispersions. However, MACHOS are considered unlikely as a major component of dark matter primarily because they are composed of baryonic matter, and are thus inconsistent with the Big Bang nucleosynthesis models discussed above which predict a component of baryonic matter smaller than the total matter component of the universe. Some groups (for example, the MACHO, OGLE, and MOA projects) have searched for MACHOs by attempting to observe gravitational microlensing of starlight when a MACHO passes between a star or galaxy and the Earth.

Most theories postulate a particle form of dark matter. Besides the WIMP, the Axion is a particle candidate which is well-motivated by particle physics. Axions are a type of particle predicted by quantum chromodynamics to solve a problem known as the strong Charge-Parity (CP) problem. CP symmetry is, to oversimplify greatly, the concept that if the charge of a particle is flipped, the antiparticle results (i.e., an electron becomes a positron), and if the spatial coordinates are flipped, the mirror image of the particle results (i.e., a particle traveling to the left travels to the right). The fact that CP symmetry is violated for weak-force interactions, though not for electromagnetic or strong interactions, accounts for the overabundance of matter relative to antimatter in the universe. A full discussion of the strong CP problem is outside the scope of this work, so let it simply be noted that Axions

arise naturally from one way of resolving the strong CP problem. Axions have the correct properties to be a dark matter candidate, and are predicted to have masses on the order of $\mu eV/c^2$ to $meV/c^2$. One method of detecting Axions, for example as implemented by the Axion Dark Matter Experiment (ADMX), is to search for their predicted decay into photons under a strong applied magnetic field.

Sterile neutrinos are another postulated particle dark matter candidate. The three known neutrinos all have a property termed "left-handed" chirality, which describes the orientation of their quantum mechanical spin relative to their direction of motion. An additional "right-handed" neutrino is predicted which would be sterile in the sense that it does not feel the weak force with full strength. A large range of masses are possible, spanning $eV/c^2$ to $GeV/c^2$ scales. Due to their lack of electromagnetic interaction, sterile neutrinos could provide a dark matter candidate [8]. Experiments such as Fermilab's MiniBooNE have attempted to detect sterile neutrinos through the X-rays emitted during their decay.

It should also be noted that other potential explanations for the apparent missing mass in the universe have been postulated, and in some cases experiments could be designed to test these theories. For instance, it is possible that current theories of gravity are missing important physics on large spatial scales and under very small accelerations, in such a way as to account for the observed galactic rotation curves. This is the motivation behind theories of Modified Newtonian Dynamics, or MOND, originally proposed by Mordehai Milgrom in 1983. MOND has been tested as a good fit to galaxy rotation curves, but fails to explain other problems such as the need for dark matter to justify CMB fluctuations and the formation of structure [10]. Other theories include the effects of additional dimensions (for example, [19]), and consequences of Grand Unified Theory physics [30]. It is also possible that dark matter is not composed of only one type of candidate, but instead is made up of multiple of the above mentioned (or other) candidates.

### 1.1.3 WIMPS

This work focuses on another particle dark matter candidate: the WIMP, or Weakly Interacting Massive Particle. WIMPS are an important dark matter candidate because they both explain the invisible mass in the universe and are strongly predicted by particle physics theories. Before developing the ideas behind WIMP detection, it is important to mention the basics of their expected characteristics and behavior.

Supersymmetry (SUSY) is a principle of particle physics which extends the Standard Model by predicting an additional superpartner particle for each Standard Model particle with spin that differs by 1/2. The lightest of these (the Lightest Supersymmetric Particle, or LSP) interacts only weakly and is required to be stable because it cannot decay into a lighter particle, and is therefore a possible dark matter candidate. The LSP is often considered to be the neutralino, making the neutralino the most commonly accepted identity of the WIMP. The details of SUSY theory which predict the properties of the LSP are beyond the scope of this paper (for further details see [14]). However, some important results which constrain WIMP properties can be discussed. For instance, SUSY predicts a spin-independent cross section for interaction with Standard Model nucleons on the order of $10^{-4} - 10^{-8}$ pb. The total interaction cross section also has a component dependent on nucleon spin, but this is usually dominated by the spin-independent component. WIMP mass is

also not well constrained, though SUSY requires that it must be less than approximately 1 $TeV/c^2$ [18]. A lower limit on mass of the order 1 $keV/c^2$ is necessitated by the need to produce the formation of the observed structure in the universe; at lower masses it would not be possible to develop sufficient local overdensities in the early universe [24]. If WIMPS were indeed produced early on in the evolution of the universe, and if their mass falls in the 10 $MeV/c^2$ to 10 $TeV/c^2$ range, this would produce the observed relic density of dark matter. This is known as the WIMP miracle [30], and is one piece of evidence which supports WIMP dark matter.

For a more detailed summary of the evidence and candidates for dark matter, see [7].

### 1.1.4 Detection Techniques

Many different techniques have been developed with the goal of detecting WIMPS, and though none have ever convincingly done so, they have together succeeded in narrowing the possible parameter space in which to find WIMPS. Detecting WIMPS would allow us to begin to understand their properties.

Detection techniques can be classified into three general categories: indirect detection, direct detection, and production of WIMPS in colliders. Indirect detection techniques attempt to observe the signatures of WIMP annihilation into Standard Model particles. Direct detection methods seek to detect dark matter by measuring the energy deposited when WIMPS collide with nuclei in a detector. In addition, using particle accelerators, some groups hope to find WIMPS through the momentum that is apparently missing during standard model particle collisions which produce dark matter. These strategies are summarized in Figure 1.2.

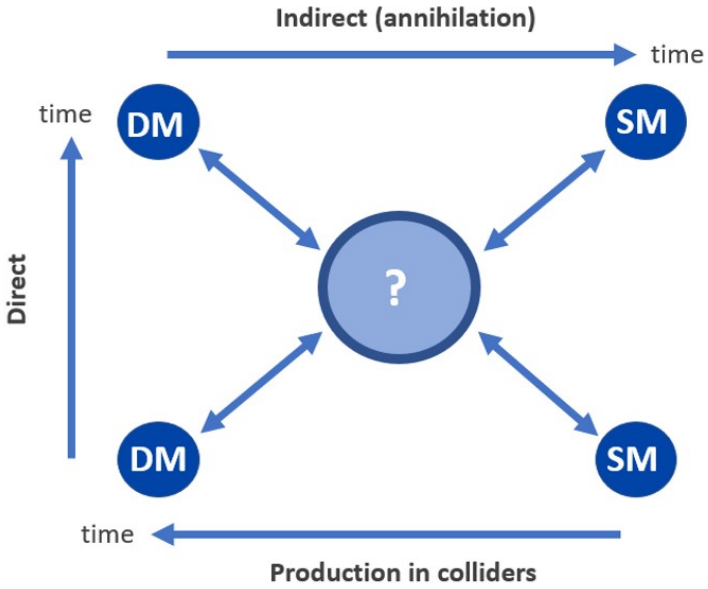



Figure 1.2: The three main types of dark matter detection techniques. Reading time in the direction of the arrow for each type summarizes the dark matter (DM) and standard model (SM) particle interactions involved in each method. Figure after [21].

Indirect detection techniques seek to measure the signals produced when dark matter particles collide and annihilate, by searching regions expected to contain a high density of dark matter (for example, the galactic center, dwarf spheroidal galaxies, and galaxy clusters). Supersymmetry predicts that dark matter particles will have a corresponding antiparticle. When a dark matter particle collides with its antiparticle, they can annihilate to form Standard Model particles such as positrons, which then annihilate to produce gamma radiation with an energy similar to that of the progenitor particles, and also neutrinos [7]. The annihilation rate $\Gamma$ should be proportional to the DM particle abundance $n_{DM}$ as well as the annihilation cross section $\langle \sigma v \rangle$ which governs the probability with which particles interact and annihilate. For particles in the early universe to be in thermodynamic equilibrium, the creation rate was equal to the annihilation rate. If the universe had simply cooled without expanding, the number density would be exponentially suppressed as time progressed. However, the universe is expanding, and so the relic density of dark matter particles $\Omega_{DM}$ was fixed (i.e., dark matter dropped out of equilibrium) at the time the expansion rate was equal to the annihilation rate. Thus the annihilation cross section can be determined from the relic dark matter density to be $\langle \sigma v \rangle \sim 10^{-26}\ cm^3 s^{-1}$. The fact that this cross section is consistent with interaction on the scale of the weak force provides another way of looking at the 'WIMP miracle', because it strongly suggests the particle physics predictions for WIMP interactions are in line with astrophysical requirements on dark matter density.

Both ground and space-based experiments have been developed with the intent to detect gamma ray signals from WIMP annihilation. Recent measurements of the galactic center from the Fermi Gamma-ray Space Telescope have been shown to be consistent with WIMP annihilation, though some authors argue the radiation can be explained by models of the galactic bulge without dark matter [1]. WIMP annihilation is a speculated source of the excess 511 keV emission line observed from the Milky Way center, and would be consistent with the annihilation of WIMPS if their mass was on the order of MeV, or with the de-excitation or decay of more massive GeV to TeV particles [9].

Instead of waiting for a signal from dark matter to reach us, another method for detecting dark matter relies on creating WIMPS on Earth in particle accelerators (such as CERN's Large Hadron Collider) through high energy collisions of Standard Model particles. The primary strategy of these experiments is to accelerate protons to several TeV energies, at which point they can be better considered as composite particles of quarks and gluons rather than elementary particles, and then to collide two protons together [21]. A series of detectors surrounds the accelerator track, which is capable of reconstructing the paths of produced Standard Model particles and measuring the momenta of all Standard Model particles resulting from a collision. Because the protons are incident along the beam path, they have no momentum transverse to the collision axis. This means that if the net measured momentum is found to have some nonzero component perpendicular to the beam axis, there must be an equal component of momentum in the opposite direction carried away by an undetected particle. This particle could then be the sought-after WIMP. Although each of these types of detection experiments is fascinating and promising in the search for dark matter, LZ is a type of direct detection experiment, and so the next section will focus on providing a brief overview of direct detection.

## 1.2 Direct Detection

Direct Detection experiments rely on the small, but nonzero, probability that dark matter particles will interact with regular baryonic matter. This class of experiments search for the energy deposited during collisions between dark matter particles and atomic nuclei in specially designed detector volumes. The expected rate of these collisions depends on the local flux of dark matter, its cross section for interaction with the target nucleus, and of course the volume of the detector itself. The flux of dark matter through a detector depends on the local density of dark matter, as well as its velocity relative to the Earth. The local density of dark matter can be determined from the observed Milky Way rotation curve by fitting the measured motions of stars at given radii from the galactic center to a best fit galaxy density profile. Though the dark matter density calculated in this way is dependent on the model chosen for the structure of the dark matter halo [21] and also on the satellite from which the data is taken, the most recent measurements using the Sloan Digital Sky survey are consistent with a local dark matter density of $\rho_{DM} \approx GeV/c^2/cm^3$. The velocity of WIMPS in the region of the Earth is also dependent on the halo model chosen. However, an instructive calculation can be made by first considering the local circular velocity of a body in orbit about the galactic center. Simple orbital mechanics for circular orbits (i.e., at radius R setting the outward centripetal acceleration $\frac{m_{DM}v^2}{R}$ equal to the inward gravitational acceleration $\frac{Gm_{DM}M_{MW}(R)}{R^2}$ gives a circular velocity

$$v = \sqrt{\frac{GM_{MW}(R)}{R^2}} \tag{1.2}$$

where $M_{MW}$ is the Milky Way mass at radius $R$, and G is the gravitational constant. A very simple, though not perfectly accurate, assumption, would be that all the mass in the Milky Way ($\sim 10^{12}\ M_o$) is inside Earth's orbit at radius $R \sim 10^{21}\ m$. This gives $v_c \sim 370\ km/s$. Calculations from the observed motions of nearby stars, using a more accurate Milky Way density profile, give $v_c \sim 232 \pm 14\ km/s$ [4]. The upper limit on velocity comes from the escape velocity at the Sun's distance from the galactic center. This is found by setting non-relativistic kinetic energy equal to the gravitational potential

$$v_{esc} = \sqrt{\frac{2GM_{MW}(R)}{R^2}} \tag{1.3}$$

The simple assumptions above gives $v_{esc} \sim 510\ km/s$, which is in reasonable agreement with reported values (based on observed local star velocities) of $v_{esc} = 544 \pm 55\ km/s$ [27]. The most commonly accepted halo model is known as the Standard Halo Model (SHM), and assumes an isothermal and isotropic dark matter distribution, with $R^{-2}$ density fall off with radius from the galactic center. This includes a Maxwellian velocity distribution for dark matter particles, and so predicts the dispersion in velocities will be $\sigma(R) = \sqrt{3/5}v_cR$. [4]. The net result of these calculations is that the expected WIMP velocity in the neighborhood of the Earth is given by the local circular velocity $v_c$, plus the dispersion in WIMP velocities $\sigma$.

An order of magnitude calculation of the expected flux of WIMPS through a detector can be calculated from the expected WIMP velocity and the local WIMP density. This gives a flux (mass per detector volume per time) on the order of $\sim 10^7\ GeV/c^2/cm^3/s$.

Additional complications come from the Sun's small deviation from the local circular velocity ($v_o \approx v_c + 12\ km/s$), and the Earth's motion around the Sun, which adds a sinusoidal modulation to the expected WIMP flux. The Earth orbits the Sun with $v_e \approx 30\ km/s$ relative to the Sun. Thus, an annual modulation in the relative velocity between WIMPS and Earth-based detectors should be on the order of $v_e/v_c \sim 5$. Such an oscillation in detected signal has been reported by the DAMA/LIBRA group [5], though this result is the subject of much skepticism in the scientific community due to the inability of subsequent experiments to replicate the result, and the lack of public availability of DAMA/LIBRA data.

We can also define a differential rate of expected events, as a function of recoil energy (the energy deposited by a WIMP during a collisions). Intuitively, this can be thought of as the expected contribution to the event rate from each bin of recoil energies. This depends on the (energy dependent) local WIMP number density, the cross section for interaction $\sigma_o$, the (energy dependent) nuclear form factor, and the velocity distribution. The differential rate can be expressed as

$$\frac{dN}{dE_{NR}} = \frac{\sigma_o \rho_{DM}}{2\mu^2 m_{DM}} F^2(q) \int_{v_{min}}^{v_{max}} \frac{f(v)}{v}\, dv \tag{1.4}$$

where $E_{NR}$ the recoil energy, $\mu$ is the reduced mass of the WIMP and target nucleus, $\sigma_o$ is the cross section for interaction, $F$ is nuclear form factor which depends on the characteristics of the target nucleus, and the integral is over the distribution of velocities of WIMPS reaching the detector. A full derivation of this expression involves particle physics beyond the scope of this paper, but can be found in, for example, [29]. For a given cross section and WIMP mass, Figure 5 shows that the differential rate takes the form of a decreasing exponential with recoil energy. For Xenon (the target material used in LZ), integrating this gives an expected rate of four events per tonne per year.

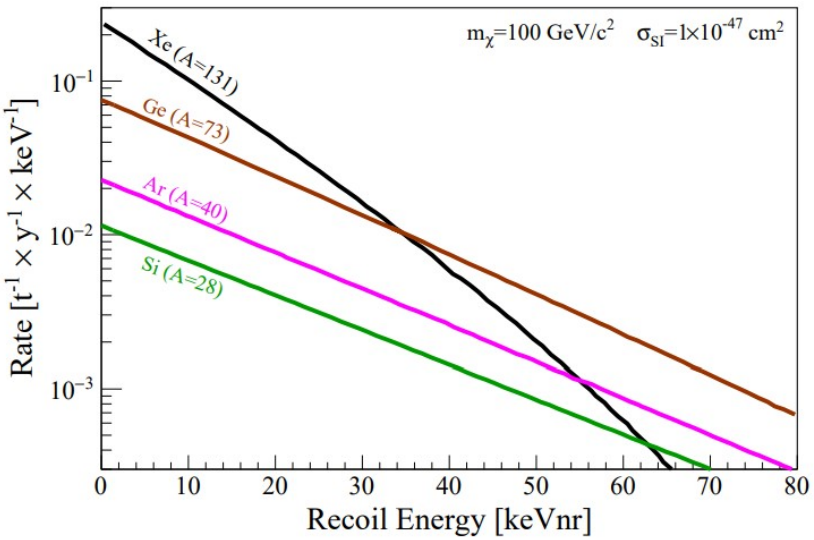



Figure 1.3: The expected event rate as a function of nuclear recoil energy, under illustrative chosen WIMP mass and interaction cross section. The form is in general a decaying exponential with recoil energy, as shown for selected detector materials. Figure from [27].

The goal of direct detection experiments is to measure the energy deposited during a WIMP – nucleus collision. This energy can be related to the initial energy of the incident WIMP and the scattering angle using simple kinematics. If the collision is modeled as 'billiard ball' elastic scattering (see figure 1.4), this relationship can be determined simply through conserving momentum (parallel and normal to the collision axis) and energy, and solving for the final velocities in terms of the initial velocities, center-of-mass scattering angle, and reduced mass of the two particle system. For each scatter event, this gives

$$E_{NR} = \frac{4M_{DM}M_T}{(M_{DM}+M_T)^2}\frac{1-cos(\theta_{COM})}{2}E_{DM} \tag{1.5}$$

where $M_T$ is the target mass, $M_{DM}$ is the WIMP mass, the energy values are measured in the lab frame, and the scattering angle measured in the center-of-mass frame can be determined from the lab-frame scattering angle by $tan(\theta_{LAB}) = \frac{sin(\theta_{COM})}{m_{DM}/M_T} + cos(\theta_{COM})$.

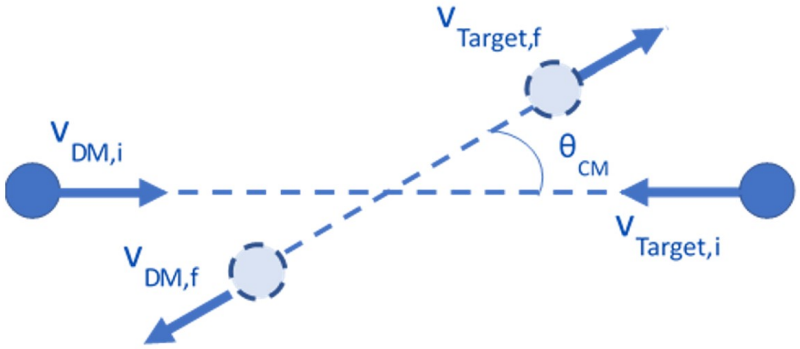



Figure 1.4: Schematic of the "billiard ball" scattering kinematics that describes the collision of a WIMP with a target nucleus, with all quantities measured in the center-of-mass frame.

# Chapter 2

# The LUX-ZEPLIN Experiment

## 2.1 LZ Design

The LUX-ZEPLIN (LZ) Experiment is the offspring of two previous experiments: Large Underground Xenon (LUX) and ZonEd Proportional scintillation in LIquid Noble gases (ZEPLIN). Its aim is to directly detect the interactions of WIMP dark matter particles with the detector, thereby providing the first ever dark matter detection, or alternatively to push to limits on sensitivity to as yet unsearched regions of parameter space. It consists of seven tonnes of liquid Xenon (LXe) buried 1.5 km underground at the Sanford Underground Research Facility in Lead, SD [12]. Due to the low expected event rate ($< 0.1$ event per day per tonne of Xe), and the low energies deposited in each interaction ($< 100\ keV$ per event), the LZ experiment employs various techniques to maximize its detection capabilities [22]. An overview of the LZ detector design is given below in order to provide context for the discussion of signal detection and calibration techniques.

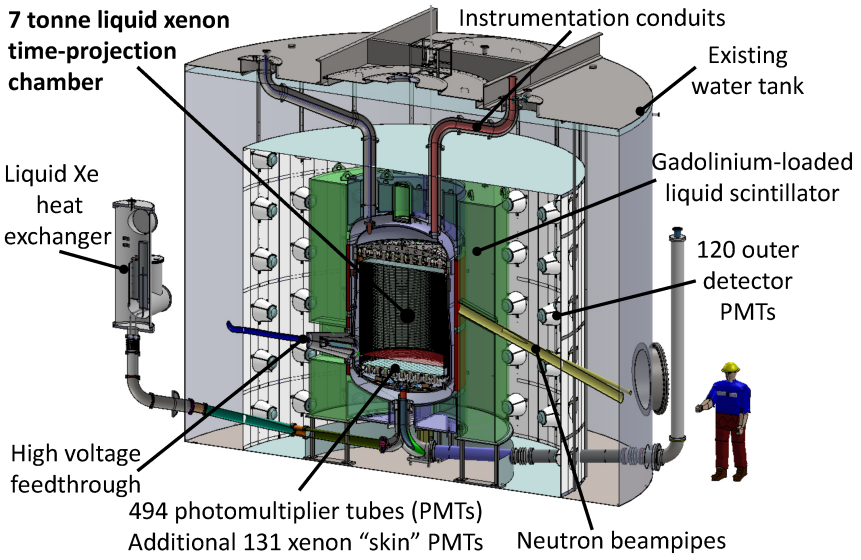



Figure 2.1: A schematic of the LZ detector configuration. Figure from [12].

The detector consists of a two-phase Xe time-projection chamber (1.5 m tall and 1.5 m in diameter), surrounded by 494 photo-multiplier tubes which detect photons

produced in the detector. Xe makes a useful detector material due to its high atomic mass and interaction cross-section, as well as its lack of radioactive isotopes that could produce background events. The inner detector is then encased in several layers of protection from outside radiation, including a Gadolinium-loaded liquid scintillator veto, and a water tank. The outer detector layers serve the dual purpose of preventing external background radiation from reaching the detector volume, and of measuring the internal background of the detector. For instance, gamma radiation and neutrons produced inside the detector by the detector materials or PMTs could be misidentified as WIMP signals unless vetoed through detection in the outer detector. An electric field is applied across the TPC in order to drift any electrons produced in the LXe to the liquid surface [22].

For the purposes of calibration, an external source of neutrons can also be connected to the detector via an evacuated conduit which directs the beam of neutrons into the detector volume. A monoenergetic neutron source is placed in line with the conduit outside the detector, and is then directed through the surrounding water tank into the Xe volume, entering with a well-defined position and direction [31].

### 2.1.1 Signal Detection in LZ

The LZ detector relies on two signal channels for detecting interactions in the detector volume. When a particle collides with a Xe nucleus in the detector, the interaction can occur either with the nucleus of the Xe (nuclear recoil events), or with the electron cloud (electron recoil events). Nuclear recoil events can be modeled with simple "billiard ball" scattering kinematics as described in section 1.2, with the deposited energy $E_{NR}$ given by equation 1.5. This energy deposition results in the production of photons through the excitement and de-excitement (scintillation) of the Xe, and electrons through ionization. The photons are detected immediately by the PMTs, to generate the prompt S1 signal. The electrons drift upwards under an applied electric field, and are subsequently extracted into the gaseous Xe phase under a different electric field where they produce secondary S2 photons via electroluminescence.

The S1 signal locates a scatter event in time, because it is detected promptly by the PMTs. Fitting the spatial profile of the S2 signal provides the (x,y) position of the event to 0.6 $cm$ accuracy. The time delay between the detection of the S1 and S2 signals indicates the depth (z) of the scattering event within the detector volume to 0.1 $cm$ [22]. This process is shown schematically in Figure 2.2. Both the S1 and S2 signals are recorded as peaks of some area, which is then interpreted as a number of photons detected.

One complicating factor is that not all scintillation photons and ionization electrons produced during a scattering event are detected. Photons and electrons can be re-absorbed by Xe nuclei or absorbed by impurities in the LXe. In addition, the 11.8% light collection efficiency of the PMTs, while high compared with similar experiments, further limits signal detection. The number of quanta produced per keV of energy deposited, or the light yield $L_y$ and charge yield $Q_y$, are thus larger than the number of quanta detected by the gain factors $g_1$ and $g_2$ respectively [31]. Another complication arises from the inherent quantization in the signals (fractional photons or electrons are obviously not possible), which means that even though the nuclear recoil energy can take on a continuum of values, the recorded signal is

quantized and subject to Poisson fluctuations.

In addition, it is important to distinguish electron recoil events from nuclear recoil events, because WIMPs will produce nuclear recoils while several background sources (such as gamma and x-ray radiation) will produce electron recoils. This can be achieved because the amount of energy detected in an electron recoil $keV_{ee}$, per keV deposited in the event, is not the same as the amount of energy detected for a nuclear recoil $keV_{NR}$ per keV deposited, and the values differ by the quenching factor QF. Because QF is different between signal channels (S1 and S2 above), measuring in multiple channels allows for discrimination between electron and nuclear recoil events [16].

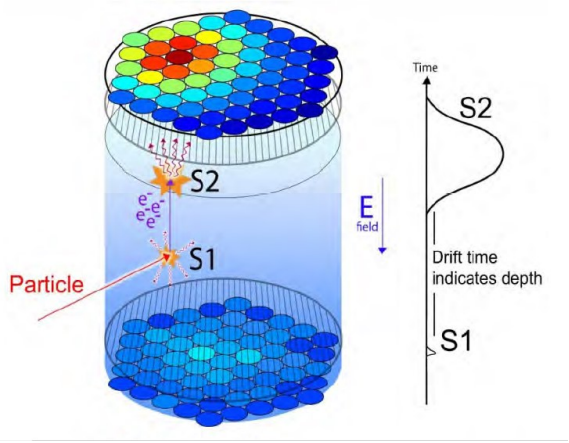



Figure 2.2: A schematic representation of the two-phase Xe TPC used in the LZ experiment. The photons produced during an interaction in the Xe are detected as a prompt S1 light signal by the PMTs; electrons drift to the surface under an applied electric field, are extracted into the gaseous phase, and produce a delayed S2 light signal through electroluminescence. Figure from [22].

## 2.2 Past Calibrations

In order to understand the response of the LZ detector to an observed WIMP-Xe interaction, it is extremely important to calibrate the expected S1 and S2 signals detected from a scattering event depositing a given energy. We first need to know the relationship between the number of photons and electrons produced and the energy deposited, and so determining the number of photons and electrons produced per keV deposited ($L_y$ and $Q_y$) is critical to interpreting the results of the LZ experiment. The relationship between the number of photons and electrons produced and the number detected (in the S1 and S2 signals) must also be understood. Several calibration strategies have been developed in past works.

In relative calibration methods, the S1 and S2 signals per keV of energy deposited are measured relative to a standard electron recoil source. Light and charge are then reported relative to the light and charge produced by this source without the need to know the gain factors g1 and g2. Absolute calibration techniques use a directed beam of incident particles or a source attached to the side of the detector to calibrate the detector response in quanta per keV deposited directly using the known g1 and g2 values. The use of a mono-energetic beam of neutrons in particular

has the benefit of allowing for calibration on an event-by-event basis as the incident particle energy is well-known. However, these techniques introduce several difficulties compared with direct calibration techniques. Most importantly, even when the incident neutron energy is well-known, the range of possible scattering angles means that only the maximum energy deposition is known, and any given scatter could produce a deposited energy between zero and this endpoint value depending on the angle of scattering, as shown in figure 3.2.

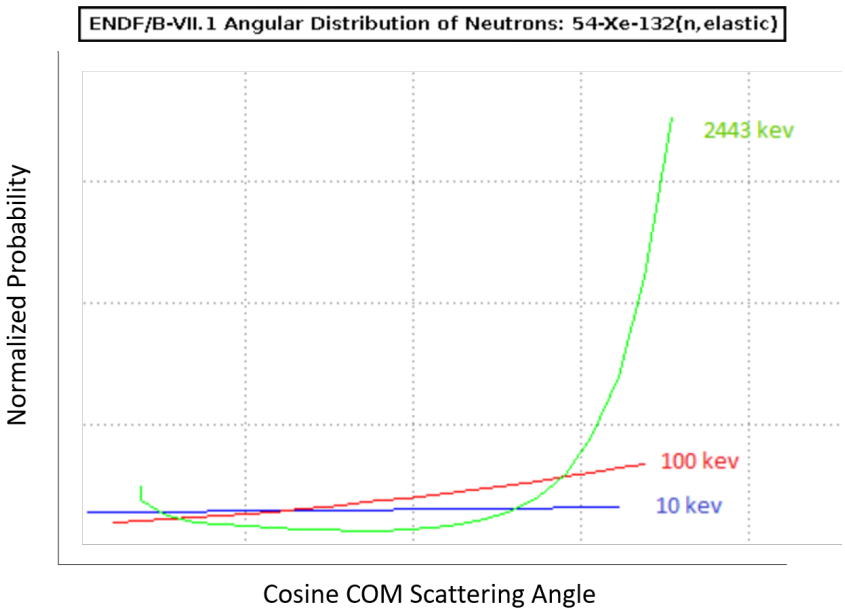



Figure 2.3: Scattering probability vs COM scattering angle on Xenon, for a few incident energies. At a given incident energy, a neutron can deposited only up to 3% of its total energy, and the likelihood of a given fractional energy deposition depends on the incident energy. Figure from National Nuclear Data Center.

In *ex situ* calibrations, a monoenergetic beam of neutrons is directed into a small Xenon-filled test cell. A second, far away detector is used to flag valid events, and the geometry of the source and detector are designed such that the scattering angle of the neutrons, and therefore the deposited energy, is known [11]. Several complications arise with *ex situ* calibration techniques. For instance, neutrons can interact with detector materials either prior to or after scattering in the Xe, which results in a loss of some unknown amount of energy and thus an inaccurate determination of the scattering angle and energy deposition. The dimensions of *ex situ* detectors are much smaller than those of the LZ detector, and put artificial limits on the range of possible scattering angles. In addition, neutrons which scatter multiple times within the test cell introduce further uncertainties by increasing the total light and charge produced over all scatters, and eliminating the simple relationship between scattering angle and observed neutron path [31].

In contrast, *in situ* calibrations are conducted in the LZ detector itself, and offer solutions to the above problems as well as certain additional benefits. In one form of *in situ* calibration, a source which produces a continuum of neutron energies is placed near the detector, and the signal spectrum is measured. Simulations are used to match the observed signal spectrum to the most likely energy deposition spectrum and set of $L_y$ and $Q_y$ values. In addition, the fact that endpoint in energy deposition

(the maximum energy deposited) is known for a given incident energy can be used to map the maximum observed S1 and S2 values to the theoretical endpoint in energy deposition (0.03 $E_{NR}$), thus providing an additional definite calibration data point at the expense of requiring a high number of incident neutrons for meaningful statistics.

In another form of *in situ* calibration, a monoenergetic beam of neutrons is directed into the detector volume via a conduit. Neutrons, having a vastly higher cross-section for interaction than WIMPS and thus a much shorter mean free path (on the order of tens of cm, depending on incident energy), often scatter multiple times within the detector. For neutrons which scatter exactly twice, the S1 and S2 signals can be used to determine the positions of each scatter vertex in the detector [32]. The relative positions of the scatters determines an angle of scattering from the first vertex, which implies an deposited energy by eq. 1.5 given the well-defined incident neutron energy. If this measurement is made for a large number of deposited energies, a comparison between energy deposited and observed S2 determines the shape of the $Q_y$ curves as a function of $E_{NR}$ over the given range of deposited energies. Then, $L_y$ can be calibrated using the energies determined from the (now calibrated) S2 signals.

Past calibrations have been conducted for LUX (which has a similar design to LZ), but LZ is, as of this writing, still under construction. Thus, though the above described methods have been defined for LZ, no calibrations have been carried out.

Past calibration strategies have been limited by the need to constrain the scattering angle, and by very low light yields at low energies. In fact, it is common to observe zero photons for a scatter event depositing only several keV of energy. In addition, in using only single or double scatter events, we discard many potentially useful events. The goal of this work is then to address these issues and expand on past methods by evaluating methods for calibrating in the LZ detector at low energies using neutrons which scatter multiple times within the detector.

# Chapter 3

# Evaluation of Calibration Methods

In an LZ direct detection experiment, all information about a potential interaction comes from the two signal sources described in chapter 2: prompt scintillation light (S1) and secondary light from ionization electrons (S2). In order to interpret an observed set of S1 and S2 signals and learn about the particle involved in the interaction, it is therefore vital to have a clear understanding of what light and charge signals are expected to be produced for a given energy deposited in a nuclear recoil with the Xe atoms. Thus, calibration experiments seek to define the relationship which translates between deposited energy ($E_{nr}$) and the light and charge yield so that, in a subsequent experiment, the amount of energy deposited can be inferred from the number of photons and electrons detected. [1]

Existing calibration methods are discussed in the previous chapter. These techniques provide useful calibrations for high incident energies. However, because the WIMP mass is not yet known, calibrating at low energies is also important. For low energies ($< 1\ keV$ deposited) where the numbers of photons and electrons produced are small and detection efficiencies are low, the light yield, and to a lesser extent the charge yield, becomes so low that this type of calibration is impractical. In this low energy regime, a new calibration method which leverages the fact that low energy neutrons can scatter multiple times within the detector is necessary. Even if an individual scatter is unlikely to produce a large enough light signal to be detected, the set of multiple scatters could. Thus, the goal of this work is to study calibration methods with low energy multi-scattered neutrons, and test whether such methods would be viable for the LZ detector.

In this work, we consider neutrons with incident energy of 40 kev. In addition, we consider 20000 incident events. This number was chosen because it is on the order of the number that would be expected during a standard length calibration, if we take the calibration length to be $\sim$1 week. Using more or less neutrons has the expected effect of increasing or decreasing the statistics, and thus decreasing or increasing the expected statistical errors in a calibration.

[1] It is important to note the distinction between light and charge yield (denoted $L_y$ and $Q_y$), and S1 and S2 signal. $L_y$ and $Q_y$ are properties of the Xe itself and represent the number of photons and electrons respectively produced per keV of energy deposited during an interaction. This is not the same as the number of photons and electrons which are detected as S1 and S2 signal, due to the less than perfect detection efficiencies. Detection efficiency is near 100% for electrons near the top of the detector, but decreases with depth due to greater chance of absorption during travel to the surface, and is significantly smaller for photons in general than for electrons.

## 3.1 Multi-scatter Neutron Calibration

Two approaches were defined and evaluated. If the scatter vertices can be ordered in time, the path of the neutron through the Xe implies a certain scattering angle at each vertex. This set of angles then implies a set of energy depositions, which can be mapped to the observed S1 and S2 signals, after a correction for the photon and electron detection efficiency. There are then two ways to proceed:

1. Calibrating on a per-neutron basis by averaging the observed signal over all vertices for a given neutron, and mapping that to the average energy deposition for that neutron.

2. Calibrating on a per-vertex basis by mapping the energy deposited at each vertex to the observed signal for that vertex.

The first method fully leverages the fact that even if a single scatter is unlikely to produce a detectable signal, the signal from multiple scatters might be measured and then divided by the number of scatters to find an average light and charge yield. Because of the much higher S2 detection efficiency, as long as some light is detected to verify that a true event has occurred, the S2 signal can be used to determine the number of scatters. The incident neutron energy remains approximately constant during scattering due to the small energy deposited (compared to incident energy) at each scatter, and so the total light and charge detected can then be simply divided by this number to determine an average. This method is less sensitive to errors in determining the energy deposited at each vertex, especially if they are correlated between vertices, but results in losing significant statistics by taking only one data point per neutron. The second method is more sensitive to errors in reconstructing the energy deposition at each vertex or in measuring the resulting signals, but offers a gain in statistics from using a data point for every scatter vertex individually.

Both of these methods require reconstructing the chronological order of scatters. The S1 signals are time-ordered as observed, but can't be easily matched to a vertex positionally. In contrast the S2 signals can be matched to vertices positionally but not easily ordered in time because the liquid electron travel time uncertainty is greater than the expected time separation between scatters. Thus, the goal of reconstructing the order of scatterings is twofold: to determine the energy deposited at each vertex from the implied geometry of the neutron's path (angle-reconstructed energy, or ARE), which is less sensitive to errors which arise from measuring the deposited energy from S2 directly, and to be able to map S1 signals to vertices and thus perform a $L_y$ calibration. This will be discussed further in section 3.4.

## 3.2 Simulations and Models

Several models were utilized to run and interpret scattering simulations. Geant4 was used with the BACCARAT package to simulate the firing of neutrons into the LZ detector volume and to produce the positions and energy depositions at each simulated scatter event. NEST was used to convert from energy deposition values into simulated detected light and charge signals.

### 3.2.1 Geant4 and BACCARAT

Geant4 is a model which simulates the interactions of particles with matter under various physics conditions. It is used in many areas of physics where simulations can help inform and interpret experimental work. Geant4 is especially useful because of its ability to treat many types of physical processes, materials and particle types, and a wide range of energies, and to do so in accordance with the most recent and complete experimental data available [3]. Given an initial problem definition, in this case a beam of low-energy mono-energetic neutrons fired in a conic beam into a cylindrical tank of liquid Xe, Geant4 can simulate the various types of interactions that may occur, and the results of those interactions, under the most accurate physics currently known. For this work, the most relevant outputs were the positions of each scatter event and the energy of the neutron before and after each scatter.

BACCARAT is package which interfaces directly with Geant4 in order to better simulate the physics relevant to low-energy interactions in liquid Xenon. A text file provides inputs for each simulation run, including the detector, a source geometry, the type of source, and the energy of the generated particles. BACCARAT produces outputs which can then be processed to access a wide variety of information about each run.

### 3.2.2 NEST

In order to translate from the energy deposited at each scatter as output by Geant4 and BACCARAT into the expected S1 and S2 signals observed, one must take into account the photon and electron production rates per keV deposited, the efficiency by which electrons are extracted to the surface and converted into photoelectrons, the efficiency with which photons are detected by the PMTs, and the position of the event within the detector. NEST (the Noble Element Simulation Technique), with the Python wrapper Nestpy, was used to model these processes. NEST relies on a large body of experimental results which determine the scintillation light and ionization charge yield of noble elements such as Xe under a given deposited energy, recoil type, and electric field [28]. In this work, NEST was used to convert the simulated deposited energies at each scatter vertex into an expected S2 signal in order to add anticipated uncertainties onto the simulated energy depositions. A new branch of the existing code repository was added in order to output nonzero numbers of electrons extracted for small energy depositions.

The charge and light yield curves produced by NEST were considered the baseline for comparison in this work, and the goal of conducting a calibration would be to improve these curves at low energies.

## 3.3 Behavior of Low-Energy Neutrons in Xe

Prior to studying new methods of calibration at low energies, it was important to evaluate the expected signals at low energies, and to develop an understanding for the behavior of low-energy neutron interactions with Xe.

Interactions between two particles are characterized by a cross-section for interaction, as described in section 1.2. The mean free path (MFP) of a neutron in liquid Xe can be determined analytically from this energy-dependent cross section $\sigma$ and

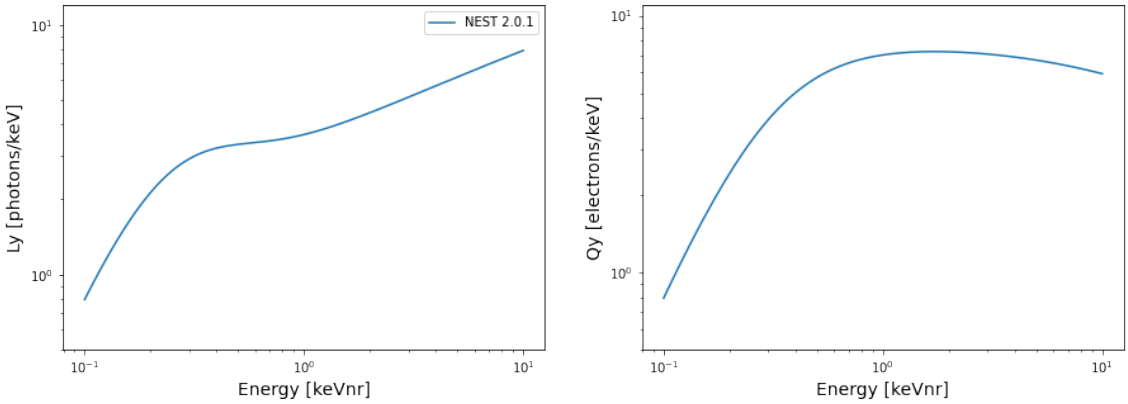



Figure 3.1: Light and charge yield curves as a function of nuclear recoil energy deposited in LXe, generated from NEST.

the density of the Xe $n$ as $MFP = 1/\sigma n$. A first check on the simulations was then to determine numerically the neutron MFP from the slope of the histogram of the distance to each first scatter at several incident energies. This is shown for several selected energies in table 3.1, and the values reported agree well with expectation.

In addition, the likelihood of scattering as a function of scattering angle was also determined from the simulations and verified to match the expected angle-dependent cross section for interaction as reported by the NNDC, as shown in figure 2.3.

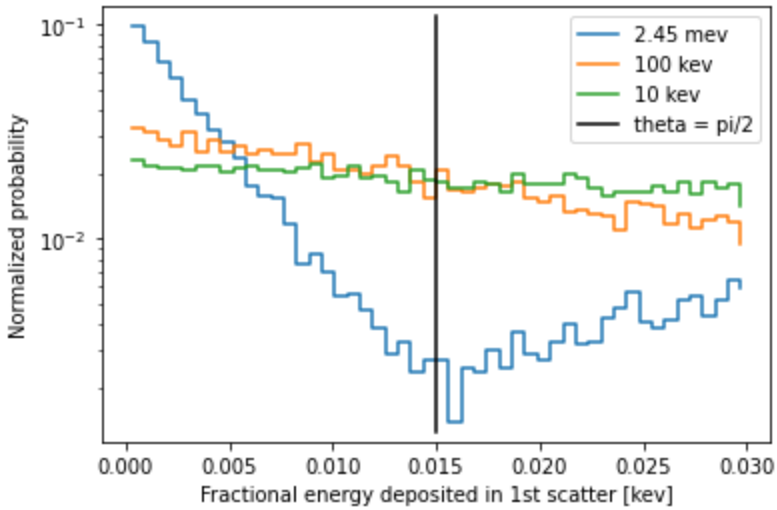



Figure 3.2: Scattering probability vs COM scattering angle, for a few incident energies. This agrees well with the NNDC reported values.

Another important initial consideration was what subset of the incident neutrons should be expected to be detectable at low energies, and whether the scatters should be expected to be separable in time from their S1 and S2 signals. This was important in order to inform the energy range of interest for a calibration, and for characterizing the expected signals.

Two scatters were considered to be time-separable if they were separated sufficiently in time such that their S1 signals, if nonzero, could be resolved as individual events. Because both S1 and S2 signals are detected as pulses with some width and

peak, if two or more occur very close together in time, the signal will be degenerate with that from fewer, higher energy, events. The threshold for time-separability in the LZ detector is approximately 40 $ns$, and can be posed as a distance threshold for a given incident neutron energy (and thus a given velocity). For 100 $keV$, 40 $keV$, and 10 $keV$ incident neutrons, this corresponds 17 $cm$, 11 $cm$, and 5.4 $cm$ respectively. The fraction of neutrons with all scatters time-separable decreases with increasing energy, and is small over the entire energy range of interest, as shown in table 3.1. Methods for distinguishing signals from multiple scatters which are not fully time-separable, such as pulse shape discrimination, have been developed in previous works, for instance [2].

Detectability was defined as whether a signal would be observed for a given event. Multiple detectability thresholds could be defined. For this work, an event was considered to be detectable if, summed over all scatters, at least two photons and two electrons would be expected to be observed. This is strongly limited by the number of photons detected, as electrons are detected much more efficiently. Various detection efficiency curves are depicted in 3.3, with each curve representing a different threshold for detectability. Detectability increases with incident energy due to the greater energy deposited in each scatter. Table 3.1 highlights the trade-off between separability and detectability with incident neutron energy. In addition, detectability increases with the number of scatters due to the greater total energy deposition while the chance of all scatters being separable decreases, exacerbating the anticorrelation between separability and detectability.

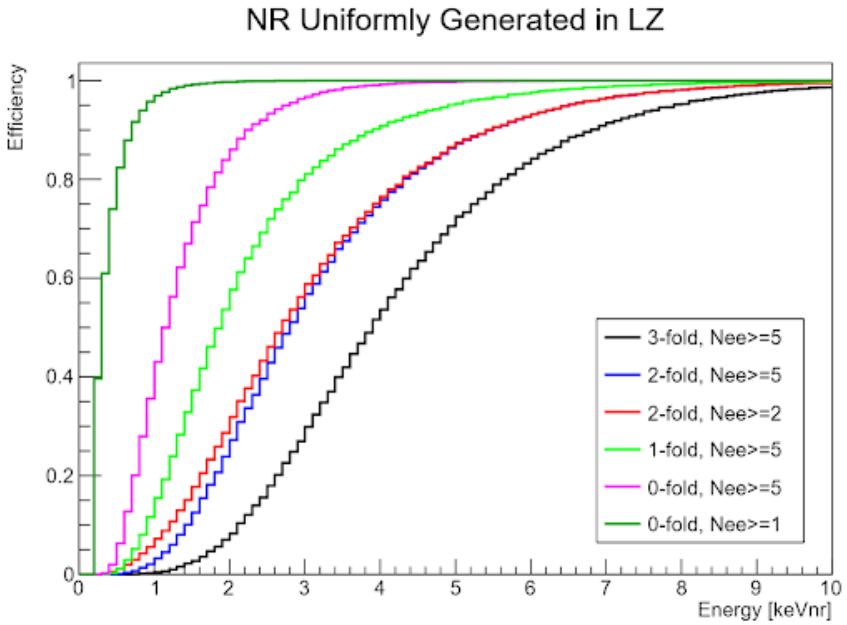



Figure 3.3: Detection efficiency as a function of deposited energy, under various detection threshold conditions. Figure courtesy of Xin Xiang.

| | MFP [cm] | % neutrons with all scatters separable | % neutrons detectable |
|---|---|---|---|
| 40 keV | 11.8 | 7 | 16 |
| 60 keV | 12.5 | 6 | 22 |
| 80keV | 12.6 | 5 | 26 |
| 100keV | 12.9 | 4 | 28 |

Table 3.1: Mean free path and measures of separability and detectability as a function of incident neutron energy. The fraction of neutrons with all scatters separable decreases with energy, while the fraction which are detectable increases.

## 3.4 Order Reconstruction Algorithm

### 3.4.1 Algorithm Definition

A major part of this work was the development of an algorithm to reconstruct the chronological ordering of multi-scatter events from observable quantities. In a real experiment, the observed S1 and S2 signals allow for a determination of the position of each scatter vertex inside the detector, within an uncertainty, and of the energy deposited at each vertex from the observed S2, again with an uncertainty resulting from fluctuations in the quantized number of electrons detected per keV deposited. As mentioned above, the chronological ordering of the scatters cannot necessarily be determined directly from the S1 and S2 signals. For the S2's, this is because the uncertainty in the travel time of electrons to reach the top of the detector is greater than the time separation between scatters. The S1 signals are often overlapping, and even if they are all time-separable for a given set of scatters, the position from which a given S1 originated is not determined from the signal itself. However, knowing the ordering of scatters would mean knowing the order of the S2 signals, and thus the order of energy depositions within the detector. This would allow for the association of deposited energies to observed S1 signals, which is the basis for a light yield calibration. Difficulties arise in this last step, as the S1 signals are often zero for low incident neutron energies and are often not time-separable.

A very simple, though ultimately not very accurate, method of reconstructing the ordering of scatter vertices is to assume given an initial first scatter, that the next scatter chronologically is the closest scatter spatially. This is motivated by the idea that given a set of possible subsequent scatter positions, it is more likely that the neutron used more steps to travel farther and fewer steps to travel less far, and thus that the closest position corresponds to the first step away from the current position. However, what this method offers in simplicity it lacks in accuracy of the final reconstruction, as very often the paths of neutrons through the detector are non-linear, with some scatters bringing the neutron back towards its entry position.

Improving on this method, an algorithm could be defined to choose, starting from a known first position, each subsequent position based on some probability of it being the next scatter. Although distance could factor into this probability, the key piece of information comes from recognizing that deposited energy, as reconstructed from the S2 signals, implies a certain scattering angle at each vertex by equation 1.5. Thus, from a given position each potential subsequent position can be assigned a likelihood based on how consistent the required scattering angle is with the S2

signal produced from that vertex. Assuming the first position can be determined by requiring that it be within the projection of the beam pipe into the detector, the scatters could be ordered in this way. However, difficulty arises when two or more scatter positions are within uncertainty of each other, and this method becomes computationally expensive for higher numbers of scatters. In addition, this method is very sensitive to misidentifying one scatter position.

Instead, a different method was taken to reconstruct the ordering of the scatter vertices. For each set of $n$ scatter positions corresponding to the path of a neutron through the Xenon, there are $n!$ possible orderings. Each of these potential orderings implies a geometry and thus a set of angles, which themselves imply a set of energy depositions corresponding to each vertex. These angle-reconstructed energy depositions (ARE) can be compared to the set of S2-reconstructed energy depositions (S2E) by vertex, and the order for which the two sets are most consistent can be identified as the most likely ordering. Thus, the set of all possible orderings is defined, and the energy deposition implied by each is calculated and compared to the energies implied by the observed S2's. We require possible orderings to have the first scatter within the projection of the beam pipe into the LXe, to account for the fact that neutrons travel along straight paths until scattering, which in many cases reduces the total number of possible orderings to evaluate. Two different methods for evaluating the agreement between the angle-reconstructed and S2-reconstructed energies, and each is expected to perform well in different cases:

1. The root-mean-square (RMS) deviation between the angle-reconstructed and S2-reconstructed energy depositions is determined for each potential ordering, and the ordering with the lowest RMS deviation chosen. This favors small absolute differences between the two energy values, regardless of the values of the energies themselves or the uncertainty in that difference.

2. The number of sigma deviations (i.e., number of factors of the uncertainty) between the two energy values is determined at each vertex, and converted into a likelihood that that sigma difference would be observed. These likelihoods are combined over all scatters for a given order, and the order with the highest overall likelihood is chosen. This favors orderings for which the uncertainty in the energy differences is large compared with the difference.

Both methods were implemented, motivated by the fact that each should fail in different cases, but the overall accuracy of the algorithm in choosing the correct ordering was not significantly dependent on the chosen method. That these methods of choosing the correct order give similar results supports the fact that difficulties in choosing the correct ordering come from the uncertainties inherent in the data, not from choice of method itself.

## 3.5 Treatment of Experimental Uncertainties

Simulated positions and recoil energies as produced by Geant4/BACCARAT do not include uncertainties. In real experiments however, the positions and deposited energies are not known exactly due to various sources of experimental uncertainties, and these uncertainties will lead to incorrect order reconstructions and to errors in reconstructing the energy depositions from the scattering geometries. In fact, if

all uncertainties are turned off, the reconstruction algorithm has an overall correct reconstruction rate of $> 99\%$. The small fraction of incorrect orderings when no uncertainties are included can be explained by the small spread in mass of the Xe isotopes which is accounted for in the simulations but not in the reconstruction algorithm. In position, the uncertainty is greater in the x and y direction (0.6 cm uncertainty) than in z (0.1 cm uncertainty) [22]. This leads to an uncertainty in the angles reconstructed from the measured positions. Positional uncertainties are incorporated into the algorithm by adding to the simulated positions an additional term picked randomly from a Normal distribution with standard deviation equal to the positional uncertainty.

There is also an uncertainty in the S2-reconstructed energy depositions which arises from the fact that S2 signal is measured, rather than recoil energy directly (simulated recoil energies are referred to as "true" energies, or TE, in this work). The quantization in number of electrons detected introduces an uncertainty in the energy deposition inferred from the measured S2, as does the statistical counting uncertainty arising from Poisson fluctuations in the number of electrons observed. This is incorporated into the algorithm through the use of NEST to convert simulated energy deposition into a number of electrons produced which is subject to uncertainty, and then to convert this number back into an energy value. The ultimate effect is a $\sqrt{n}$ uncertainty in number of electrons, which means the distribution of observed electrons for a given true value is Poisson.

The energy error between the true and angle-reconstructed energies is generally dominated by the positional uncertainties unless the order is chosen incorrectly. This is because even though the angle-reconstructed energy values are selected based on the best match to the S2-reconstructed energies (rather than the "true" simulated energy values), if the order is chosen correctly the angle-reconstructed energies will actually better match the true values than the S2-reconstructed values. Thus, the energy uncertainty only indirectly affects the final energy error by producing incorrect orderings.

## 3.6 Calibration Method Evaluation

### 3.6.1 Metric of Confidence

Before beginning to characterize the accuracy of the reconstruction algorithm and define cuts on the data to improve accuracy, it was necessary to define a metric which provides a measure of how likely a particular reconstruction is to be correct. This metric must be definable from purely observable quantities, rather than simulation data that would not be available in a real experiment. The metric of confidence was designed to be a predictive measure for reconstruction accuracy on which cuts could potentially be applied to exclude neutrons whose reconstructions are less likely to be correct.

For the first method of order choosing described above (which was considered the standard method for this work), the metric of confidence was defined to be the difference between the RMS deviation of the S2-reconstructed energies (S2E) to angle-reconstructed energies for the picked order and the second best order as

$$MOC \equiv \frac{RMS(ARE - S2E)_{chosen\ order}}{RMS(ARE - S2E)_{next-best\ order}} \tag{3.1}$$

For the second method of order picking, the metric of confidence was defined to be difference between the likelihoods associated with the picked and next best orderings.

For the first definition (used for the remainder of this work), it is important to note that a metric of confidence closer to zero indicates greater confidence in the correct ordering having been chosen.

### 3.6.2 Preliminary Cuts

Several preliminary cuts had to be applied to the simulated data before an order reconstruction could be attempted. First, any neutron that experienced an inelastic scatter event was discarded, because the signal from inelastic events would obscure the desired signal from elastic events. This is a very small effect for low energy incident neutrons, with $< 0.1\%$ of 40 $keV$ neutrons undergoing an inelastic event.

Next, the location of each scatter position was determined relative to a conic projection of the beam pipe into the detector, and any neutron having no scatters within the beam pipe projection was discarded. The true first scatter position must occur within the beam pipe projection, and so a reconstruction could not be attempted without at least one scatter meeting this criteria. In a real experiment, this could result from the neutron interacting with the detector materials before entering the Xe. In the simulations, this should not occur except due to the addition of positional uncertainties as described above, and this is indeed the effect observed. $< 2\%$ of neutrons had their true first scatter outside of the beam pipe projection, and $< 1\%$ of neutrons had no scatters within the beam pipe projection and were thus discarded.

Finally, no reconstruction was needed for neutrons with fewer than two scatters, and an upper limit on the number of scatters was also imposed to limit the run-time of the algorithm. Though some neutrons can scatter up to tens of times before exiting the detector, the number of events with a given number of scatters decreases sharply with increasing number of scatters (see table 3.2).

### 3.6.3 Failure Modes and Accuracy Cuts

A major part of this work was the evaluation of the reconstruction algorithm, both in terms of the fraction of events correctly ordered and the resulting discrepancy between angle-reconstructed and true energy depositions, and the development of cuts which could improve the overall accuracy of the algorithm.

The algorithm's accuracy was first defined in terms of the fraction of neutrons whose order was correctly reconstructed, of the neutrons not discarded in the preliminary cuts described above. As shown in table 3.2, ordering accuracy drops off with increasing number of scatters $n$ due to the increased complexity and greater number of possible incorrect reconstructions [2].

The simplest case, with n=2, was first considered in isolation in order to better understand the cases in which the algorithm fails. The n=2 case is particularly

[2]It is important to realize that the "worst case" accuracy, and thus the benchmark for comparison, is dependent on the number of scatters $n$. In no case should 0% accuracy be expected, because even if the algorithm were no better than random guessing, there is a $1/n!$ chance of choosing the correct ordering. If an accuracy lower than this were observed, it would imply the algorithm was correctly identifying the more likely orderings, and then preferentially choosing incorrect orderings.

simple because only two orderings are possible, and each is characterized by a single angle-reconstructed energy (no angle can be reconstructed for the last scatter). Several failure modes were identified. The simplest failure case occurs when the true first scatter is not located within the beam pipe projection, and so the correct ordering is never considered so that correct reconstruction can never be produced. This accounts for $5.1 \pm 1.1\%$ of the incorrect n=2 reconstructions, and no cut can be applied to remove this effect. Another possible failure case occurs when the two scatters interchange positions under the addition of positional uncertainties. Although the nature of the positional uncertainty means there is no distinct threshold at which this is guaranteed to occur, a cutoff could be placed such that when two scatters are within positional uncertainty of each other their order reconstruction is unlikely to be correct. $3.8 \pm 0.05\%$ of n=2 neutrons meet this condition and were incorrectly reconstructed.

The energy uncertainties also play an important role in motivating data cuts. One possible class of failures occurs when the true geometry is such that the true exit angle is close to the reconstructed entry angle of the order which is exactly the reverse of the true ordering. In other words, the energy deposited at the last scatter and the angle implied if it were instead the first scatter are consistent with each other within uncertainty. For n=2, this is accounts for $12 \pm 1.1\%$ of incorrect reconstructions.

These failure modes can then be extrapolated to higher $n$, although the increasing number of possible geometries makes these cases more complex. Failure can occur when the true first scatter is not in the beam pipe projection, when two vertices are close enough that positional uncertainty allows the angle reconstruction to interchange them, and when the true exit angle matches the implied entrance angle if the path were traversed in precisely the reverse of the true ordering.

This analysis was the basis for the definition and application of several cuts designed to improve the accuracy of the reconstruction algorithm. First, a cut can be defined on the metric of confidence (MOC) as defined above. Initially, based on the MOC value at which the overall ordering accuracy begins to drop off more rapidly, a threshold of MOC $< 0.3$ was applied. In addition, to address the failures which result primarily from the interchange of angles under the addition of positional uncertainties, a cut on the minimum between points was also initially applied. Because the dominant positional uncertainty is in the xy-plane rather than in z, the threshold was set to cut neutrons for which the minimum separation between scatter positions projected in the xy-plane was greater than 0.6 cm. The acceptance of these cuts, and their impact on the reconstruction accuracy, is shown in table 3.2. It is clear that although this set of cuts improves the fraction of events correctly ordered, this comes at a very high cost in acceptance.

This set of initial cuts was defined based on the ordering accuracy, and further analysis described in the next section reveals that they do not in fact optimize the final energy reconstruction. A more rigorous statistical analysis was then carried out to determine the optimal set of cuts.

More important than ordering accuracy is whether the energy depositions were correctly recovered. While this is certainly correlated to the order being correctly reconstructed, there are cases in which the order could be incorrect while the implied energies are still close to the true energies. The reconstruction algorithm picks the ordering with the smallest RMS deviation between the S2-reconstructed and

| n | **Pre-cut % of neutrons** | **Pre-cut % correctly ordered** | **Acceptance of cuts (%)** | **Post-cut % correctly ordered** |
|---|---|---|---|---|
| **all** | 100 | 57 | 31 | 91 |
| **2** | 33 | 85 | 70 | 92 |
| **3** | 23 | 64 | 30 | 88 |
| **4** | 16 | 40 | 7 | 82 |
| **5** | 12 | 31 | 2 | 75 |
| **6** | 9 | 28 | <1 | 80 |

Table 3.2: Accuracy (fraction correctly ordered) and acceptance of cuts as a function of multiplicity n, under and initial set of cuts based on order-reconstruction accuracy. Note that all percentages are referenced relative to all neutrons for which a reconstructed was attempted (i.e., which passed the preliminary set of cuts).

angle-reconstructed energies. This means even if positional uncertainty produces an incorrect ordering, the chosen energy values might still be fairly accurate. In addition, even though the algorithm uses only the S2-reconstructed energy to pick the ordering, the angles should, except for positional uncertainty, match the true energies rather than the S2-reconstructed energies once the correct order is chosen. Thus the energies implied by the set of angles should most closely match the true energy values. Thus, if the ordering is chosen correctly, the angle-reconstructed energies can actually better reproduce the true energies than they do the S2-reconstructed energies. If the ordering is chosen incorrectly, this does not necessarily mean the angle-reconstructed energies will be very far from the true energies because the chosen order was the one which best matched the S2-reconstructed energies. Thus if the S2-reconstructed energies were close to the true energies (as they should be, except for the large uncertainty in number of electrons produced), the overall energy reconstruction can still be reasonably accurate. For higher multiplicities, it is often still useful to rely on these energy-equivalent tracks even when the ordering is chosen incorrectly.

### 3.6.4 Statistical Data Cut Optimization

As the aggressiveness of data cuts is increased, there is inherently a trade-off between the post-cut accuracy (or how well the energies are reconstructed after the cuts are applied) and the acceptance (or what fraction of events are admitted through the cuts). The goal is then to find the optimal balance between these factors. In order to optimize the set of cuts, we define a final expected statistical error. This is a measure of the mean statistical error expected across the dataset for broad comparison between each calibration method. In the per-vertex case, this is defined as the per-vertex RMS error between the angle-reconstructed and true energies, divided by the total number of vertices used:

$$expected\ final\ error\ =\ \frac{RMS_{per-vertex}}{\sqrt{N}} \tag{3.2}$$

with

$$RMS_{per-vertex}\ =\ \sqrt{\frac{\sum_{i=1}^{N}\left(ARE_i\ -\ TE_i\right)^2}{N}} \tag{3.3}$$

where $N$ is the number of vertices used.

For the per-neutron case, this is defined as the per-neutron RMS error between the summed angle-reconstructed and summed true energies, divided the total number of neutrons used:

$$expected\ final\ error\ =\ \frac{RMS_{per-neutron}}{\sqrt{N}} \tag{3.4}$$

with

$$RMS_{per-neutron}\ =\ \sqrt{\frac{\sum_{i=1}^{N}\left(\sum_{j=i}^{n} ARE_i - \sum_{j=i}^{n} TE_i\right)^2}{N}} \tag{3.5}$$

where $N$ is the number of neutrons used and $n$ is the number of scatter vertices for the $i^{th}$ neutron.

For both methods, the metric of confidence (MOC) threshold cut was first optimized by iterating over a range of threshold values for each multiplicity, and calculating the expected statistical error at each value. In this way a set of MOC threshold values was defined corresponding to each multiplicity by determining the minimum of each curve in figure 3.4. Then, using the optimized set of MOC threshold cuts, a similar analysis was carried out to determine the optimum set of thresholds on minimum allowed xy-plane separation.

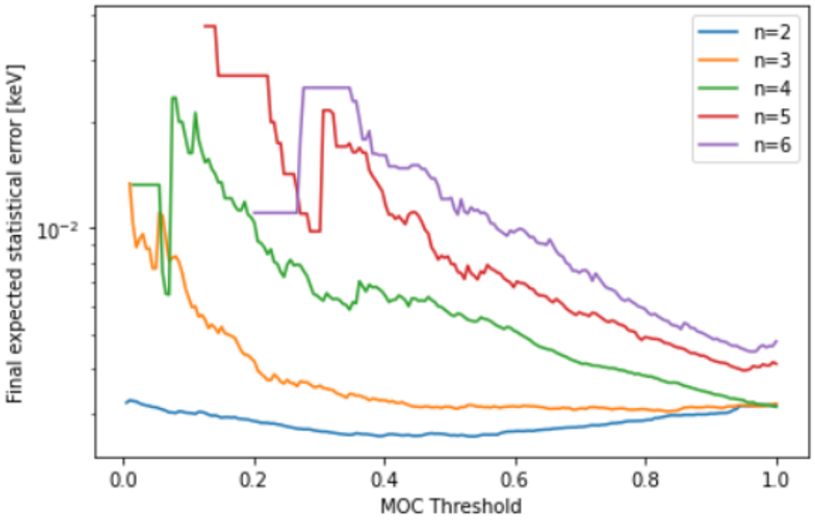



Figure 3.4: Optimization of $n$-dependent MOC threshold cuts for the per-vertex case. Similar plots were produced for the MOC threshold optimization for the per-neutron case, and for optimizing the threshold on minimum xy-plane separation for both cases.

The optimized set of cuts (shown in table 3.3) was ultimately determined to be quite minimal, especially at higher multiplicities. This highlights the trade-off between acceptance and accuracy: from a purely statistical standpoint, the set of cuts which minimizes the expected statistical error is very minimal at higher multiplicities because in that regime the gains in acceptance dominate the loss of accuracy. Under this set of cuts, the expected final statistical error is similar between the per-neutron ($0.00095\ keV$) and per-vertex ($0.0014\ keV$) cases, though somewhat higher in the later case.

| $n$ | Per-vertex | | Per-neutron | |
|---|---|---|---|---|
| | **Optimal MOC Threshold** | **Optimal Separation Threshold** | **Optimal MOC Threshold** | **Optimal Separation Threshold** |
| **2** | 0.54 | 4.00cm | 0.53 | 3.9cm |
| **3** | 0.82 | 0.34cm | 0.83 | 0.2cm |
| **4** | 0.99 | 0.04cm | 1.0 | 0.1cm |
| **5** | 0.99 | 0.08cm | 0.95 | 0.0cm |
| **6** | 1.00 | 0.00cm | 0.96 | 0.0cm |

Table 3.3: Optimized MOC and minimum xy-plane separation thresholds as a function of multiplicity.

# Chapter 4

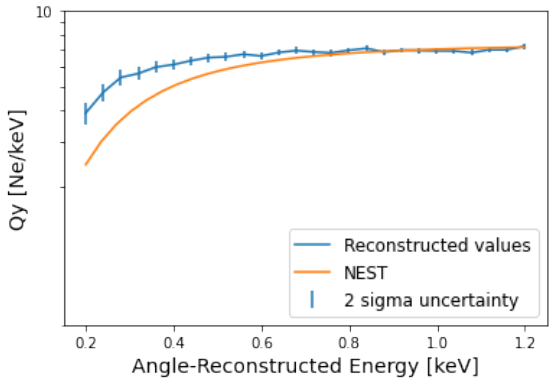
# Results and Analysis

In order to evaluate the expected performance of the two calibration methods described in the chapter 2, charge yield ($Q_y$) and light yield ($L_y$) curves were produced under each method from the energy values reconstructed from the simulations. These were then compared to the expected $Q_y$ and $L_y$ curves produced directly from NEST, in order to determine how closely the energy reconstruction algorithm reproduces the expected NEST values.

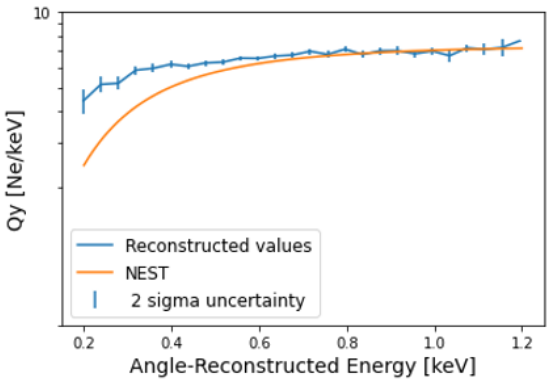
## 4.1 Expected Charge Yields

For the per-vertex case, charge yield was calculated for each vertex as the number of electrons produced divided by the angle-reconstructed energy. For the per-neutron case, the average charge yield was calculated as the average number of electrons produced divided by the average angle-reconstructed energy for each neutron across all vertices. Note that, as a simplification, the number of quanta produced as determined by NEST was used for this calculation. In a real calibration, this would be determined from the observed number of quanta, divided by the gain values. Charge yield was then binned by energy deposition, and the standard error of the mean calculated for each bin for a measure of the statistical error within each bin. In addition, the average fractional difference between the reconstructed curve and the NEST curve was also calculated as a metric of the systematic error. The average standard error of the mean was found to be 0.13 electrons/keV for the per-vertex and method 0.19 electrons/keV for the per-neutron method, and the average fractional difference to the NEST curves was found to be 0.15 times the curve values and 0.17 times the curve values in the per-vertex and per-neutron cases respectively.

The charge yield curve for the per-vertex and per-neutron methods are shown in figure 4.1. The reconstructed charge yield largely follows the shape of the expected curve, with small statistical errorbars. At higher deposited energies (greater than approximately 0.6 $keV$, the charge yield curve is approximately flat, and drops off steeply with decreasing deposited energies. However, there is a systematic offset which causes the reconstructed values to overestimate the expected values at low energy depositions. This offset stems from a systematic underestimation of the angle-reconstructed energy values at higher multiplicities for incorrectly ordered events, and can be accounted for in moving from the true values to the quantized observable quantities. The systematic errors are discussed further in section 4.3.

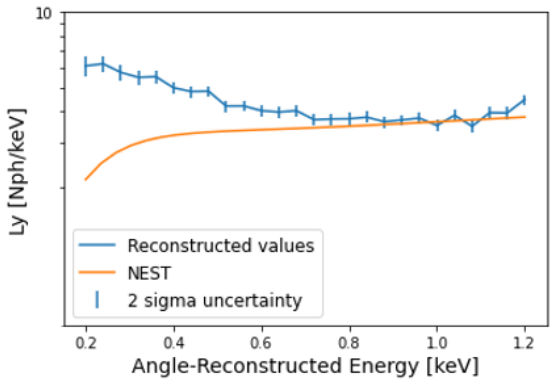



(a) Per-vertex $Q_y$

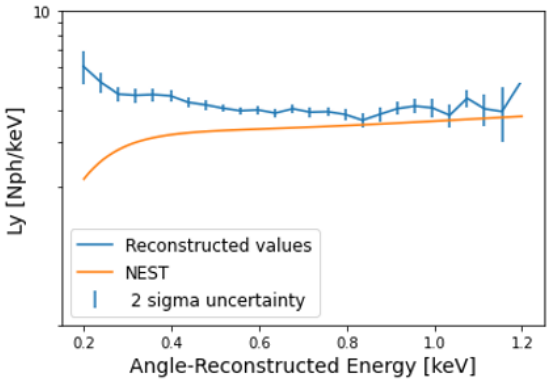



(b) Per-neutron $Q_y$

Figure 4.1: Reconstructed and expected charge yield curves for the per-vertex and per-neutron calibration methods.




(a) Per-vertex $L_y$




(b) Per-neutron $L_y$

Figure 4.2: Reconstructed and expected light yield curves for the per-vertex and per-neutron calibration methods.

## 4.2 Expected Light Yields

The light yield curves were calculated analogously to the charge yields. For the per-vertex case, light yield was calculated for each vertex as the number of photons produced divided by the angle-reconstructed energy. For the per-neutron case, the average light yield was calculated as the average number of photons produced divided by the average angle-reconstructed energy for each neutron. As for charge yield, light yield was then binned by energy deposition, and the standard error of the mean for each bin and average fractional difference to the NEST curves calculated as measures of the statistical error within each bin and the systematic errors respectively. The average standard error of the mean was found to be 0.18 photons/keV for the per-vertex method and 0.26 photons/keV for the per-neutron method, and the average fractional difference to the NEST curves was found to be 0.88 times the curve values and 0.87 times the curve values in the per-vertex and per-neutron cases respectively.

The light yield curve for the per-vertex and per-neutron methods are shown in figure 4.2. The reconstructed light yield curves show similar statistical errors as the charge yield curves, but with a much more pronounced systematic offset. As for the charge yields, the systematic errors are discussed further in section 4.3.

Note that both the reconstructed $Q_y$ and $L_y$ curves are produced using all neutrons which were not cut in the statistically optimized set of data cuts. However, in a real calibration there is the added complication that some neutrons may produce zero S1 signal total, which means the timing of a set of scatters, and thus depth of the scatters, is not easily determined. The primary effect of limiting to neutrons which produce an S1 signal is to decrease the overall statistics available, and thus lengthen the errorbars in the above plots, while largely maintaining the shape of the curves. This issue is discussed further in section 4.4, and a potential method of circumventing it is suggested in section 4.6.

## 4.3 Systematics in Reconstructed Charge and Light Yields

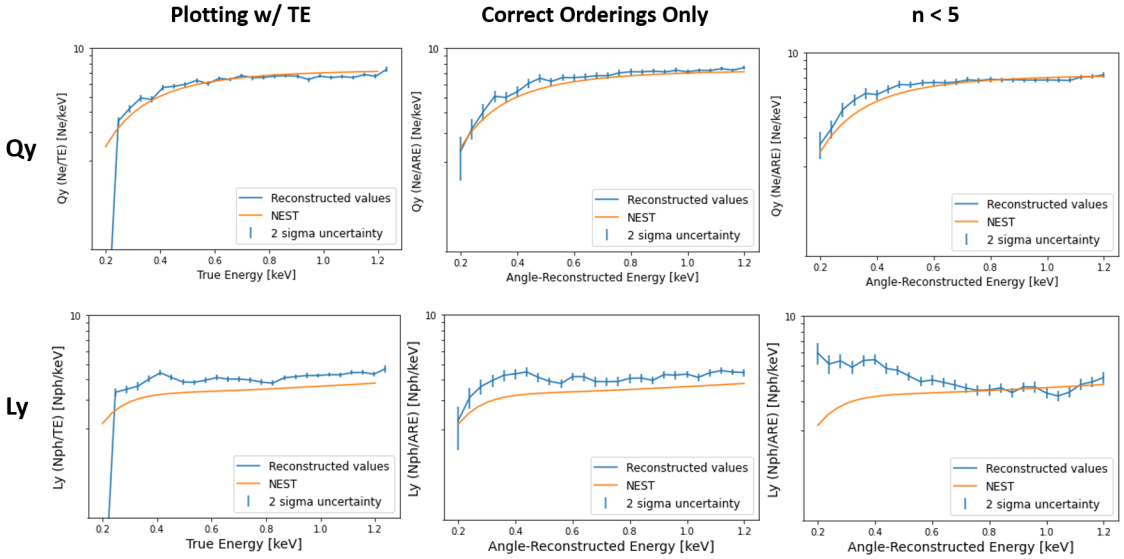
There is a systematic offset apparent in both the reconstructed $Q_y$ and $L_y$ curves for both the per-vertex and per-neutron methods. As mentioned above, this is partially caused by a systematic underestimation of the angle-reconstructed energies compared to the true energies for higher multiplicity events, especially in the case where the order is reconstructed incorrectly. One possible explanation for this comes from the quantization of the electron signal at low energies: for true energy deposition values $< 0.2\ keV$, zero electrons are produced and thus the S2-reconstructed energy is zero. Because the angle-reconstructed energies are chosen to best match the S2-reconstructed energies, this would lead to underestimation of angle-reconstructed energies especially when the order is incorrect (if order is correct, correct angles mean angle-reconstructed energy matches the true energies better than the S2-reconstructed energies). However, if this threshold on the number of electrons produced is removed from the NEST model, the systematic persists and thus this effect cannot explain the majority of the offset.

Other effects of the quantization and Poisson fluctuations in the observed electron and photon signals can account for the systematic. The systematic arises in the move from true quantities (i.e., the simulated values) to measured quantities, even when no order-reconstruction is attempted. This can be demonstrated in plotting the charge and light yield curves using the true rather than angle-reconstructed energy values, which reproduces the expected NEST curves much more closely (almost perfectly for $Q_y$, and somewhat less well for $L_y$, see left hand column of figure 4.3). Thus the systematic is introduced in moving to measured quantities from the true values, due to the inherent uncertainties in the quantization of those quantities.

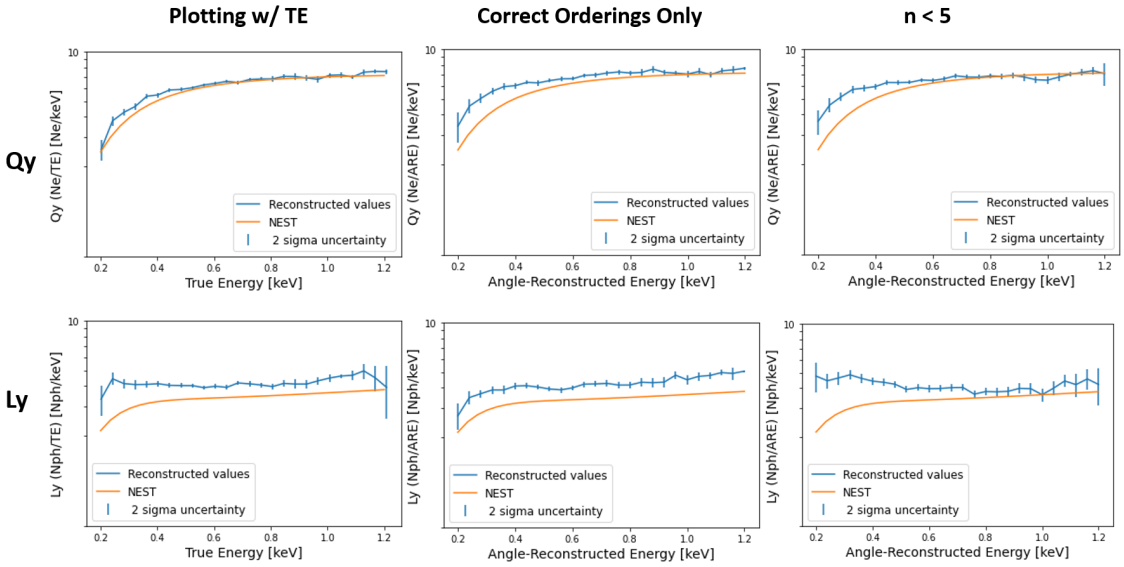
Enforcing the correct scatter orderings can also artificially remove much of the systematic (reduces the effect by factor of 1.8 for per-vertex case and 1.5 for per-neutron case). However, this is obviously not possible to enforce directly in a real calibration experiment. Instead, a harsh cut on the metric of confidence could be applied. The effect of this is shown in the middle column of figure 4.3. However, the cost of imposing an MOC threshold which eliminates the observed systematic is drastically reduced acceptance especially at higher multiplicities.

In addition, limiting to multiplicities of fewer than five also significantly reduces the systematics, by a factor of 2.0 in per-vertex case and 1.3 in per-neutron case. This can be directly cut for, though as for the MOC thresholds, more aggressive cuts leads to significant loss of statistics. The effect of limiting to $n < 5$ is shown in

the middle column of figure 4.3.




(a) Per-vertex Case




(b) Per-neutron Case

Figure 4.3: Per-vertex (a) and per-neutron (b) $Q_y$ and $L_y$ curves (top and bottom rows of respective sub-figures), under different cuts and plotting conditions in order to account for the source of the systematics.

## 4.4 Effect of Zero S1 Events

In the previous sections, we have analyzed the calibration techniques under the simplifying assumption that we are able to locate each set of scatter vertices in time using the S1 signal. This is helpful in directly ascertaining whether the NEST curves can be reproduced. However, as demonstrated in section 3.3, in a real calibration a significant fraction of neutrons may not produce enough photons to be detected in the S1 signal, due to the low detection efficiency for the photons produced. Though

utilizing multi-scattered neutrons increases the chance of detectability, many neutrons still produce zero S1. In general, a neutron producing zero S1 must be discarded from the calibration because the depths of the scatter vertices cannot be easily defined.

The primary effect of this is to significantly limit the available statistics. Overall, approximately one quarter of the neutrons which passed the above defined preliminary cuts produce nonzero S1 signals, which leaves about one third of the total vertices. As expected, the effect is much greater for lower multiplicity events, as each vertex provides an additional chance for photons to be detected. In fact, only $7 \pm 1$ % of $n = 2$ neutrons produce nonzero S1's, while $51 \pm 4$ % of $n = 6$ neutrons produce nonzero S1's. Because lower multiplicity events are then preferentially discarded by a cut on nonzero S1 signal, this exacerbates the observed systematic effect described above. The systematic comes predominantly from higher multiplicity events, which are more likely to produce nonzero S1 signals. The ultimate effect on the reconstructed charge and light yield plots is shown in figure 4.4.

This is currently a major limiting factor in performing any low-energy calibration, because of the significant reduction in available events. A possible method for determining the depth of scatter vertices under zero S1, which could mitigate this issue in many cases, involves using the H-reflector's timing tag to provide a time for the set of scatter vertices. This is briefly described in section 4.6.

## 4.5 Conclusions

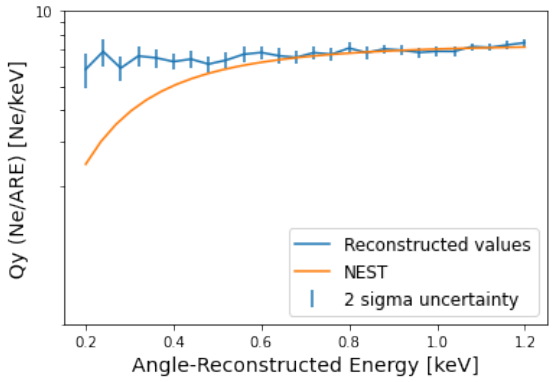
In this work, we have demonstrated the necessity of carefully calibrating the charge and light yields of interactions with Xe in order to understand the results of the LZ experiment. We have evaluated two methods of overcoming the issues of unknown scattering angle and low light yields at low energies by utilizing for the first time neutrons which scatter more than twice within the detector volume.

We have evaluated two different low-energy multi-scattered neutron calibration techniques: calibrating on a per-vertex and per-neutron basis.

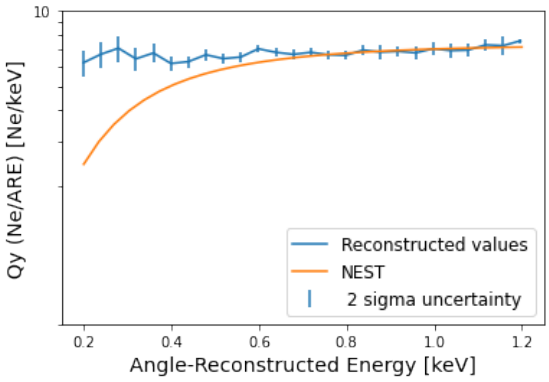
We have developed a set of cuts which optimize the expected statistical error between the angle-reconstructed and true energy depositions, and found the optimal cuts are minimal at higher multiplicities. This emphasizes the necessary trade-off between accuracy and acceptance: more aggressive cuts would of course produce lower energy error for each neutron or vertex, but the overall statistical error is dominated by the acceptance.The optimized set of cuts remains largely unchanged under the limitation to events which produce nonzero S1 signals.Both methods perform similarly from a statistical standpoint, though the per-neutron case results in a somewhat lower expected statistical error.

However, the reconstructed charge and light yield curve are dominated by the systematic errors. We have identified and accounted for sources of systematics which arise in moving from true values to measured quantities, due to the inherent uncertainties from the quantization of these quantities. Minimizing systematic errors requires much more aggressive cuts than the statistically optimized set, for instance limiting ourselves to fewer than 5 scatters, or imposing much stricter cuts on minimum allowed xy-plane separation or MOC. Limiting to nonzero S1 events increases the systematic due the higher acceptance of high multiplicity events.

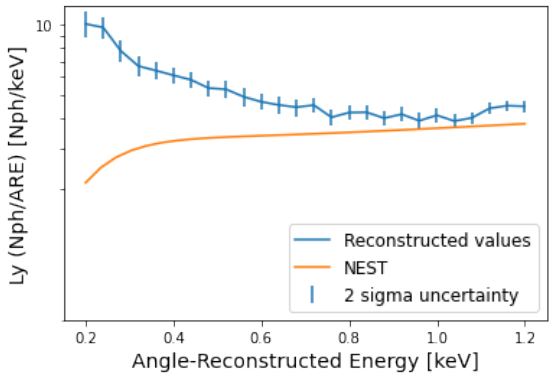
The systematic and statistical errors are low enough that our results are sensitive



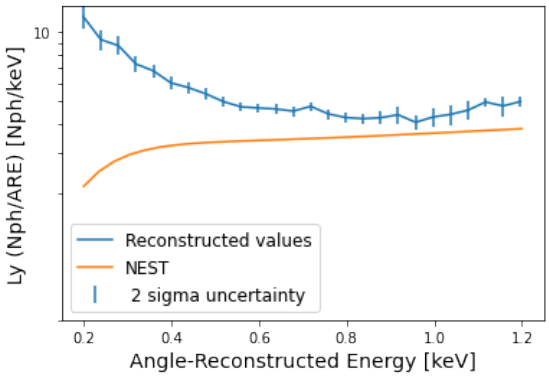

(a) Per-vertex $Q_y$, under only nonzero S1.




(b) Per-neutron $Q_y$, under only nonzero S1.




(c) Per-vertex $L_y$, under only nonzero S1.




(d) Per-neutron $L_y$, under only nonzero S1.

Figure 4.4: Light and charge yield curves after cut for events with zero S1 signals. The effect of this cut is to increase the length of the errorbars by a factor of, on average, 1.7 in the per-vertex case and 2.0 in the per-neutron case. In addition, the systematic becomes more pronounced due to the more significant reduction in statistics for lower multiplicity compared with higher multiplicity events.

to variations in the underlying light and or charge yields. In a real calibration, this implies we need to simulate many interactions under different assumptions for $Q_y$ and $L_y$, and statistically match the observed results to the expected results under each underlying assumption to determine the $Q_y$ and $L_y$ curves which best match observation.

## 4.6 Future Work

Several priorities for future work can be laid out.

First, it is important to extend analysis to include the effect of moving from electrons and photons produced, to electrons and photons detected. As discussed in chapter 2, the detection efficiencies are much lower for photons than for electrons (though can still be far from 100% for electrons produced deep in the detector), and so for many low energy scatter events no photons are detected. Using multi-scattered neutrons has the benefit of allowing for more total S1 signal to be produced so that even if an individual scatter is not detectable due to low photon signal, the set of scatters could still produce a total signal which is detectable. However, this leads to several difficulties.

First, as described above many neutrons produce zero S1 signal even summed over all scatter vertices. Such neutrons would have to be discarded from the analysis unless another method was developed to locate the set of vertices in time. One such method might involve using the timing tag from the H-Reflector (the surface from which neutrons are scattered prior to entering the beam pipe, which determines their incident energy). This provides another way of defining the timing of the neutron entering the detector. However, this method would be susceptible to false coincidence triggers due to the long travel time of low energy neutrons, which is on the order of the H-Reflector trigger rate within the applied signal cuts.

In addition, mapping the deposited energies to S1 signals, even once the scatters are time-ordered so that such a mapping is possible, is challenging. In the rare case in which all scatters produce a nonzero S1, and all S1's are separable in time, each S1 signal can be clearly associated with a vertex and thus with a deposited energy. However, if even one vertex produces zero S1 signal, it will not necessarily be clear which vertex, or whether it did in fact produce S1 but that signal was not separable from another S1 signal resulting from a different scatter. There are some special cases in which we could circumvent this to maximize the number of viable events. For instance, if all the scatters are time-separable, we can be confident that no two S1 signals are misidentified as a single signal. The time interval between each scatter is known from the determined path geometry, and so if we can determine which vertex even one S1 signal was associated with, we can infer which vertices produced the other S1 signals by looking forwards and backwards in time from that vertex. This could be achieved, for example, by finding one very large energy deposition among a set of smaller depositions and associating that with one very large observed S1 signal. Future work will need to devise a more complete method of mapping S1 signals to vertices in more complicated cases.

The next priority is to extend the simulations to full LZ geometry and materials, in order to more precisely understand the effects of the experimental set-up. For instance, the current simulations do not allow neutrons to scatter out of and then back into the detector.

# References


[1] Kevork N. Abazajian et al. "Strong constraints on thermal relic dark matter from Fermi-LAT observations of the Galactic Center". In: *Physical Review D* 102.4 (Aug. 2020). ISSN: 2470-0029. DOI: 10.1103/physrevd.102.043012. URL: http://dx.doi.org/10.1103/PhysRevD.102.043012.

[2] D. S. Akerib et al. "Liquid xenon scintillation measurements and pulse shape discrimination in the LUX dark matter detector". In: *Phys. Rev. D* 97 (11 June 2018), p. 112002. DOI: 10.1103/PhysRevD.97.112002. URL: https://link.aps.org/doi/10.1103/PhysRevD.97.112002.

[3] J. Allison et al. "Recent developments in Geant4". In: *Nucl. Instrum. Meth. A* 835 (2016), pp. 186–225. DOI: 10.1016/j.nima.2016.06.125.

[4] L. Baudis and Profumo. "Dark matter". In: *he Review of Particle Physics* (Aug. 2017).

[5] R. Bernabei et al. "Final model independent result of DAMA/LIBRA–phase1". In: *The European Physical Journal C* 73.12 (Nov. 2013). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-013-2648-7. URL: http://dx.doi.org/10.1140/epjc/s10052-013-2648-7.

[6] Gianfranco Bertone and Dan Hooper. "How dark matter became a particle". In: *CERNCOURIER* (2017). DOI: arXiv:1711.01693.

[7] Gianfranco Bertone, Dan Hooper, and Joseph Silk. "Particle dark matter: evidence, candidates and constraints". In: 405.5-6 (Jan. 2005), pp. 279–390. DOI: 10.1016/j.physrep.2004.08.031. arXiv: hep-ph/0404175 [hep-ph].

[8] A. Boyarsky et al. "Sterile neutrino Dark Matter". In: *Progress in Particle and Nuclear Physics* 104 (Jan. 2019), pp. 1–45. ISSN: 0146-6410. DOI: 10.1016/j.ppnp.2018.07.004. URL: http://dx.doi.org/10.1016/j.ppnp.2018.07.004.

[9] Man Ho Chan. "Electron–positron pair production near the Galactic Centre and the 511 keV emission line". In: *Monthly Notices of the Royal Astronomical Society: Letters* 456.1 (Dec. 2015), pp. L113–L116. ISSN: 1745-3933. DOI: 10.1093/mnrasl/slv187. URL: http://dx.doi.org/10.1093/mnrasl/slv187.

[10] Adrian Cho. "Can dark matter vanquish controversial rival theory?" In: *Science* 355 (Jan. 2017), pp. 337–337. DOI: 10.1126/science.355.6323.337.

[11] LUX Collaboration et al. *Low-energy (0.7-74 keV) nuclear recoil calibration of the LUX dark matter experiment using D-D neutron scattering kinematics*. 2016. arXiv: 1608.05381 [physics.ins-det].

[12] LZ Collaboration. *The status and science of the LZ dark matter experiment*. URL: https://lz.lbl.gov/.

[13] The LZ Collaboration et al. *LUX-ZEPLIN (LZ) Conceptual Design Report*. 2015. arXiv: 1509.02910 [physics.ins-det].

[14] John Ellis, Andrew Ferstl, and Keith A. Olive. In: *Space Science Reviews* 100.1/4 (2002), pp. 235–246. ISSN: 0038-6308. DOI: 10.1023/a:1015834615268. URL: http://dx.doi.org/10.1023/A:1015834615268.

[15] Brian D. Fields and Keith A. Olive. "Big bang nucleosynthesis". English (US). In: *Nuclear Physics, Section A* 777 (Oct. 2006), pp. 208–225. ISSN: 0375-9474. DOI: 10.1016/j.nuclphysa.2004.10.033.

[16] Richard J. Gaitskell. "Direct Detection of Dark Matter". In: *Annual Review of Nuclear and Particle Science* 54.1 (2004), pp. 315–359. DOI: 10.1146/annurev.nucl.54.070103.181244. eprint: https://doi.org/10.1146/annurev.nucl.54.070103.181244. URL: https://doi.org/10.1146/annurev.nucl.54.070103.181244.

[17] Dongqing Huang. "Ultra-Low Energy Calibration of the LUX and LZ Dark Matter Detectors". PhD thesis. Brown U., 2020. DOI: 10.26300/zvs6-fx07.

[18] Gerard Jungman, Marc Kamionkowski, and Kim Griest. "Supersymmetric dark matter". In: *Physics Reports* 267.5-6 (Mar. 1996), pp. 195–373. ISSN: 0370-1573. DOI: 10.1016/0370-1573(95)00058-5. URL: http://dx.doi.org/10.1016/0370-1573(95)00058-5.

[19] M. E. Kahil and T. Harko. "Is Dark Matter an Extra-Dimensional Effect?" In: *Modern Physics Letters A* 24.09 (Mar. 2009), pp. 667–682. ISSN: 1793-6632. DOI: 10.1142/s0217732309028667. URL: http://dx.doi.org/10.1142/S0217732309028667.

[20] John Marr. "Entropy and Mass Distribution in Disc Galaxies". In: *Galaxies* (2020).

[21] Enrico Morgante. *Aspects of WIMP Dark Matter Searches at Colliders and Other Probes*. Springer Theses, 2017.

[22] B. J. Mount et al. *LUX-ZEPLIN (LZ) Technical Design Report*. 2017. arXiv: 1703.09144 [physics.ins-det].

[23] Del Popolo. "Dark matter, density perturbations, and structure formation". In: *Astronomy Reports* 51 (2017), pp. 169–196.

[24] N. Prantzos et al. "The 511 keV emission from positron annihilation in the Galaxy". In: *Reviews of Modern Physics* 83.3 (Sept. 2011), pp. 1001–1056. ISSN: 1539-0756. DOI: 10.1103/revmodphys.83.1001. URL: http://dx.doi.org/10.1103/RevModPhys.83.1001.

[25] Vera Rubin and Kent Ford. "Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions". In: *Astrophysical Journal* 159 (1970), p. 379.

[26] Barbara Ryden. *Introduction to Cosmology, 2nd Edition*. Cambridge University Press, 2003.

[27] Marc Schumann. "Direct detection of WIMP dark matter: concepts and status". In: *Journal of Physics G: Nuclear and Particle Physics* 46.10 (Aug. 2019), p. 103003. ISSN: 1361-6471. DOI: 10.1088/1361-6471/ab2ea5. URL: http://dx.doi.org/10.1088/1361-6471/ab2ea5.

[28] M Szydagis et al. “Enhancement of NEST capabilities for simulating low-energy recoils in liquid xenon”. In: *Journal of Instrumentation* 8.10 (Oct. 2013), pp. C10003–C10003. ISSN: 1748-0221. DOI: 10.1088/1748-0221/8/10/c10003. URL: http://dx.doi.org/10.1088/1748-0221/8/10/C10003.

[29] Flip Tanedo. *Defense Against the Dark Arts: Notes on dark matter and particle physics*. Lecture Notes, Institute for High Energy Phenomenology, Aug. 2011.

[30] Charles Thorpe-Morgan et al. “Annihilating dark matter search with 12yr of Fermi LAT data in nearby galaxy clusters”. In: *Monthly Notices of the Royal Astronomical Society* 502.3 (Jan. 2021), pp. 4039–4047. ISSN: 1365-2966. DOI: 10.1093/mnras/stab208. URL: http://dx.doi.org/10.1093/mnras/stab208.

[31] J.R. Verbus et al. “Proposed low-energy absolute calibration of nuclear recoils in a dual-phase noble element TPC using D-D neutron scattering kinematics”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 851 (Apr. 2017), pp. 68–81. ISSN: 0168-9002. DOI: 10.1016/j.nima.2017.01.053. URL: http://dx.doi.org/10.1016/j.nima.2017.01.053.

[32] James Verbus. “An Absolute Calibration of Sub-1 keV Nuclear Recoils in Liquid Xenon Using D-D Neutron Scattering Kinematics in the LUX Detector”. PhD thesis. Brown University, 2016.

[33] Fritz Zwicky. “Die Rotverschiebung von extragalaktischen Nebeln [The Redshift of Extragalactic Nebulae]”. In: *Helvetica Physica Acta* 6 (1993), pp. 110–127. DOI: arXiv:1711.01693.

# Appendix

| Average Per-Vertex Energy Error, before cuts [eV] | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | All uncertainties | | No position uncertainty | | No energy uncertainty | | No uncertainties | |
| | All | Correct orderings | All | Correct orderings | All | Correct orderings | All | Correct orderings |
| **all** | 160 | 86 | 82 | 12 | 100 | 80 | 12 | 12 |
| **2** | 120 | 79 | 28 | 11 | 90 | 74 | 11 | 11 |
| **3** | 160 | 84 | 72 | 12 | 100 | 79 | 12 | 12 |
| **4** | 170 | 89 | 120 | 12 | 100 | 74 | 13 | 13 |
| **5** | 200 | 94 | 140 | 13 | 110 | 84 | 13 | 13 |
| **6** | 220 | 120 | 130 | 13 | 130 | 110 | 13 | 13 |

Table A1: Per-Vertex Energy error (per-neutron RMS deviation of TE to ARE, averaged over all neutrons in each sample) after turning on and off each source of uncertainty (energy and positional), and selecting for correctly ordered reconstructions. Turning off uncertainties decreases the energy error, making up the bulk (>90%) of total error in most cases. For n<4, error is driven by position uncertainty (turning this off results in a greater drop in error). Energy uncertainty (causing incorrect orders) only provides the dominant error when N>=4 without MOC cuts. Positional uncertainties dominate per-vertex energy reconstruction, and can be minimized with an additional distance separation cut.

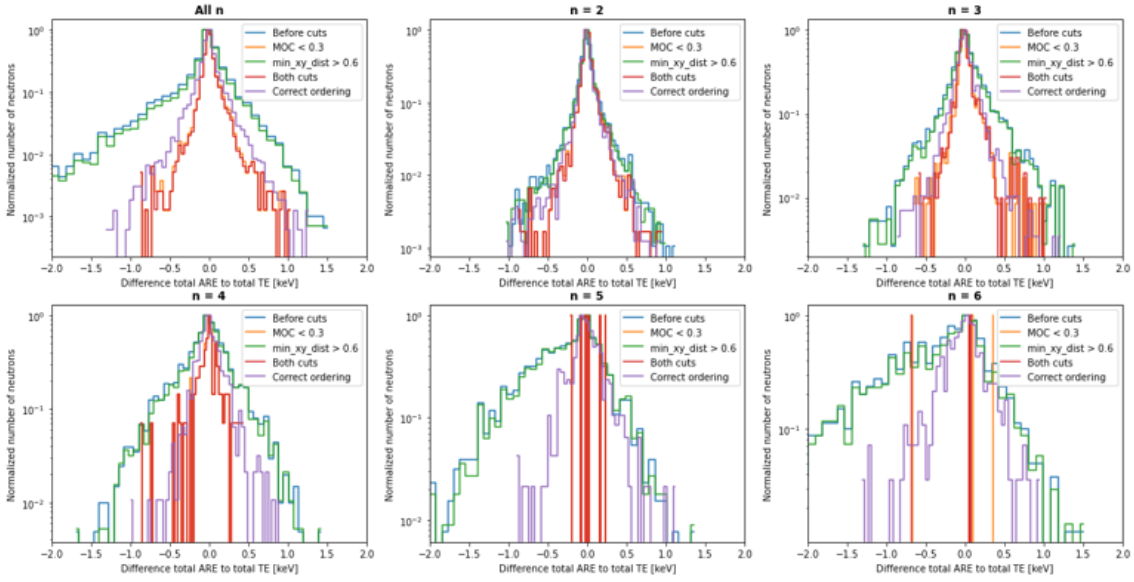



Figure A1: Histogram of abs(total TE - total ARE) (sum over all vertices), by multiplicity. Curves are normalized to their max value in order to compare the shapes. This shows the systematic bias at high multiplicity for underestimating the ARE when the order is chosen incorrectly.