

Dark Matter Searches and Measurement of the Cosmic Muon Flux with the LUX-ZEPLIN (LZ) experiment

Dissertation

zur

**Erlangung der naturwissenschaftlichen Doktorwürde
(Dr. sc. nat.)**

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät

der

Universität Zürich

von

Harvey J. Birch

aus

dem Vereinigten Königreich

Promotionskommission

Prof. Dr. Björn Penning (Vorsitz)

Prof. Dr. Laura Baudis

Prof. Dr. Ben Kilminster

Zürich, 2025

Abstract

Abstract goes here

Dedication

Always for you.
KJCB

Declaration

I declare that..

Acknowledgements

I want to thank...

Contents

List of Figures

List of Tables

1 Introduction

1.1 Evidence for Dark Matter

1.2 Candidates for Dark Matter

1.3 Direct Detection of Dark Matter

2 The LUX-ZEPLIN Dark Matter Experiment

The LZ experiment is currently the leading dark matter direct detection experiment in the search for WIMPs. The detector is located on the 4,850 ft level (4,300 mwe) of the Sanford Underground Research Facility (SURF) in the Homestake Mine (Lead, SD) [1]. At the core of the experiment is a dual-phase Time Projection Chamber (TPC) which is sensitive to low-energy nuclear recoils (NR), the signal which is produced through WIMPs interacting with liquid noble gases. One of the main backgrounds in a WIMP search are neutrons as they also interact through nuclear recoils and thus LZ employs an active veto system to remove them. Theoretically WIMPs will only interact with the xenon target however neutrons would interact in both the TPC and veto detectors. The TPC is housed within a vacuum insulated cryostat with a layer of liquid xenon (Skin) which acts as high voltage stand-off, this region is also instrumented with PMTs and is part of the active veto system. The LXe skin is used to veto mostly gamma ray interactions within the TPC volume, also being sensitive to neutrons. The cryostat is surrounded near hermetically by ten acrylic vessels filled with Gadolinium-loaded liquid scintillator (GdLS). The GdLS is observed by 120 PMTs and stands within 238t of DI water which provides shielding to the detector. A schematic of the detector is shown in ??.

2.1 Liquid Xenon Time Projection Chamber

The LZ TPC holds 7 t (5.6 t fiducial) of liquid xenon (LXe) above its cathode, there is an additional thin layer (8 mm thick) of gaseous xenon (GXe) at the top of the liquid. The volume measures approximately 1.5 m in height and diameter and the walls of the TPC are made from PTFE to improve light collection efficiency [1]. The TPC, Skin and Xe payload are housed within the Inner Cryostat Vessel (ICV) and the Outer Cryostat Vessel (OCV) provides a vacuum jacket for insulation. Both cryostat vessels are made from low radioactivity titanium [2]. When a particle scatters off a LXe atom a prompt scintillation signal (S1) is produced alongside free electrons, via ionisation of the LXe atom. Through the application of an electric field, the free electrons drift to the LXe surface and are extracted in the GXe layer. As the electrons accelerate through the GXe layer, a proportional amount of scintillation light (S2) is produced. Light produced from these particle interactions is observed by a top and bottom array of 3-inch Hamamatsu R11410-22 PMTs, 494 in total. Using both the S1 and S2 signals, position reconstruction techniques can be used to determine the xyz -position of the particle interaction. The time difference between the S1 and S2 signals combined with the drift velocity is used to determine the z -position of the interaction whilst the hit pattern of the S2 signal in the top PMT array provides xy -position. This operating principle of a TPC can be seen in ??.

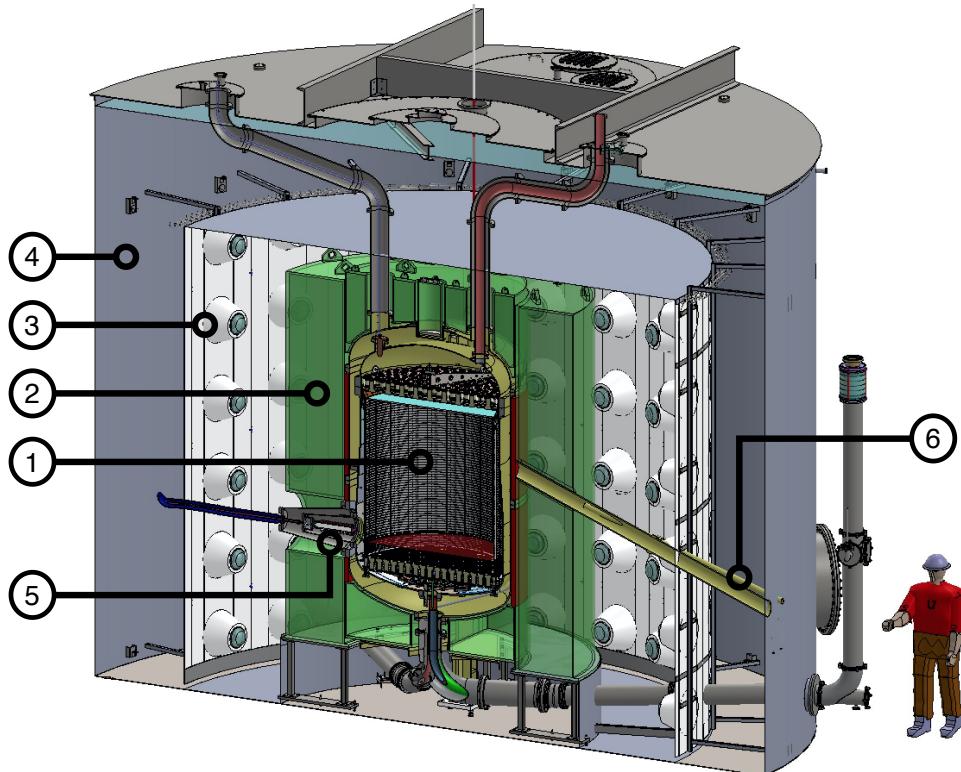
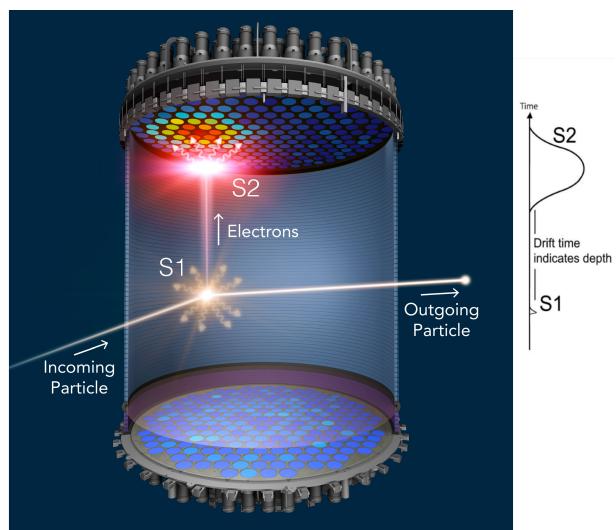


Figure 2.1: Schematic of the LZ detector showcasing the major subsystems. At the center is the liquid xenon TPC (1), monitored by two arrays of PMTs and serviced by various cable and fluid conduits (upper and lower). The TPC is contained in a double-walled vacuum insulated titanium cryostat and surrounded on all sides by a GdLS Outer Detector (2). The cathode high voltage connection is made horizontally at the lower left (5). The GdLS is observed by a suite of 8" PMTs (3) standing in the water (4) which provides shielding for the detector. The pitched conduit on the right (6) allows for neutron calibration sources to illuminate the detector [1].



(a) Operating principle of a dual phase liquid noble TPC. [3]

$$(b) y = 3 \sin x$$

2.1.1 Particle-Xenon interactions within a TPC

2.2 Xenon Skin

2.3 Outer Detector

2.3.1 Scintillator

2.3.2 PMT system

2.3.3 Optical Calibration System

2.4 Calibration Sources

2.5 Data Acquisition System

2.6 Simulation Techniques

2.7 Assembly and Operation of the LZ Detector

3 Outer Detector Commissioning and Monitoring

3.1 Single Photo-electron Calibration

3.2 Reconstructed Gain

3.2.1 Gain Curves

3.2.2 Monitoring PMT Health Over Time

3.3 Trigger Efficiency

3.4 Optical Calibration System Development

3.4.1 UV LED commissioning

3.4.2 Monitoring PMT

4 Veto Efficiency Studies for WS2024

A WIMP scatter should only deposit a small amount of energy (few keV) within the LXe volume of the experiment without any simultaneous energy deposit in the surrounding materials. Neutrons produced through radioactive decays within detector materials would mimic a WIMP interaction when they scatter off Xe atoms.

4.1 Simulation Matching

4.1.1 Tuning OD

Geometry Edits

Prior to WS2024 result, there were a number of major differences between simulations and data in the Outer Detector. As such, an effort was made to correct this. The changes made were;

1. A gap was added between the acrylic tanks to account for the gap seen during installation.
2. Water was added to the foam volume which is between the OCV and acrylic tanks; each
3. The acrylic tanks were moved further away from the OCV; this matches what adding the gaps was supposed to solve.

The percentage of water in the foam, and the outward movement of the acrylic tanks were looped over. The water percentage of the foam was increased in 1% steps, and the acrylic tanks were moved by 10mm steps. The 'best' value was found to be 30mm and 6%. Additional details on the discrepancies between simulations and data are shown in the subsequent subsections.

Capture Time

Following the geometry changes discussed above, the neutron capture timing using AmLi was studied. Prior to the geometry changes, there was a distinct discrepancy between data and simulation when considering neutron capture timing following single scatters within the TPC. Initial skims of both data and simulation were made using ALPACA, selecting all events which were classified as single scatters by LZAP. OD pulse information was also skimmed considering 100 keV (24 phd), 200 keV (49 phd) and 1 MeV (251 phd) OD thresholds. The simulation was modified in two different ways, moving the SATs out radially and adding water to the foam. All possible configurations were visually examined to determine which variation of simulation matched the data. An example of the comparison plot can be seen in ??, the "baseline simulation" was the initial configuration of the geometry prior to this study. All plots were examined side by side in a large scale canvas configuration seen in ???. It was found that from this study that 30 mm to 50 mm movement of the SATs alongside 5% to 7% increase in the percentage of water in the foam provides the best agreement between data and simulation at a 200 keV threshold.

4.2 Veto Selection Optimisation

In this section, details on how the Skin and OD cuts used in SR3 were created are defined. The Skin and OD veto cuts used in SR3 were based on those developed for SR1, described

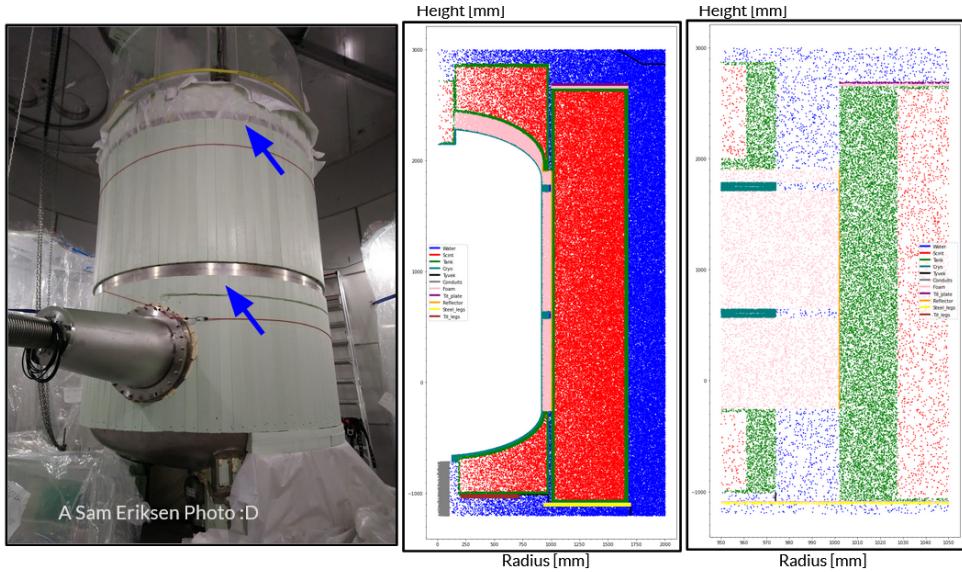


Figure 4.1: Geometry in the simulation used for SR3.

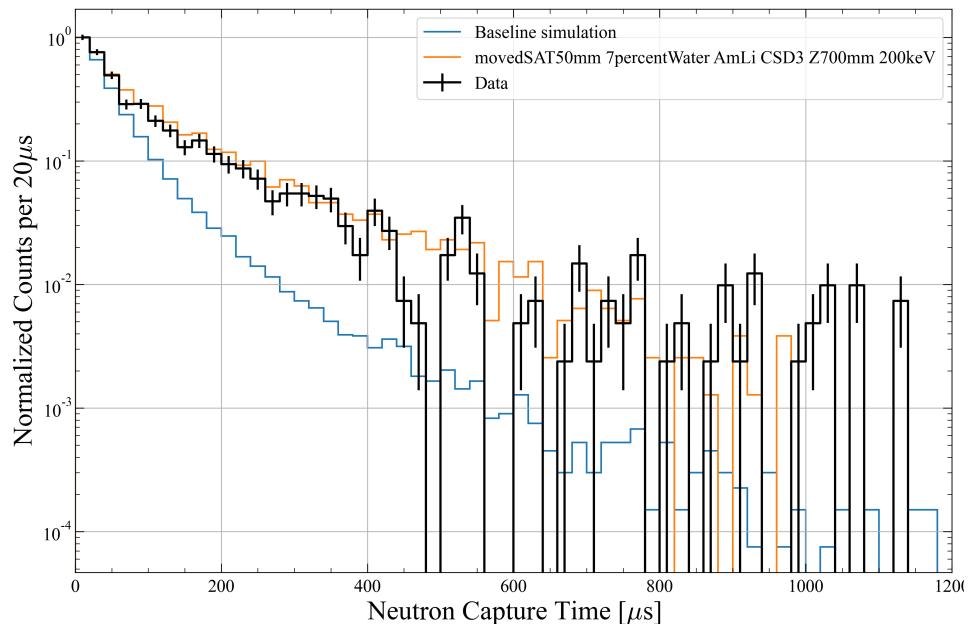


Figure 4.2: An example of the plot used to compare neutron capture timing in data with the baseline simulation and the modified simulation.



Figure 4.3: Large scale canvas of all possible simulation configurations. Each plot is similar in style to ???. Here AmLi at 700mm in CSD3 has been shown as an example.

in[LZ_AnalysisDocDB000071:xe127_skin_tagging, LZ_AnalysisDocDB000106:sr1_analysis, LZ_AnalysisDocDB000122:sr1_tuning_analysis]. This boils down to four cuts;

- Skin-prompt: For tagging γ -rays in the Skin detector,
- OD-prompt: For tagging γ -rays and neutron proton recoils in the OD detector,
- Skin-delayed: For tagging γ -rays from post-neutron capture deexcitation,
- OD-delayed: For tagging γ -rays from post-neutron capture deexcitation.

Using SR1 (and pre-SR2) as a baseline, three studies were performed to adapt these cuts for SR3;

1. Determine the detector stability, to establish if the energy scale for the detectors has changed; using this the SR1 cuts can simply be scaled for SR3, and acts as a starting point.
2. Implement a position correction to the OD.
3. Perform cut optimisation based on the above studies.

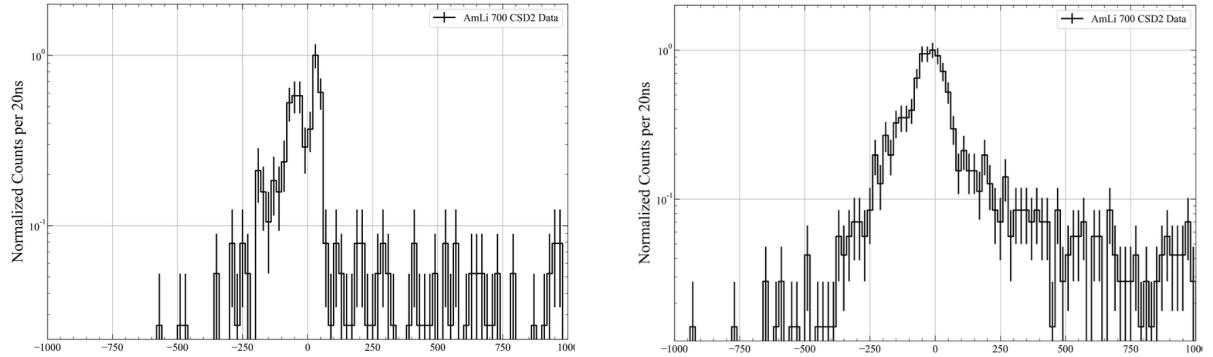
The cuts used for the vetoes were selected at the same time as AmLi neutron tagging efficiency calculation, and were set to maximise the tagging efficiency whilst reducing deadtime¹. As such, this section should be considered simultaneously with ??

On a skim of the AmLi calibrations which pass the SR3 core-cuts described in ??, the efficiency (described in ??) was calculated with the pulse area, and coincidence thresholds of the Skin and OD varying in integer steps. This then produces a heatmap of the coincidence vs. threshold vs. efficiency of each cut, examples of these are shown in ??-??. For the delayed cuts, the veto time window was also varied, and heatmaps of threshold vs. time window vs. efficiency were produced for a fixed coincidence. Heatmaps with dead time rather than efficiency were also produced, and are linked in ???. The dead time is discussed in greater detail in the next section.

The windows for the prompt cuts were selected by looking at the DD and AmLi calibrations (the run numbers and the LZAP version are listed in ??) and by measuring the time difference between the TPC single-scatter and OD and Skin pulses. For AmLi, this is shown in ??.

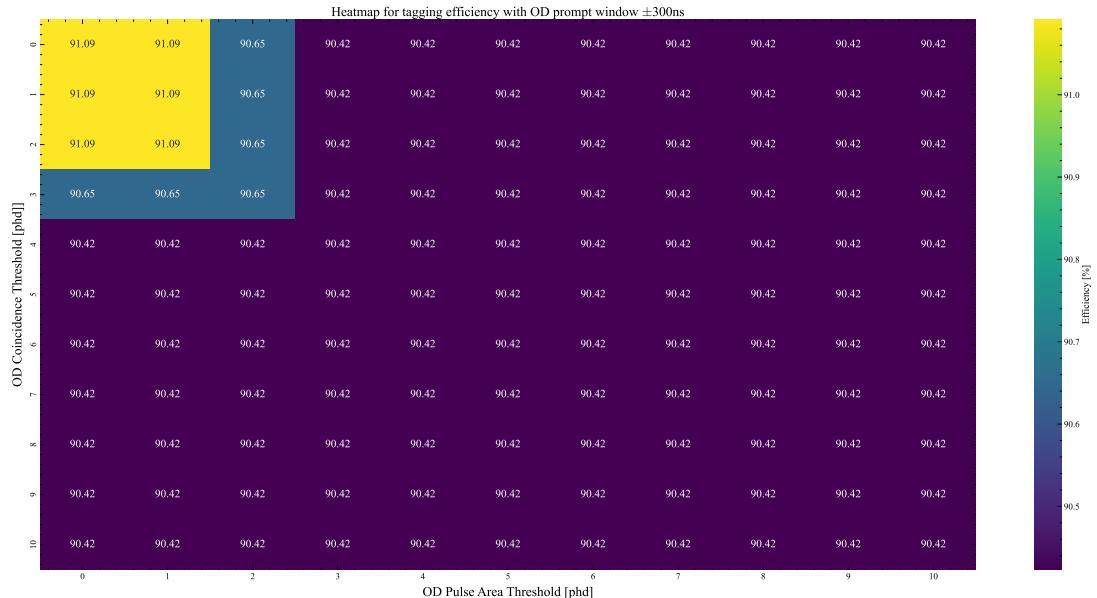
Compiling the heat maps of efficiency and deadtime for a variety of veto windows, a choice of thresholds were chosen to maximise efficiency whilst minimising dead time. An example plot we used to determine this is shown in ???. For SR3, we took the approach of trying to maintain the efficiency of SR1 veto of $\sim 90\%$, and to reduce the livetime impact. From the heatmaps, we determined that we could achieve an efficiency that at least matched the SR1 efficiency, but with a much lower deadtime. The final cuts are shown in ??.

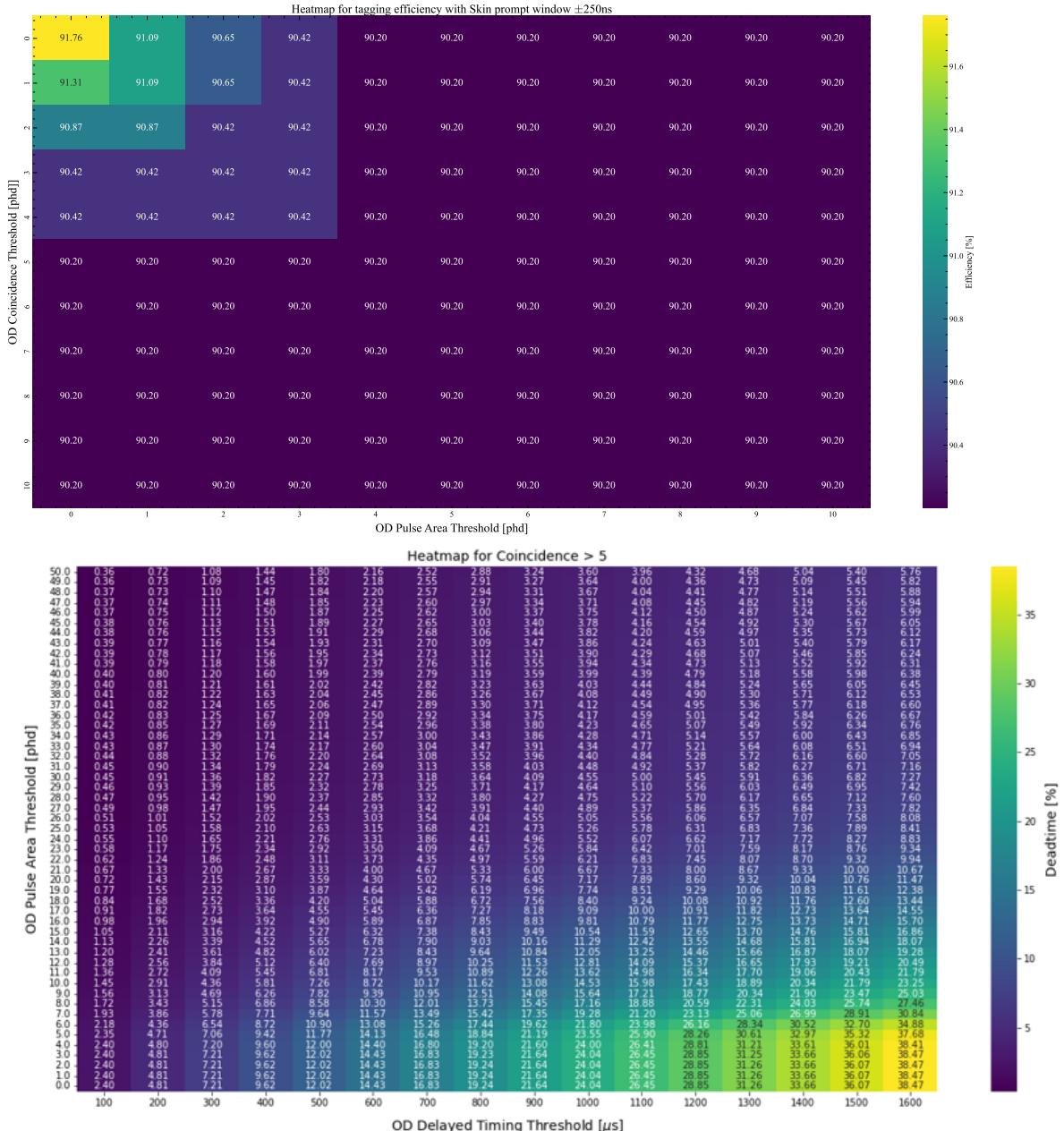
¹in future, a PLR study should be done to determine at what point this is optimised



- (a) Time difference (in ns) between TPC single-scatter and Skin pulses. To reduce the noise in the plot, a pulse requirement of greater than 2 phd and greater than 2 coincidence has been applied to the Skin pulses
 (b) Time difference (in ns) between TPC single-scatter and OD pulses. To reduce the noise in the plot, a pulse requirement of greater than 5 phd and greater than 22 coincidence has been applied to the OD pulses

Figure 4.4: Time difference between TPC single-scatter and Veto pulses. This is used to determine the prompt veto windows.





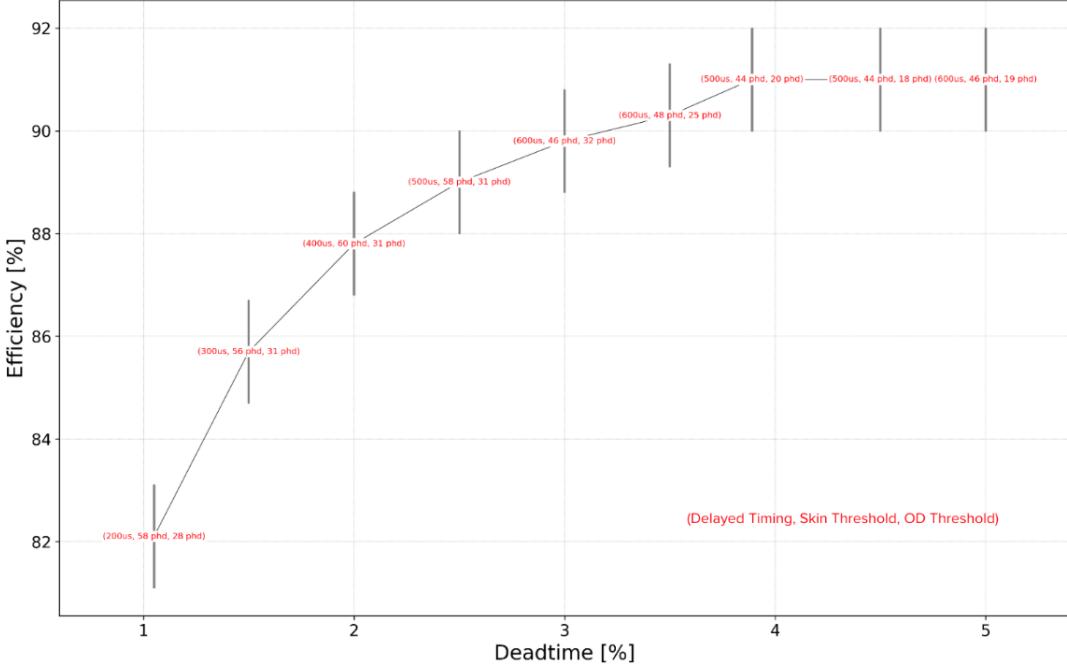
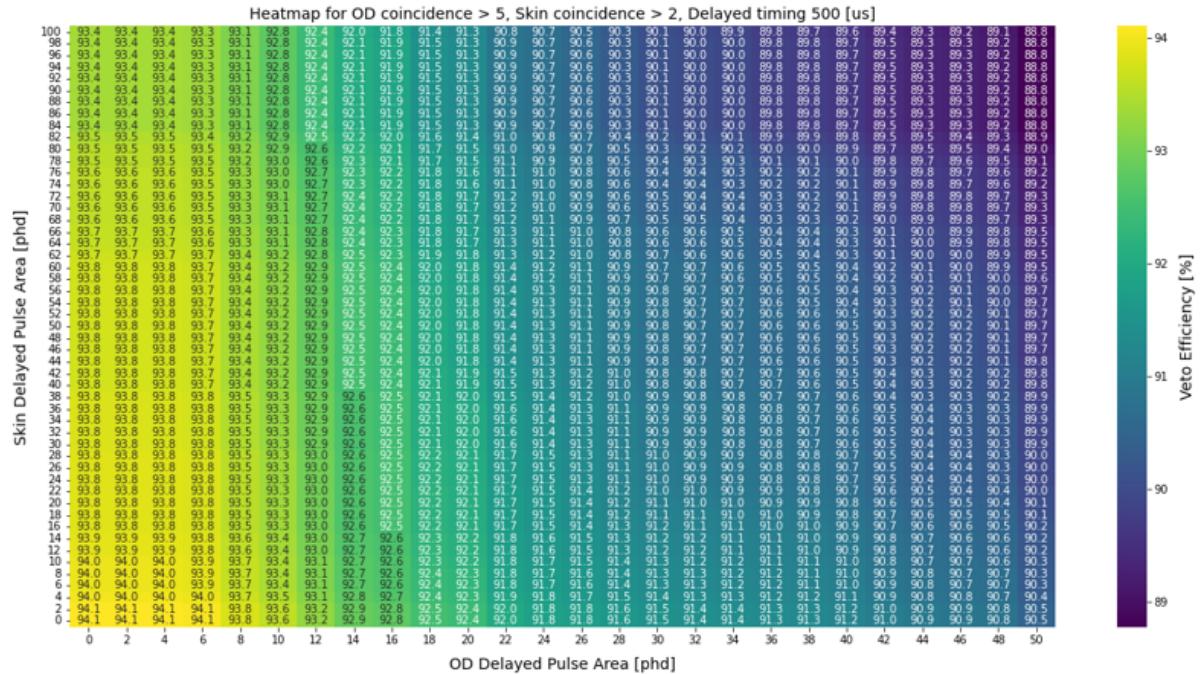


Figure 4.5: Deadtime-vs-efficiency highlighting a number of considered cuts. At each point, the numbers in brackets are; the delayed veto window length after the TPC single-scatter, the Skin threshold pulse area, and the OD threshold pulse area.



4.2.1 Deadtime Stability

The deadtime for each of the veto cuts was checked that it is stable over SR3. This was done by breaking up the SR3-WS file list, in this case, **SR3-WSv7-LZAP-5.8.0**, into month-by-month chunks. On each month, the rate of pulses in the Skin and OD from the second half of Random Trigger data was used; and the rate above threshold - where the threshold is the veto cut - is recorded. The stability of this over SR3 is shown in ??.

Skin Prompt		OD Prompt	
SR3	SR1	SR3	SR1
Window:	[−250, 250] ns	[500, 500] ns	
Coincidence:	> 2	> 2	
Pulse Area:	> 2.5	> 2.5	

Skin Delayed		OD Delayed	
SR3	SR1	SR3	SR1
Window:	[250ns, 600us]	[500ns, 1200us]	[300ns, 600us]
Coincidence:	> 2	> 55	> 5
Pulse Area:	> 46	> 50	> 32.0
			> 37.5

Figure 4.6: Cuts determined to be optimal for SR3. The SR3 thresholds have a position correction in Z. The SR1 cuts are also included.

The deadtime is then calculated as;

$$\text{Dead Time [\%]} = 1 - e^{-\lambda(x)} \quad (4.1)$$

where λ is the background rate and x , the veto threshold.

The conclusion that the dead-time is stable over SR3 was also checked by looking at the ODHealth PREM module, shown in ???. The PREM module shows the rate of OD pulses above 200 keV (as defined by the SR1-200 keV not from the SR3 energy calibration), with no significant fluctuation during SR3.

4.3 Neutron Veto Efficiency

In this section, the efficiency of tagging background neutrons using the Skin and OD detectors is calculated. This is performed by calculating the efficiency on AmLi and DD calibration data. This is then compared to AmLi and DD simulations. The difference between these simulations and data are then used to calculate an offset. The efficiency is then calculated for detector NR simulations and the offset applied (either a subtraction or a scaling, discussed later). This gives the neutron tagging efficiency for background neutrons.

4.3.1 Efficiency Calculation

The efficiency is defined as;

$$\epsilon[\%] = \frac{\text{N. Events passing Analysis Cuts} + \text{Veto Cuts}}{\text{N. Events passing Analysis Cuts}} \times 100 \quad (4.2)$$

and the inefficiency is defined as $100 - \epsilon$.

4.3.2 Neutrons From Calibration Sources

AmLi

The AmLi calibration runs are from the May-2023 calibration period before SR3, with run control operation name **Amlib-B**. The runs used are; 13004–13026. All files were processed with LZap-5.8.0.

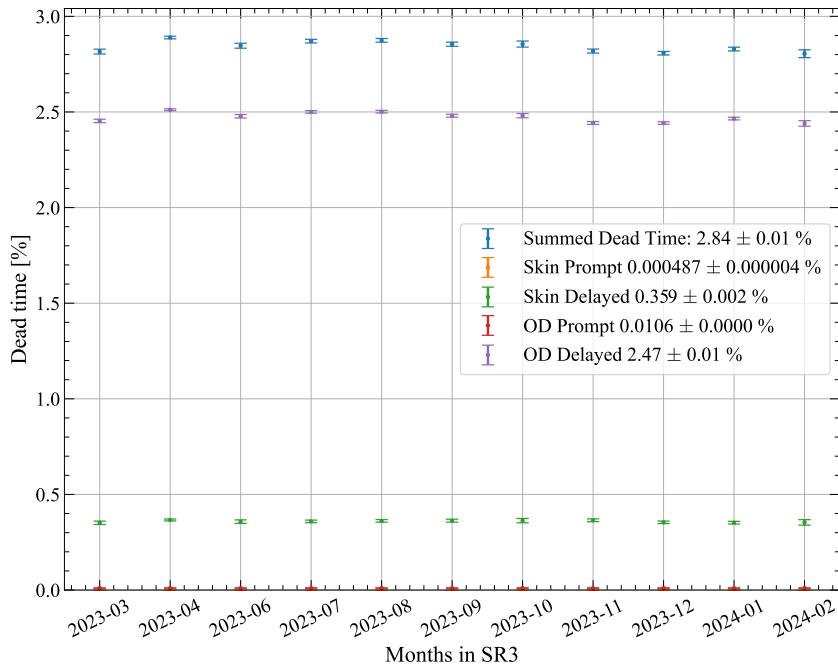


Figure 4.7: Deadtime from the Skin and OD veto cuts during each month of SR3. The error shown in purely statistical.

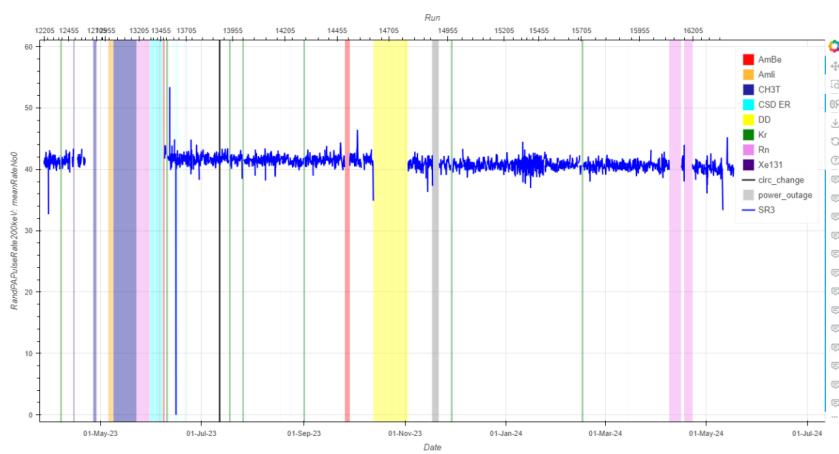


Figure 4.8: Rate of OD pulses above 200keV as defined by SR1 not SR3.

Livetime cuts	Physics cuts	S1 cuts	S2 cuts
Burst noise cut	Single scatter	S2 width vs drift time	S1 prominence cut
Muon holdoff	S1 and S2 threshold	Narrow S2	Stinger event cut
Sustained rate cut	Fiducial Volume	S2 rise time	S1 TBA vs drift time
High S1 rate exclusion		S2 early peak	S1 HSC cut
Bad buffer cuts		S2 XY quality	S1 shape
Excess Area cut		S2 TBA (above-anode gas)	S1 photon timing

Table 4.1: ALPACA-Core SR3 cuts used on AmLi calibration data for determining the efficiency.
Each cut is from SR3-cuts-v3.

On these events, the selection of cuts were applied. The majority of which are ‘standard’ SR3 cuts which are listed in ???. In addition to the cuts listed in ???, two other categorises of cuts were applied;

1. CSD tube: This is a circular cut on the reconstructed position of the SS in the TPC so that events from just one CSD tube are selected at a time. To make the comparison between data and simulation this cut is required because in simulation only one source in one tube was simulated as opposed to deploying a source in each CSD tube simultaneously. This is useful for examining the fluctuation in neutron tagging efficiency with (x, y, z) . When the efficiency is averaged across different heights and CSD tube, the average can then be replicated in both data and simulation. The cut uses the following logic;

```
def CSDSelection(x: float, y: float, which_csd: int=1):
    cent_csd_x = [87.6, -43.9, -43.9]
    cent_csd_y = [-0.13, 75.7, -75.7]
    x_el = pow((x-cent_csd_x[which_csd]), 2)
    y_el = pow((y-cent_csd_y[which_csd]), 2)
    r2_el = x_el + y_el
    return r2_el < 50**2
```

How this looks in (x, y) is shown in ???. A concern of this cut is that events towards the centre of the TPC are excluded. However, the position averaged efficiency when the CSD cut and not is $(88.21 \pm 1.03)\%$ and $(88.25 \pm 1.22)\%$ respectively. Across the $600\mu\text{s}$ window, there is no greater than 2% difference between when the cut is and isn’t applied. This comparison can be seen in ???

2. NR-band: This is a 1-, 3-, or $5-\sigma$ cut around the NR band median. The purpose of this cut is to improve the purity of the selection. In this note, the $1-\sigma$ is used. The band can be found on NERSC at /global/cfs/cdirs/lz/physics/NEST_Bands/SR3/20240313/Calibration, if the band has been, it can be recreated with the Woods parameters listed below;

The Woods Function Fit Parameters for the Band Mean are: [-6.7099328752961
The Woods Function Fit Parameters for the -1 Sigma Line are: [-7.995805026
The Woods Function Fit Parameters for the +1 Sigma Line are: [-5.428217339

The three different NR bands for the three calibration sources can be seen in ???.

AmLi Accidental Correction

When determining the veto tagging efficiency a correction must be applied to the measured efficiency to account for the accidental coincidences from AmLi gammas (and neutrons) with single scatter nuclear recoils in the TPC which can artificially enhance the measured tagging efficiency. First, veto inefficiency must be defined. Every time a single scatter nuclear recoil is

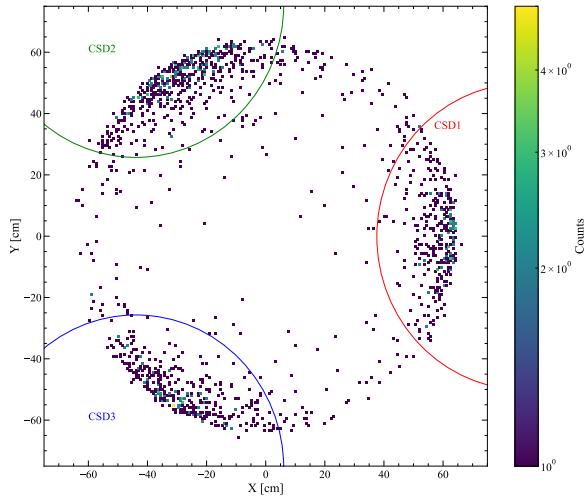


Figure 4.9: X-Y distribution of events in the TPC for AmLi sources positioned at 700mm. Each of the circular cuts are overlaid onto the plot.

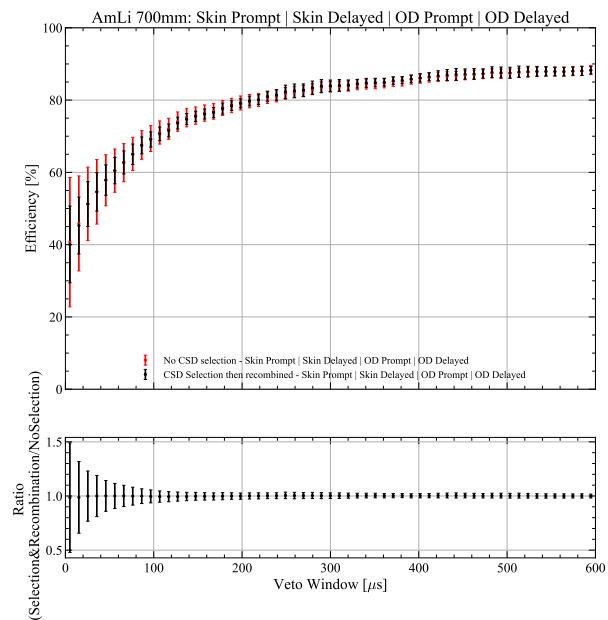


Figure 4.10: The veto efficiency (from all analysis cuts) at using AmLi at 700mm. The ratio plot shows the less than 2% difference between method 1 and method 2.

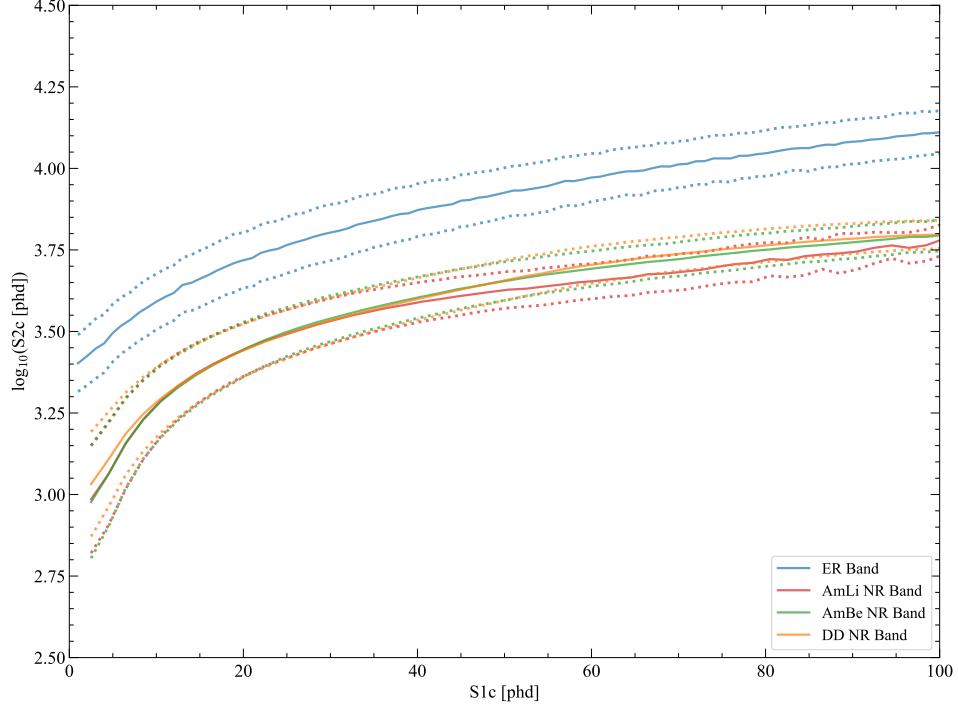


Figure 4.11: The three different NR bands used for the NR-band cut for the respective sources, AmLi, DD and AmBe.

observed in the TPC, a veto window is opened, this can be recorded using a counter, N every time this happens. If there is no pulse observed in the Skin or the OD that satisfies the veto requirements, this event is not vetoed and recorded using a counter, M . Resulting in a total inefficiency, M/N , and a total efficiency, $1 - M/N$. The effect of the accidentals be a result of the following process. If a neutron enters the TPC, scatters and is not detected by the vetos, the counter M should be iterated but there is an accidental coincidence of a gamma or a neutron from the AmLi sources with the TPC scatter so the count is not iterated. The probability of this happening can be written as $1 - P_a(0)$, where $P_a(0)$ is the probability of seeing zero accidental pulses in the Skin and OD coincidence windows. Using probability and the counters, the true inefficiency is described below,

$$\begin{aligned}
 M_{\text{observed}} &= M_{\text{true}} - M_{\text{true}}(1 - P_a(0)) \\
 M_{\text{observed}} &= M_{\text{true}}P_a(0) \rightarrow M_{\text{true}} = M_{\text{observed}}/P_a(0) \\
 \text{Ineff}_{\text{true}} &= M_{\text{true}}/N \text{ and } \text{Ineff}_{\text{observed}} = M_{\text{observed}}/N \\
 \Rightarrow \text{Ineff}_{\text{true}} &= \text{Ineff}_{\text{observed}}/P_a(0)
 \end{aligned}$$

Due to the logic of the inefficiency calculation is that the M counter iterates in an AND condition such that all coincidence windows in all detector have to be empty. This leads to a final inefficiency of,

$$\text{Ineff}_{\text{true}} = \text{Ineff}_{\text{obs}}/(1 - P_a(> 0)_{\text{any window}}) \quad (4.3)$$

$P_a(> 0)_{\text{any window}}$ can be determined directly using the post-trigger window of randomly triggered events (GPS events) and count any pulses in **any** of the veto detectors. The accidental correction is correlated with the length of the veto window, scans over the entire delayed veto window in $10\mu\text{s}$ steps, the change in the correction factor over time can be seen in ???. The impact of the accidental correction when applied to the efficiency can be seen in ???.

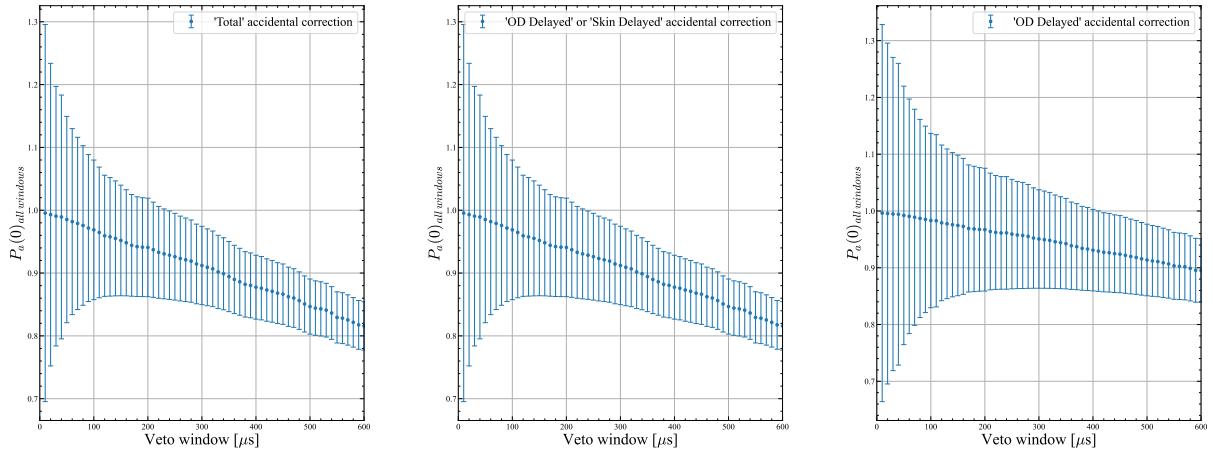


Figure 4.12: $P_a(> 0)_{\text{all window}}$ AmLi Accidental correction factors for varying veto window size for different windows of interest.

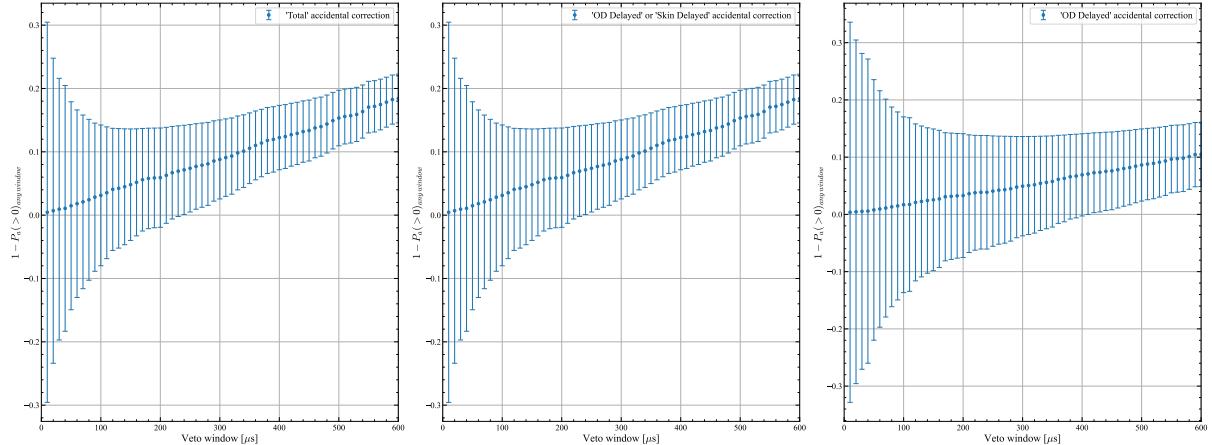


Figure 4.13: $1 - P_a(> 0)_{\text{anywindow}}$ AmLi Accidental correction factors for varying veto window size for different windows of interest.

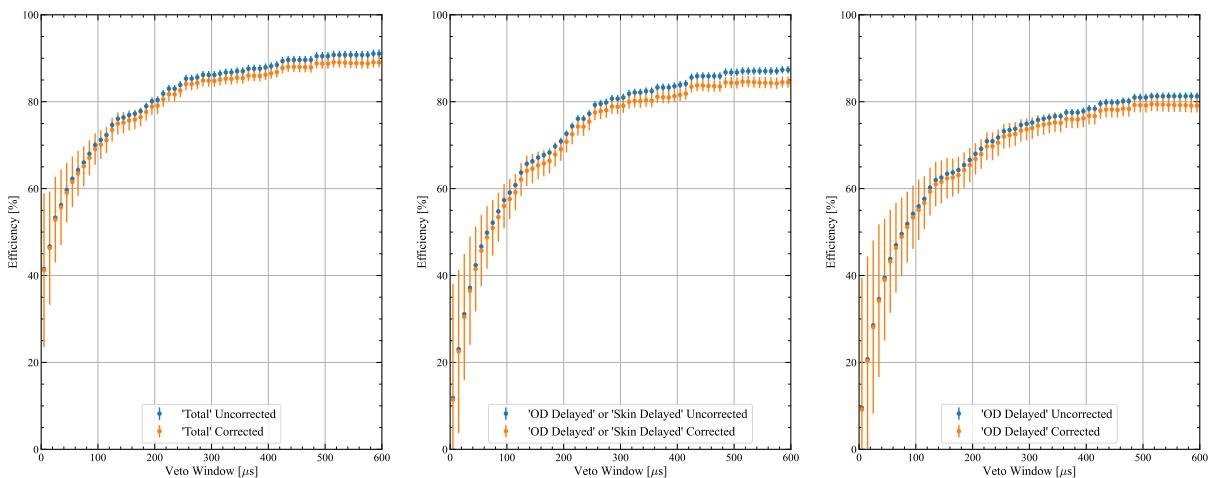


Figure 4.14: The impact of the accidental correction applied to the different veto efficiencies for the given windows of interest. AmLi data at a height of 700 mm in CSD1 has been used here as an example.

Physics cuts	S1 cuts	S2 cuts
Single scatter	S2 width vs drift time	S1 prominence cut
S1 and S2 threshold	Narrow S2	Stinger event cut
Fiducial Volume	S2 rise time	S1 TBA vs drift time
	S2 early peak	S1 HSC cut
	S2 XY quality	S1 shape
	S2 TBA (above-anode gas)	S1 photon timing

Table 4.2: ALPACA-Core SR3 cuts used on DD calibration data for determining the efficiency. Each cut is from SR3-cuts-v3.

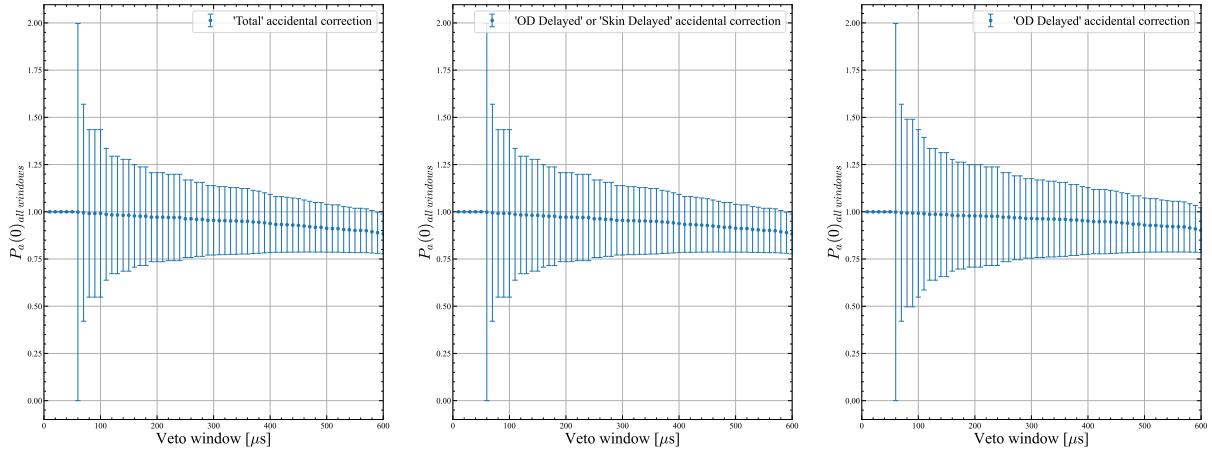


Figure 4.15: $P_a(> 0)_{all\ windows}$ DD Accidental correction factors for varying veto window size for different windows of interest.

DD Direct

The DD calibration runs are from the October-2023 calibration period during SR3, with run control operation name **DD Plasma**. The runs used are; 14631–14654. All files were processed with LZap—5.8.0.

The analysis for DD follows a logic similar to the AmLi described in ???. The cuts used for DD are listed in ???. In addition to these cuts, an NR band cut was also applied. It is important to note that this NR band is different to the one used for AmLi. The three different NR bands for the three calibration sources can be seen in ???. The band can be found on NERSC at [/global/cfs/cdls/lz/physics/NEST_Bands/SR3/20240313/Calibration](https://global.cfs.cdls/lz/physics/NEST_Bands/SR3/20240313/Calibration), if it has been moved, it can be recreated with the Woods parameters listed below;

The Woods Function Fit Parameters for the Band Mean are: [-19.937161961102728,
The Woods Function Fit Parameters for the 10% CL Line are: [-17.63553702315951
The Woods Function Fit Parameters for the 90% CL Line are: [-32.02840007751605

DD Accidental Correction

The DD calibration data was also corrected for accidentals. The same method was used which was previously discussed in ???. The correction factors used as a function of veto window can be seen in ???. The impact of the accidental corrections on the DD veto efficiency can be seen in ???.

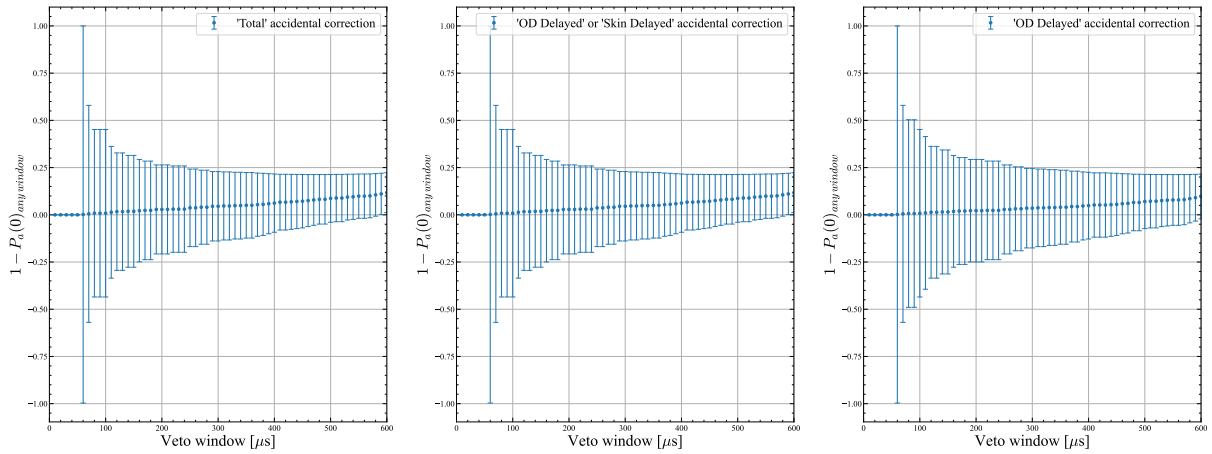


Figure 4.16: $1 - P_a(> 0)_{\text{any window}}$ DD Accidental correction factors for varying veto window size for different windows of interest.

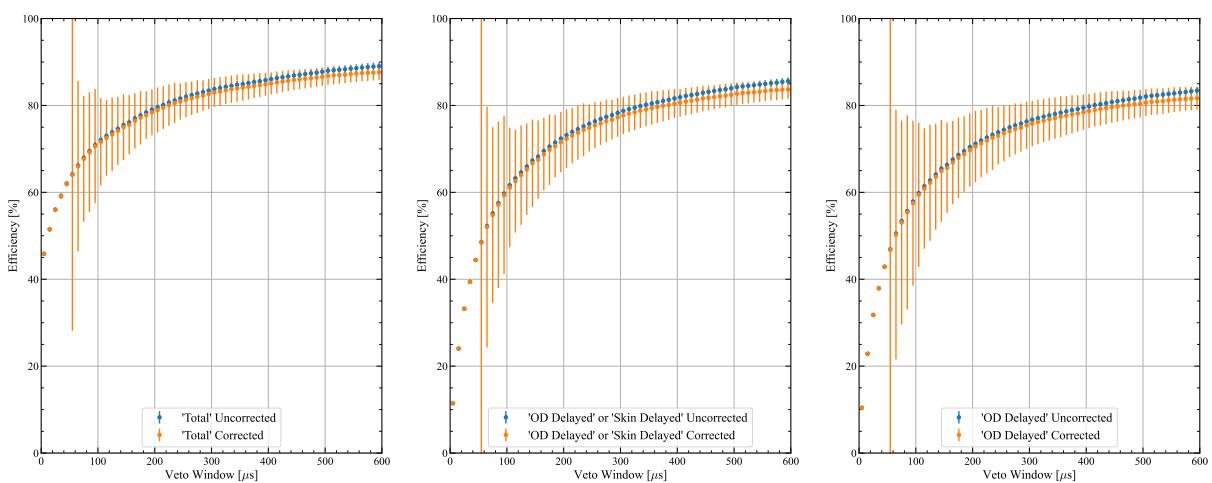


Figure 4.17: The impact of the accidental correction applied to the different veto efficiencies for the given windows of interest.

Physics cuts
Single scatter
S1 and S2 threshold
Fiducial Volume
CSD Selection (AmLi Only)

Table 4.3: ALPACA-Core SR3 cuts used on AmLi and DD simulations for determining the efficiency. Each cut is from SR3-cuts-v3.

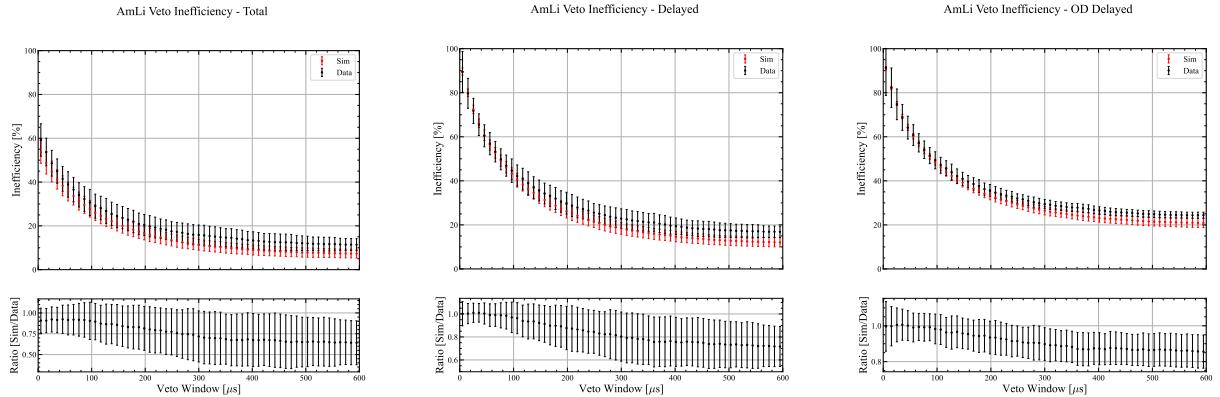


Figure 4.18: Inefficiency plots for AmLi, comparing simulations to data. Left: Total. Middle: Delayed Only. Right: OD Delayed.
In each case, the average from all CSD positions are used.

4.3.3 Simulated Neutrons From Calibration Source

AmLi and DD were simulated using BACCARAT–6.3.5 and LZLAMA–3.5.3. On each set of simulation, the cuts listed in ?? were used. No accidental correction was applied to the simulation data as there are no accidental gammas or neutrons present in the simulation. How these compare to the data measurements are shown in ??-??. More detailed plots on how the calibration simulations compare to calibration data is in ??.

4.3.4 Background Neutrons

The detector-NR simulations (described later, along with normalisations in ??) were simulated as part of the official production of SR3-bg. Each of the 644 components were simulated using BACCARAT-6.3.5 and LZLAMA-3.5.3. On these events, the cuts listed in ?? were applied.

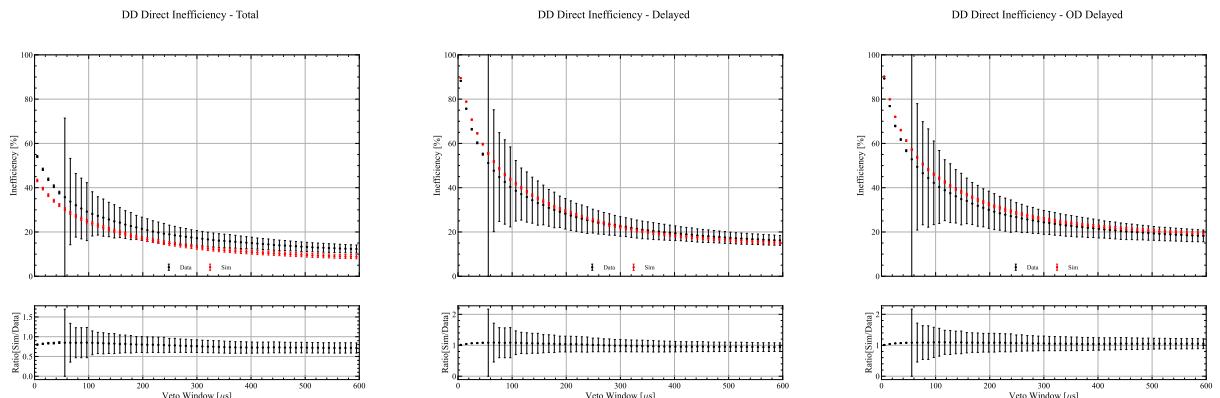


Figure 4.19: Inefficiency plots for DD, comparing simulations to data.
Left: Total. Middle: Delayed Only. Right: OD Delayed.

Physics cuts
Single scatter
S1 and S2 threshold
Fiducial Volume

Table 4.4: ALPACA-Core SR3 cuts used on Detector-NR simulations for determining the efficiency. Each cut is from SR3-cuts-v3. The contribution of each detector component is scaled after the cuts are applied and is discussed in ??.

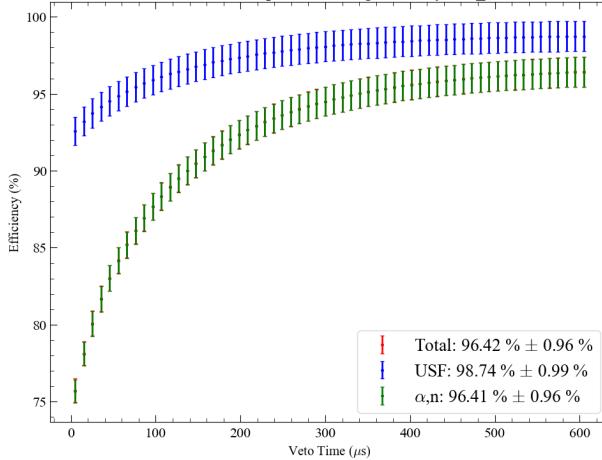


Figure 4.20: Efficiency for tagging a neutron on Detector-NR simulations.

The efficiency of tagging neutrons from USF and (α, n) events are shown in ???. Also shown is the total efficiency.

4.3.5 Background Neutron Efficiency From Data

Shown in ?? is the neutron veto efficiency from each calibration source from data and simulation. Also shown is the efficiency from simulated background neutrons. The Z-position of each point is calculated from the mean(driftTime) of events which pass the selection. For the Detector-NR, the events were split up into Z-sections by the following;

```
def SR3ZThirds_Cut(driftTime_us: float, position='top'):
    if position == 'top':
        return ((driftTime_us < 350.) & (driftTime_us > 71.))
    elif position == 'mid':
        return ((driftTime_us < 700.) & (driftTime_us > 350.))
    elif position == 'bot':
        return ((driftTime_us < 1030.) & (driftTime_us > 700.))
```

This driftTime splitting was performed so that comparing to calibration points was easier, but in evaluating the total efficiency, it is not used.

Shown in ?? are the veto efficiencies of all calibration sources for both simulations and data. The value we care about is the background neutron efficiency, $\epsilon_{\text{bg data}}$, or inefficiency, $\zeta_{\text{bg data}}$. There are a number of ways in which this can be calculated, which are described below. The proposed option is no.2, of $7.8 \pm 4.3\%$; or quote as 8 ± 4 . ?? contains the results of each approach.

Option 1 : We take the average simulation calibration efficiency, so the mean of AmLi (92.8%) and DD (91.8%), then we have 92.3%. We take the average data calibration efficiency, 88.6% for AmLi and 87.7%, for DD we have 88.2%. The ratio, Δ , of the simulation calibrations and the data calibration can then be used as a scaling factor on the simulation background. This process is expressed in ??, and gives an efficiency of 92.1%. In this case, the uncertainty is 4.5%, made up from a statistical error from the simulations (0.96%), and 4.3 from the scaling.

$$\epsilon_{\text{cal. data}} = \frac{\epsilon_{\text{data DD}} + \epsilon_{\text{data AmLi}}}{2} \quad (4.4)$$

$$\epsilon_{\text{cal. sims}} = \frac{\epsilon_{\text{sims DD}} + \epsilon_{\text{sims AmLi}}}{2} \quad (4.5)$$

$$\Delta = \frac{\epsilon_{\text{cal. data}}}{\epsilon_{\text{cal. sims}}} \quad (4.6)$$

$$\epsilon_{\text{bg data}} = \Delta \times \epsilon_{\text{bg sims}} \quad (4.7)$$

Option 2 : We take the average simulation calibration efficiency, so the mean of AmLi (92.8%) and DD (91.8%), then we have 92.3%. We take the average data calibration efficiency, 88.6% for AmLi and 87.7%, for DD we have 88.2%. The difference, Λ , between the simulation calibrations and the data calibrations can then be used as a systematic uncertainty on the simulation backgrounds. This process is expressed in ??, and gives an efficiency of 92.2%. In this case, the uncertainty is 4.3%, made up from a statistical error from the simulations (0.96%), and 4.22 from the subtraction.

$$\epsilon_{\text{cal. data}} = \frac{\epsilon_{\text{data DD}} + \epsilon_{\text{data AmLi}}}{2} \quad (4.8)$$

$$\epsilon_{\text{cal. sims}} = \frac{\epsilon_{\text{sims DD}} + \epsilon_{\text{sims AmLi}}}{2} \quad (4.9)$$

$$\Lambda = \epsilon_{\text{cal. sims}} - \epsilon_{\text{cal. data}} \quad (4.10)$$

$$\epsilon_{\text{bg data}} = \epsilon_{\text{bg sims}} - \Lambda \quad (4.11)$$

Option 3 : We take the average simulation calibration inefficiency, so the mean of AmLi ($100 - 92.8 = 7.2\%$) and DD ($100 - 91.8 = 8.3\%$), then we have 7.7% . We take the average data calibration inefficiency, 11.5% for AmLi and 12.3% for DD, we have 11.9% . The ratio, Δ , of the simulation calibrations and the data calibration can then be used as a scaling factor on the simulation background. This process is expressed in ??, and gives an inefficiency of 5.5 . In this case, the uncertainty is 2.8% , made up from a statistical error from the simulations (0.96%), and 1.8 from the subtraction.

$$\zeta_{\text{cal. data}} = 100 - \frac{\epsilon_{\text{data DD}} + \epsilon_{\text{data AmLi}}}{2} \quad (4.12)$$

$$\zeta_{\text{cal. sims}} = 100 - \frac{\epsilon_{\text{sims DD}} + \epsilon_{\text{sims AmLi}}}{2} \quad (4.13)$$

$$\Delta = \frac{\zeta_{\text{cal. data}}}{\zeta_{\text{cal. sims}}} \quad (4.14)$$

$$\zeta_{\text{bg data}} = \zeta_{\text{bg sims}} \times \Delta \quad (4.15)$$

5 Muons

5.1 Veto for WIMP Search

5.1.1 Motivation

5.1.2 Cut Description

5.2 Flux Measurement

Our initial muon model described in Ref. [4] has been used to calculate muon fluxes that can be compared with previous measurements. Vertical muon intensity has been calculated as $5.18 \times 10^{-9} \text{ cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1}$, in good agreement with the value of $(5.38 \pm 0.07 \text{ (stat)}) \times 10^{-9} \text{ cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1}$ [5] measured in the same cavern using the veto system of the Davis experiment (corrected to include multiple muon events). Total muon flux through a spherical detector with unit cross-sectional area has been calculated as $6.16 \times 10^{-9} \text{ cm}^{-2} \text{ s}^{-1}$, slightly higher than the recent measurements with the veto system of the Majorana Demonstrator located in the nearby cavern $(5.31 \pm 0.17 \text{ (stat)}) \times 10^{-9} \text{ cm}^{-2} \text{ s}^{-1}$ [6].

The difference between the two measurements and our initial model, although relatively small (within 20%), pointed to a need for new measurements with the LZ experiment to re-normalise the muon model.

The new muon flux measurements estimate the average density of the rock between the surface and the cavern and will be used to inform the muon model for the DUNE experiment[7], which uses the same MUSUN code and the same rock density.

5.2.1 Outline of Model

The muon simulations use two Monte-Carlo codes: MUSIC (MUon SImulation Code) and MUSUN (MUon Simulations UNderground)[8, 9], adapted here for the LZ experiment. Initially, muons with different energies on the surface of the Earth were transported through various distances in rock using MUSIC. The rock composition has been taken from Ref. [10, 11]. Measurements of several rock samples have been reported [10, 11] and the average rock parameters have been calculated as $\langle Z \rangle = 12.09$ and $\langle A \rangle = 24.17$. The average density of rock was assumed to be 2.70 g/cm^3 [11] in the MUSIC simulation. Other measurements [12], including the measurement of the muon flux with the veto system of the Majorana demonstrator, suggest that the density may be larger ($2.8\text{-}2.9 \text{ g/cm}^3$ [6]).

The resulting energy spectra of survived muons have been convoluted with muon spectra on the surface of the Earth for different zenith angles and slant depths (see [9] for the procedure's description). The distance from the underground laboratory and the Earth's surface for each azimuth and zenith angle has been obtained from the surface map and the position of the laboratory in the global coordinate system [13]. Muon energy spectrum and zenith angle distribution on the surface of the Earth were calculated using the parametrisation first suggested in Ref. [14] modified for the curvature of the Earth to include large zenith angles. Other corrections, such as muon decay and energy loss in the atmosphere, muon production via charmed meson decay, muon intensity dependence on altitude and geomagnetic rigidity cut-off were not included due to their negligible effect on high-energy muons ($> 1 \text{ TeV}$) capable of reaching the SURF location. As a result of muon transport and convolution with the surface fluxes, muon energy spectra and different angles at the detector site have been calculated and stored in a file. The model gives

the total muon flux through a sphere as $6.16 \times 10^{-9} \text{ cm}^{-2} \text{ s}^{-1}$, mean muon energy of 283 GeV, mean zenith angle of 27° and the mean slant depth of 4500 m w. e. [15].

The Muon generator MUSUN inside BACCARAT reads muon distributions from this file and samples muons on the horizontal (top) and vertical surfaces of a box that encompasses the laboratory hall. The top horizontal surface of the box is located 7 metres above the cavern boundary. Vertical surfaces of the box are positioned 5 m away from the boundary. The box extends to 3 m below the cavern floor. Moving the box surfaces into the rock, away from the cavern, ensures the development of muon-induced cascades in rock or shotcrete before muons and their secondaries enter the cavern. The primary energy spectrum and the angular distributions of these muons are illustrated in ??.

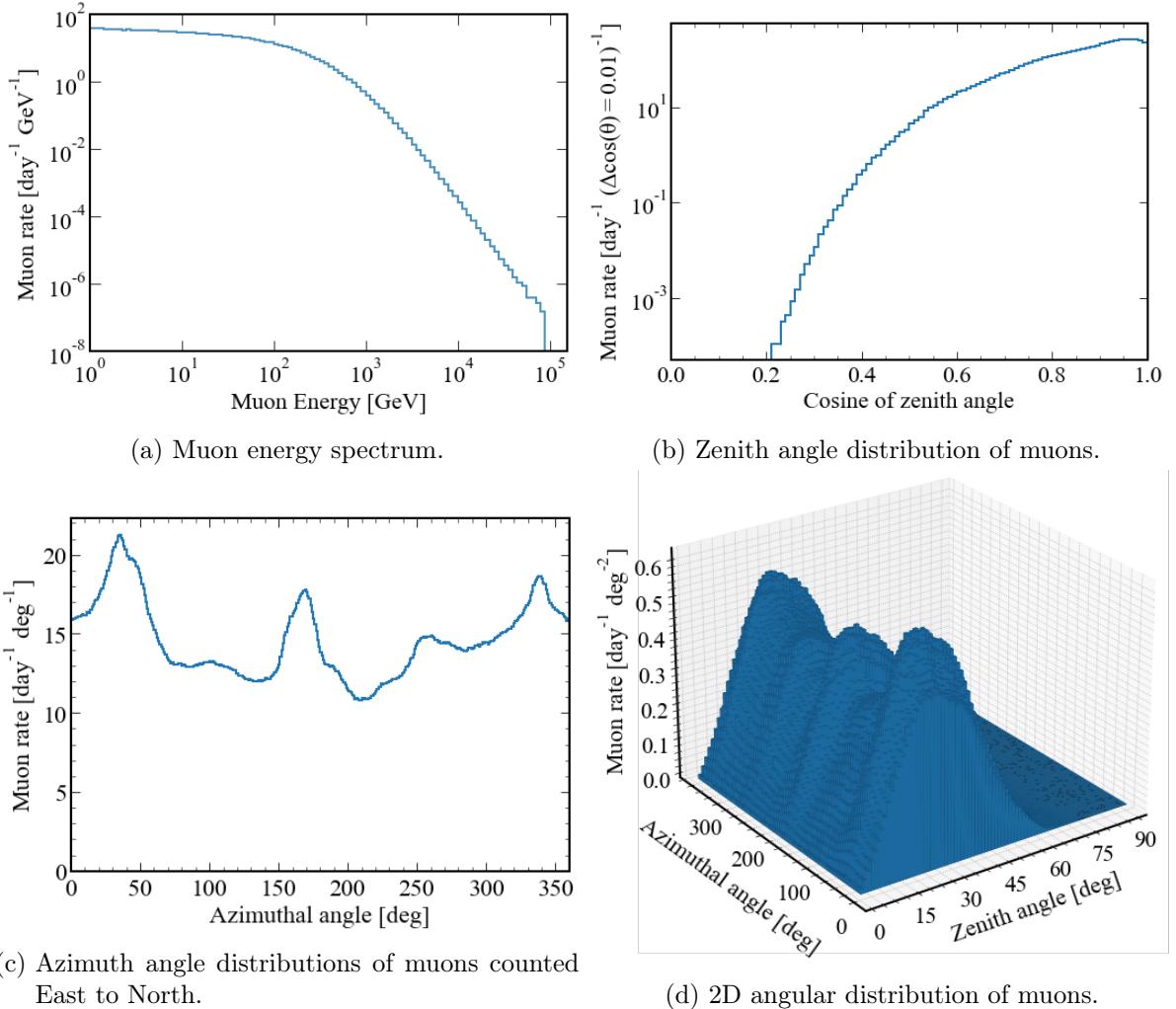


Figure 5.1: Angular distributions and the kinetic energy spectrum of muons at SURF, Davis Campus, as generated by the MUSUN model.

Muons generated by MUSUN have been transported through the rock, laboratory hall and detector geometry using GEANT4 (inside BACCARAT) and their energy depositions in all "sensitive volumes" have been recorded. In this simulation, 4.82×10^7 muons have been propagated which corresponds to a live time of 9147 days (0.0609874 muons per second), providing about 30 times higher statistics than for data collected during WIMP search 2022 and WIMP search 2024 (WS2022 & WS2024).

5.2.2 Data selection

This study considers the both WS2022 and WS2024 datasets, specifically using the file lists, SR1-WSv5_LZAP-5.4.6.files.list, taken from the SR1WS ALPACA module and SR3-WSv8_LZAP-5.8.0_all.files.list, taken from the SR3LENR ALPACA module.

Skims of the data sets were performed using the MuonPhys ALPACA module which applied a loose initial selection to reduce the size of the initial dataset for later analysis. This module was also used for determining the live time of the two datasets. The live times are summarised in ???. For the rate calculation, DAQ live time was chosen as the appropriate live time to use as this analysis is does not consider the same decision logic as used in the physics live time.

Table 5.1: Summary if relevant durations for the WIMP search runs. For this analysis, DAQ livetime was used in the rate calculation work.

Run	Run Time [days]	DAQ Live Time [days]	Physics Live Time [days]
WS2022	99.5	98.8	96.4
WS2024	268.0	267.6	255.4
Total	367.5	366.4	351.8

Typical Muon Event

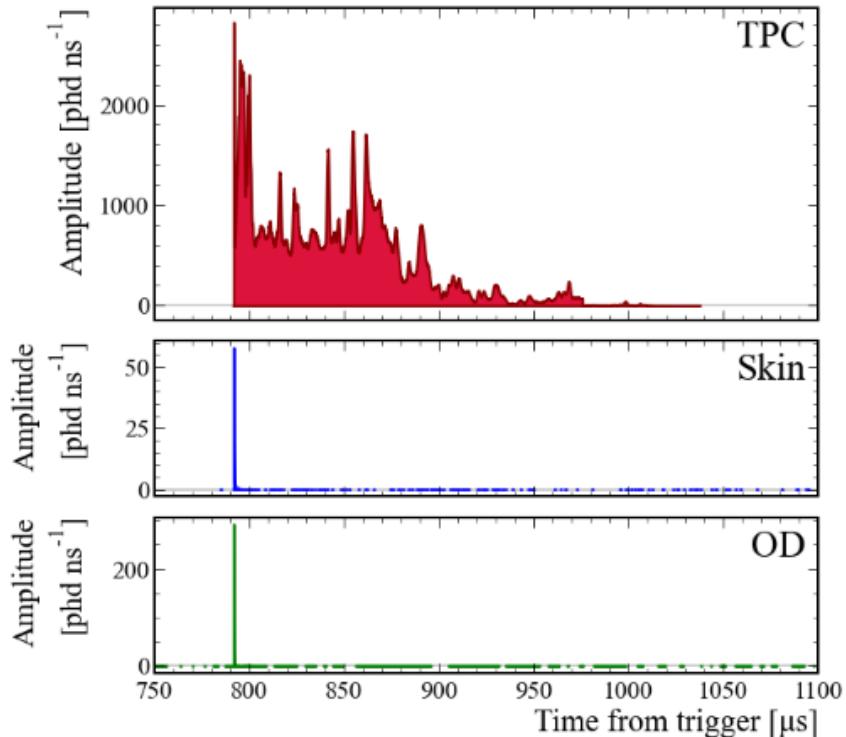


Figure 5.2: Event-viewer output of typical muon event. PMT saturation is evident towards the end of the muon ‘tail’ in the TPC.

OD Selection

An OD energy threshold was implemented to reduce the probability of random coincidence between the three detectors. The energy deposited in the OD by a muon is the largest, physical signal we will see in the OD. The only other large physical signals we anticipate seeing are neutron captures, the most frequent of which are on gadolinium. Therefore, the OD threshold

was set at the endpoint of the gadolinium neutron capture at ~ 8 MeV which corresponds to 2000 phd. An OD ‘noise cut’, initially developed for the muon veto and ‘hold off’ for the WIMP search, has been used to reduce the impact of burst noise in our flux measurement [16]. During tests of the muon veto, it was found that pulses produced due to the burst noise phenomena could imitate a muon-like signal in the OD. The ratio of pulse area to pulse amplitude enabled the creation of a custom variable called ‘Pulse Shape’. Two distinct distributions can be seen in ???. Events from both distributions were hand scanned using the event-viewer and it was found that pulses with a pulse shape less than 0.003 were noise pulses with the same characteristics as previously observed burst noise.

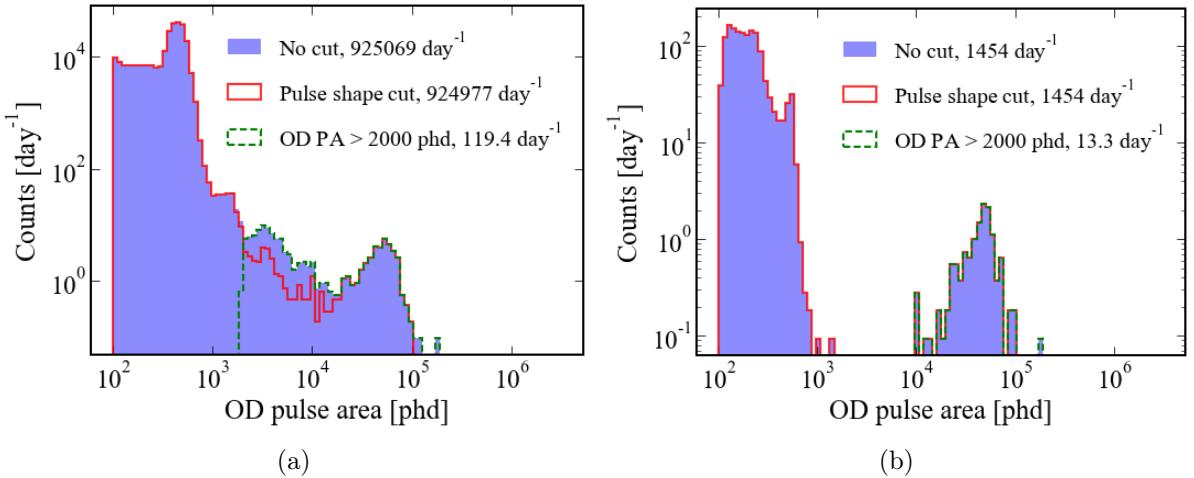


Figure 5.3: Histograms of the largest OD pulse areas in ~ 12.5 million background events from WS2022 before applying inter-detector timing cuts. Pulse shape has been used to differentiate between pulses produced by a muon and pulses produced by noise phenomena observed in the detector. The peak at around 450 phd results from neutron captures on the nuclei of hydrogen atoms and the subsequent emission of a single 2.22 MeV gamma ray. The plots show the impact of the 2000 phd cut and pulse shape cut (a) before and (b) after inter-detector timing coincidences are applied.

Inter-detector Timing Selection

Cosmic-ray muon events in the LZ experiment can be uniquely identified by their large energy deposits that are coincident in the three detector systems: OD, skin and TPC. As muons move at relativistic speeds, the time difference between pulses in each detector will be small compared to other backgrounds such as neutrons and gamma rays. Broad scans were conducted to make use of the coinciding detector signals to determine the inter-detector timing selection. As shown in ???, the time difference distributions between detector volumes allow for distinguishing between muon interactions through all volumes and signals unrelated to muons. Following the scan, three different considerations were made:

1. $-200 \text{ ns} < \Delta t_{\text{OD} - \text{Skin}} < 200 \text{ ns}$
2. $-1200 \text{ ns} < \Delta t_{\text{OD} - \text{TPC}} < 200 \text{ ns}$
3. $-1200 \text{ ns} < \Delta t_{\text{Skin} - \text{TPC}} < 200 \text{ ns}$

Here the pulse time, t , is described using the combination of two RQs, `pulseStartTime_ns+areaFractionTime`. When traversing the TPC, a muon does not produce a singular pulse as observed in the OD and Skin. The muon ionises xenon atoms along the track, and a subsequent series of S1 and S2 pulses are produced. The pulses combine to produce a ‘tail’ as seen in ???. The first pulse in the

tail is used in the inter-detector timing selection. ?? highlights the effectiveness of the timing cut when used with the Outer Detector selection, reducing the event rate per day from 119.4 to 13.3.

TPC Energy Selection

The effects of cuts differ between simulations and data due to uncertainty associated with GEANT4 modelling of the muon shower, particularly at lower energies (see ??). As the cuts change, the rates change accordingly, but by different factors for data and simulations. This difference can be observed by examining the ratios of muon flux between the data and the simulations. Therefore, two decisions are required: which gain (high gain, HG, or low gain, LG) should be used for data to normalise the muon model, and what threshold is necessary for energy deposition in the TPC to ensure that the data can be accurately compared with the simulations. To accomplish this, we worked backward from the flux results described in ???. Comparing HG ratios of data rates to simulated rates with the corresponding LG ratios, as displayed in ??, it is clear that the LG ratios remain steady within statistical errors above 10 MeV. A steady ratio of data and simulation rates was one reason we measured the flux using LG data. A low threshold such as < 1 MeV, would not be useful because these events are not full muon events as seen through handscanning events around this threshold. They tend to be the end of a muon cascade in both data and simulations. ?? demonstrates this by showing that most events > 10 MeV are muons, not muon secondaries, as it separates the events with a muon passing through the TPC from those that do not. A higher energy threshold such as < 50 MeV would also not be useful, as ?? illustrates, the effect of PMT saturation can be seen above this energy in data.

Lack of Skin Threshold

This analysis considers a triple coincidence between the OD, Skin and TPC. However, all that is required in the Skin detector is a pulse which is classified as a ‘max pulse’ in LZap and the pulse must have an amplitude greater than 1 phd/ns. Due to the lack of calibration and subsequent energy reconstruction in the skin, it was chosen not to set an energy threshold in the Skin as it would be difficult to make comparisons to the energy-only simulations produced using BACCARAT.

5.2.3 NEST - Muon Simulations and Light-Energy Conversion

NEST has been used to reconstruct the energy of muons traversing the TPC. Typically particle interactions are observed through an ‘S1’ pulse and an ‘S2’ pulse and using methods described here [17], the energy deposit can be calculated. When a muon traverses the xenon space, a series of S1 and S2 pulses are produced which can not be separated in data. NEST, however, has a minimum ionising particle (MIP) module in which the linear energy transfer can be set. This is $2.1 \text{ MeV cm}^2 \text{ g}^{-1}$ [18] for a muon in liquid xenon with an average energy of 283 GeV. Two final inputs to NEST to simulate a muon are the start and end points of the track.

Sampling the start position and direction vector produced by the MUSUN simulations, all 48.2 million muons were projected down towards the LZ detector. If the muon passed through the xenon space, the point at which it entered and exited the TPC was saved and later used as the input into NEST. 67918 muons from BACCARAT passed through the xenon volume and were used towards producing the conversion between energy deposit and light produced due to the interaction. NEST outputs the energy, total S1 light and total S2 light and for the conversion used in this analysis, the S1 and S2 light was summed. Two separate sets of conversion factors were determined through running NEST with the WS2022 LZ detector configuration and WS2024 detector configuration. An example of the comparison can be seen in ???. A spline fit was fitted to the NEST data and extrapolated down to the minimum data point at 574 phd that passed our OD energy and inter-detector timing selection. The spline fit was then used to

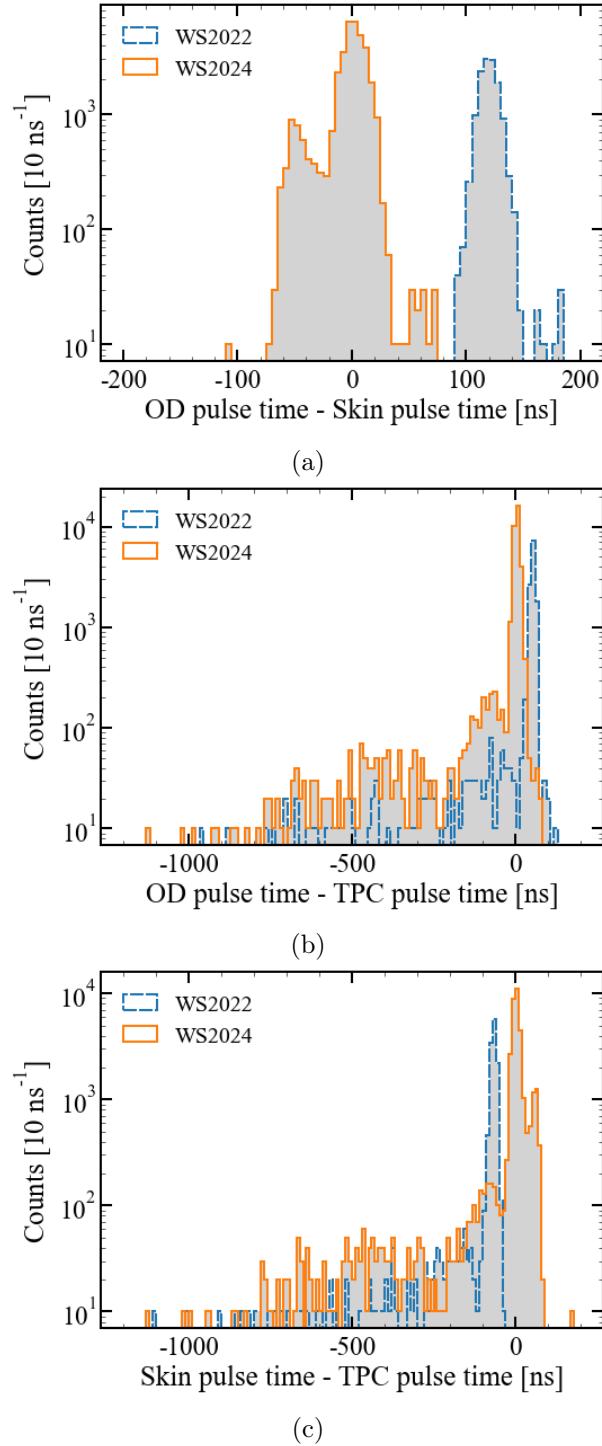


Figure 5.4: Inter-detector timing plots, these distributions were used in defining the timing selection following the skim of the WS2022 and WS2024 datasets looking at events with a pulse greater than 2000 phd in the Outer Detector. The figures depict (a) the time difference between the largest pulse in the Outer Detector and the largest pulse in the Skin; (b) the time difference between the largest pulse in the Outer Detector and the start of the muon tail in the TPC; and (c) the time difference between the largest pulse in the Skin and the start of the muon tail in the TPC.

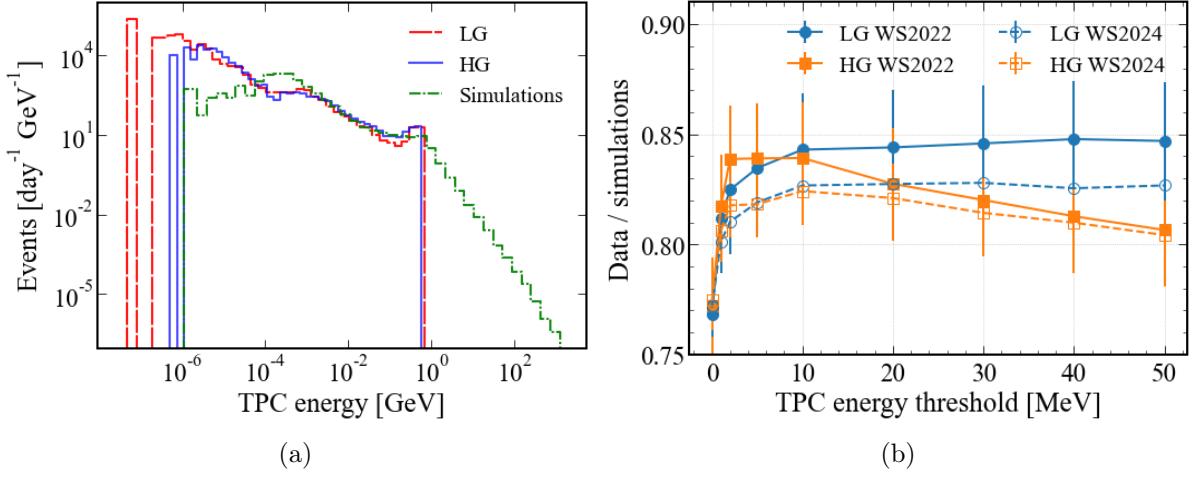


Figure 5.5: (a) Histograms comparing the spectrum of deposited energy in the TPC from BAC-CARAT muons with the spectrum of deposited energy in the TPC from WS2022 and WS2024 HG (LG) data (converted from TPC total pulse area). (b) The ratios of HG (LG) WS2022 and WS2024 data to BACCARAT simulations as a function of the energy threshold in the TPC.

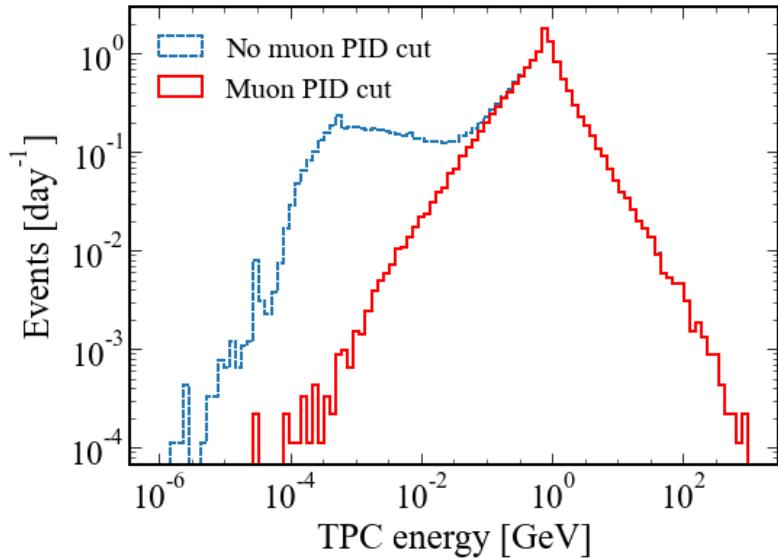


Figure 5.6: Energy deposition spectra of muon events from BACCARAT simulations with and without a muon particle ID (PID) cut. The cut selects events that had a muon depositing energy in the TPC directly.

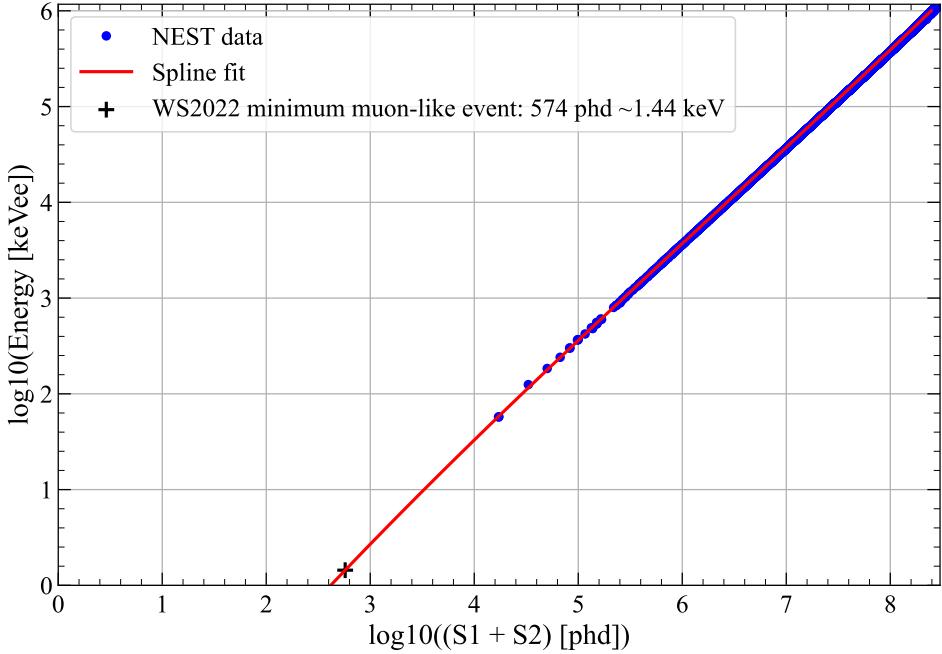


Figure 5.7: Deposited energy from BACCARAT simulations versus the number of detected photons from NEST simulations. The minimum muon-like event found in data is overlaid at 574 phd and the spline fit is extrapolated down to this value.

convert the Total Pulse Area observed in a muon event window to energy. To reproduce this conversion using the following parameters in a spline fit, SciPy's `interpolate.BSpline`.

WS2022 Parameters

```
t: [0.08294199 0.08294199 0.08294199 0.08294199 4.14709934 8.21125669
8.21125669 8.21125669 8.21125669]
c: [-2.70501286 -1.03606704  1.88505502  4.49568078  5.91208229  0.
0.          0.          0.          ]
k: 3
```

WS2024 Parameters

```
t: [0.08294199 0.08294199 0.08294199 0.08294199 4.14709934 8.21125669
8.21125669 8.21125669 8.21125669]
c: [-2.62556966 -0.95670208  1.96474676  4.57569338  5.99200743  0.
0.          0.          0.          ]
k: 3
```

where `t` is the knots, `c` is the spline coefficients and, `k` is the B-Spline degree factor. The spline fit was extrapolated down to 574 phd to obtain a corresponding energy deposition of 1.44 keV. This energy was used as our minimum threshold when determining the TPC threshold.

5.2.4 Comparison Between Data and Simulation

LZLAMA processed the BACCARAT files to convert energy depositions into S1s, S2s and other observed quantities that we could directly compare with the WIMP Search data. However, the decision was made to only use BACCARAT simulations for two reasons: Firstly, LZLAMA is not tuned for highly ionising tracks made by muons through the TPC. The spectrum of the total pulse area in the data was very different from that which LZLAMA produced from BACCARAT, in both the TPC and OD. Conversely, the BACCARAT energy spectra in the TPC and OD had a very comparable shape to that of the data. Secondly, LZLAMA does not account for

secondary particles such as pions and kaons. Initially, LZLAMA halted the processing of a file if it found an unrecognised particle ID in an event, and would move on to the next root file, which caused a great loss in the total number of events. The particle IDs for these error-triggering events were integrated into the LZLAMA source code. As a result, instead of encountering an error, a warning message was generated, quoting the particle ID, and the respective event was skipped over, thus ensuring that the remaining events in the file were not missed. A total of 767 events were missed causing slight bias since this cannot happen in data. In summary, the BACCARAT simulations were found to bear more resemblance to the spectra in data. They were also determined to have less systematic uncertainty than the LZLAMA data for processing muons and their secondary particles. Hence, we shall hereafter focus only on comparing the BACCARAT output to data.

Recording the photon production from muon interactions in the OD, skin and TPC would be too computationally intensive to use NEST as muons are highly ionising, so BACCARAT records only the energy deposited by a muon as it travels through the detector volumes. Consequently, to compare the muon flux of data and simulations, the conversion of pulse area to energy had to be determined. In previous investigations, the relationship between the photons detected by OD PMTs and the energy deposited in the OD was found to be linear at higher energies [19]. On this basis, our conversion factor was estimated by superposing the two spectra and matching the broad peaks of the WS2022 and WS2024 total pulse area spectra with the broad peak of the simulated energy spectrum (in energy units). The conversion factor for both WS2022 and WS2024 is 6.5×10^{-6} GeV/phd. The superposition is depicted in ?? where it can be seen how closely the different features of the data spectra (summed) line up with their simulated counterparts.

The cuts introduced in the simulations were designed to mimic those applied to data, namely a coincidence cut where a muon had to pass through each of the OD, skin and TPC within an event; and the 8 MeV energy threshold, previously explained in Section 4.1.

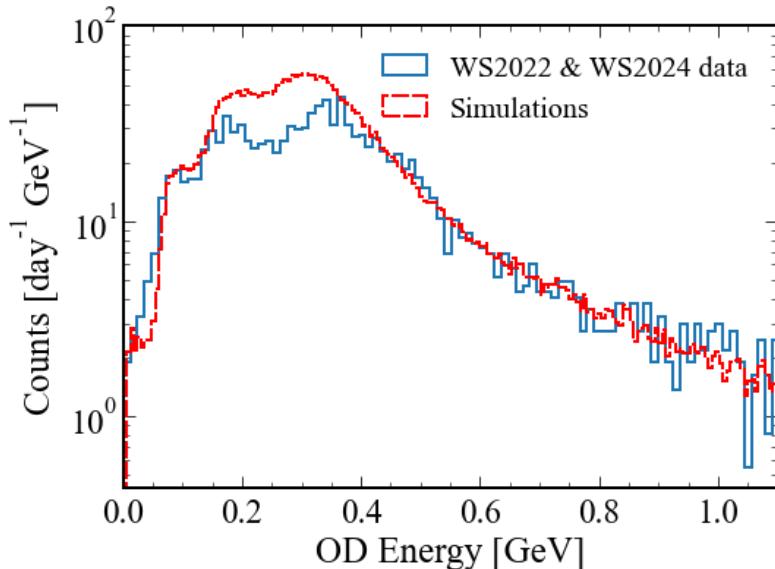


Figure 5.8: Energy deposited by BACCARAT simulated muons in the OD and the energy of events in data converted from the total pulse area in the OD.

After converting both TPC and OD pulse area to energy we were able to directly compare their combined energy maps between simulated results and data. The plots displaying the full energy distributions (after coincidence and OD energy cuts) and similar distributions after the TPC 10 MeV cut is imposed are in ?. Since we cannot distinguish between secondaries of

muons and actual muons, we need the cut to ensure that this systematic uncertainty, which is higher at lower energies, is avoided as much as possible. The energy distributions show that the 10 MeV cut removes events with a small amount of energy deposited in the TPC and a large amount of energy deposited in the OD. This is characteristic of muon secondaries or muons that may have skimmed the edge of the TPC.

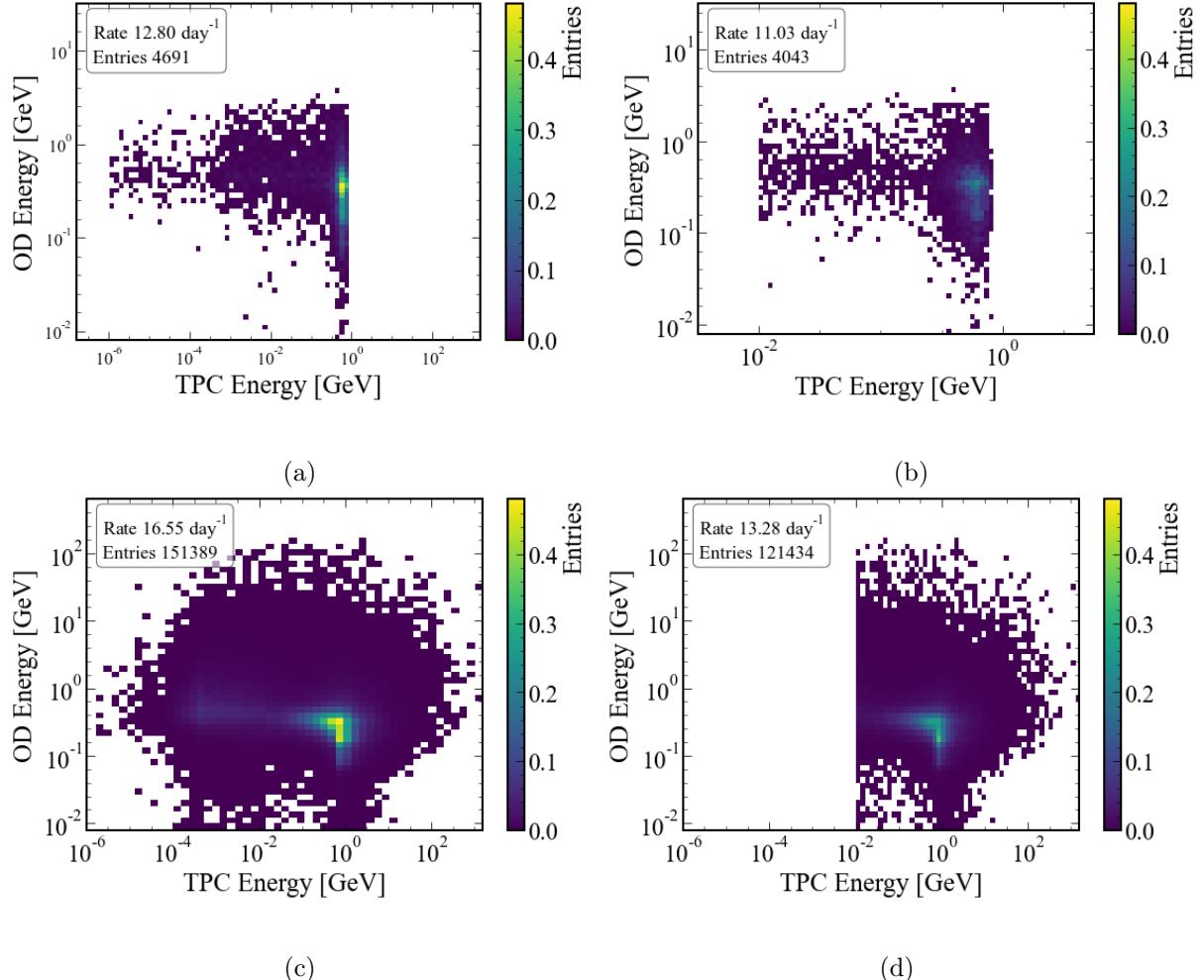


Figure 5.9: Energy depositions from events in the OD and TPC are shown, with the total pulse area from WS2022 and WS2024 data converted to energy. In (a), only OD and inter-detector timing cuts are applied, while in (b), the 10 MeV cut is also included. Similarly, BACCARAT simulations are plotted with the 8 MeV OD energy threshold and coincidence cut in (c), and with the additional 10 MeV cut included in (d).

5.2.5 Muon Rate Results

The measured muon rates from simulations, HG and LG data for WS2022 and WS2024 are presented in, respectively, ?? and ?? with increasing TPC energy thresholds. By varying the TPC energy threshold (translated from total pulse area via the NEST conversion) the change in the ratio of the simulated muon rate to the muon rate from data can be tracked. Columns 4 and 6 in each table show these ratios for LG and HG data respectively, and they are also displayed in ??b. The final column in each table lists the differences between these HG and LG ratios for every threshold. As described in Section 4.3, the steadiness of the LG ratio compared to the HG ratio above 10 MeV was the major factor in our decision to implement a 10 MeV threshold and use LG data for reconstructing the muon flux (see Section 9).

Table 5.2: Muon rates from simulations and WS2022 data. The data has the OD cuts and timing selections applied. The simulations have a 3-fold detector coincidence criteria and the 8 MeV OD energy cut applied. The TPC energy thresholds represent the total energy deposited by a muon event in the TPC.

TPC threshold	Simulation rate [day ⁻¹]	Data LG rate [day ⁻¹]	Data/sims LG ratio	Data HG rate [day ⁻¹]	Data/sims HG ratio	LG ratio - HG ratio
TPC > 1.44 keV	16.550 ± 0.043	12.709 ± 0.359	0.768 ± 0.022	12.780 ± 0.360	0.772 ± 0.022	-0.0
TPC > 1 MeV	15.055 ± 0.041	12.224 ± 0.352	0.812 ± 0.023	12.305 ± 0.353	0.817 ± 0.024	-0.0
TPC > 2 MeV	14.477 ± 0.040	11.940 ± 0.348	0.825 ± 0.024	12.143 ± 0.351	0.839 ± 0.024	-0.0
TPC > 5 MeV	13.760 ± 0.039	11.485 ± 0.341	0.835 ± 0.025	11.546 ± 0.342	0.839 ± 0.025	-0.0
TPC > 10 MeV	13.275 ± 0.038	11.191 ± 0.337	0.843 ± 0.025	11.141 ± 0.336	0.839 ± 0.025	0.0
TPC > 20 MeV	12.840 ± 0.037	10.837 ± 0.331	0.844 ± 0.026	10.625 ± 0.328	0.827 ± 0.026	0.0
TPC > 30 MeV	12.597 ± 0.037	10.655 ± 0.328	0.846 ± 0.026	10.331 ± 0.323	0.820 ± 0.026	0.0
TPC > 40 MeV	12.413 ± 0.037	10.524 ± 0.326	0.848 ± 0.026	10.088 ± 0.320	0.813 ± 0.026	0.0
TPC > 50 MeV	12.258 ± 0.037	10.382 ± 0.324	0.847 ± 0.027	9.886 ± 0.316	0.807 ± 0.026	0.0

Table 5.3: Muon rates from simulations and WS2024 data. See the ?? caption for descriptions of the applied cuts and energy thresholds.

TPC threshold	Simulation rate [day ⁻¹]	Data LG rate [day ⁻¹]	Data/sims LG ratio	Data HG rate [day ⁻¹]	Data/sims HG ratio	LG ratio - HG ratio
TPC > 1.44 keV	16.550 ± 0.043	12.769 ± 0.218	0.772 ± 0.013	12.821 ± 0.219	0.775 ± 0.013	-0.0
TPC > 1 MeV	15.055 ± 0.041	12.063 ± 0.212	0.801 ± 0.014	12.137 ± 0.213	0.806 ± 0.014	-0.0
TPC > 2 MeV	14.477 ± 0.040	11.730 ± 0.209	0.810 ± 0.015	11.838 ± 0.210	0.818 ± 0.015	-0.0
TPC > 5 MeV	13.760 ± 0.039	11.270 ± 0.205	0.819 ± 0.015	11.259 ± 0.205	0.818 ± 0.015	0.0
TPC > 10 MeV	13.275 ± 0.038	10.975 ± 0.203	0.827 ± 0.015	10.941 ± 0.202	0.824 ± 0.015	0.0
TPC > 20 MeV	12.840 ± 0.037	10.624 ± 0.199	0.827 ± 0.016	10.542 ± 0.198	0.821 ± 0.016	0.0
TPC > 30 MeV	12.597 ± 0.037	10.430 ± 0.197	0.828 ± 0.016	10.258 ± 0.196	0.814 ± 0.016	0.0
TPC > 40 MeV	12.413 ± 0.037	10.246 ± 0.196	0.825 ± 0.016	10.052 ± 0.194	0.810 ± 0.016	0.0
TPC > 50 MeV	12.258 ± 0.037	10.134 ± 0.195	0.827 ± 0.016	9.858 ± 0.192	0.804 ± 0.016	0.0

Table 5.4: Muon rates from simulations, WS2022 and WS2024 data. See the ?? caption for descriptions of the applied cuts and energy thresholds.

TPC threshold	Simulation rate [day ⁻¹]	Data LG rate [day ⁻¹]	Data/sims LG ratio
TPC > 1.44 keV	16.550 ± 0.043	12.753 ± 0.187	0.771 ± 0.011
TPC > 1 MeV	15.055 ± 0.041	12.106 ± 0.182	0.804 ± 0.012
TPC > 2 MeV	14.477 ± 0.040	11.787 ± 0.179	0.814 ± 0.013
TPC > 5 MeV	13.760 ± 0.039	11.328 ± 0.176	0.823 ± 0.013
TPC > 10 MeV	13.275 ± 0.038	11.033 ± 0.174	0.831 ± 0.013
TPC > 20 MeV	12.840 ± 0.037	10.681 ± 0.171	0.832 ± 0.014
TPC > 30 MeV	12.597 ± 0.037	10.490 ± 0.169	0.833 ± 0.014
TPC > 40 MeV	12.413 ± 0.037	10.321 ± 0.168	0.832 ± 0.014
TPC > 50 MeV	12.258 ± 0.037	10.201 ± 0.167	0.832 ± 0.014

5.2.6 Reconstruction of the Muon Flux

Reconstruction of the muon flux from the measured muon rate has been done simply by scaling the simulated muon flux by the same ratio as that for the measured-to-simulated rate:

$$F_m = F_s \times \frac{R_m}{R_s}, \quad (5.1)$$

where F_m and F_s are the measured (reconstructed from the measured rate) and simulated muon fluxes through a surface of a sphere (unit detection efficiency at all angles), respectively, and R_m and R_s are the measured and simulated muon rates through the detector.

This simple scaling is based on the common inputs to the simulation of the muon flux and muon rate, namely, muon energy spectra and angular distributions at SURF. It also relies on the assumption that muon transport through the detector and detector response are correctly simulated in the LZ simulation framework BACCARAT (based on GEANT4 toolkit). Simulating muon-induced cascades and their development in and outside the detector has associated uncertainties that are difficult to estimate. By requiring the energy deposition in the TPC to be greater than 10 MeV, we effectively select events when a muon passes through the TPC and remove a relatively small contribution of events without a muon, when only low-energy secondary particles enter the TPC. No other energy cut is included so all events with a muon in the TPC (both in data and simulations) are considered in the analysis. Simulation of muon track passing through the detector and muon energy loss from ionisation along the track should be handled accurately by GEANT4. The second assumption about an accurate simulation of detector response becomes non-critical because the rate of events does not include information about energy deposition apart from the energy threshold. The ratio of measured-to-simulated rates does not depend much on the energy threshold around and above 10 MeV and the small difference in this ratio with changing the threshold can serve as an estimate of the systematic uncertainty.

A very low energy threshold is used for the skin, only to remove noise events. Similarly, energy threshold and pulse shape analysis in the OD only remove radioactive background and noise events leaving muon events intact. Hence, we do not anticipate any uncertainty linked to the thresholds in the skin or the OD.

The ratio of measured-to-simulated muon rates for low gain in ?? remains constant within statistical uncertainty for TPC energy thresholds of 5-50 MeV. For the TPC threshold of 10 MeV, this ratio is 0.831 ± 0.013 (stat.) ± 0.008 (syst.), where systematic uncertainty is estimated from the change in the ratio with changing the threshold by a factor of 2-3. (Note that reducing the threshold will increase the fraction of events when a muon does not enter the detector thus increasing the dependence on simulation details).

Using ??, the ratio of 0.831 ± 0.013 (stat.) ± 0.008 (syst.) and the simulated muon flux of $6.16 \times 10^{-9} \text{ cm}^{-2} \text{ s}^{-1}$, we derive the reconstructed muon flux from the rate measurements as $(5.119 \pm 0.080 \text{ (stat.)} \pm 0.049 \text{ (syst.)}) \times 10^{-9} \text{ cm}^{-2} \text{ s}^{-1}$. This flux agrees well with the measurement in the nearby cavern reported in Ref. [6] $((5.31 \pm 0.17) \times 10^{-9} \text{ cm}^{-2} \text{ s}^{-1})$.

5.2.7 Evaluation of Average Rock Density

The difference in the reconstructed and simulated muon fluxes is primarily due to a different density of rock compared with the initial muon model. We assume here that the surface profile is known with sufficient accuracy and so is the laboratory position. There is a small dependence of the flux on the position within the laboratory but it is below the statistical uncertainty. The dependence of the flux on the rock composition has the second-order effect. If we attribute the lower muon flux in the measurements to the higher average rock density, we can evaluate what the realistic average rock density above and around the LZ location is.

We have calculated the muon flux at SURF with different densities of rock above the laboratory to match the measured (reconstructed from measurements) value. In this case, we used a simple approximation of a ‘flat’ surface profile above the lab taking into account only the Earth curvature (though with negligible effect for the flux at this depth). The simulated muon flux is matched with the measured one assuming the average density of $(2.78 \pm 0.01) \text{ g/cm}^3$ which is 3.0% higher than initially assumed in the muon model (see ?? for the flux dependence on the rock density). The statistical uncertainty of the muon rate measurements dominates the error here. This density is smaller than that reported in Ref. [6] $((2.89 \pm 0.06) \text{ g/cm}^3)$ despite almost the same muon flux. The muon flux has been reconstructed for the place where the measurements of the rate were carried out, namely in the Davis cavern. The flux depends on the position in and around the cavern. For example, the muon flux 7 m above the cavern from where most muons in the simulations originated will be about 3% higher.

5.3 Modulation

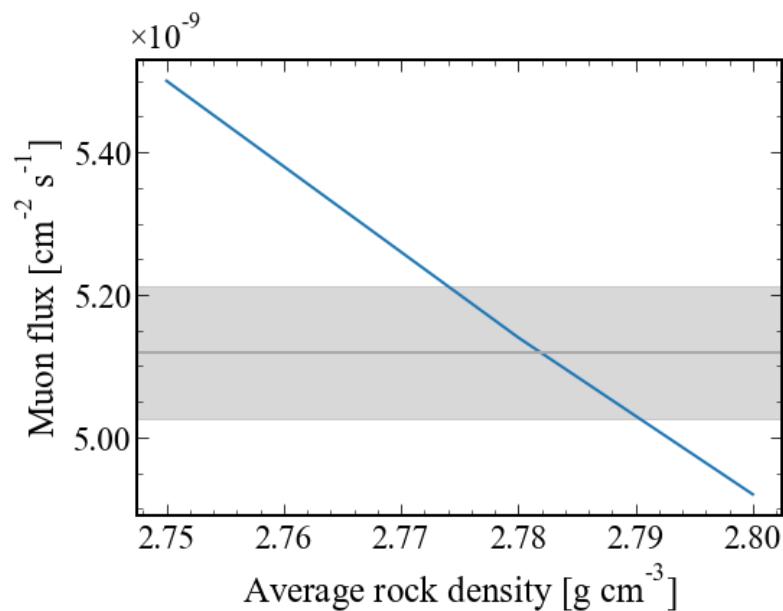


Figure 5.10: The muon flux dependence on average rock density within the muon model. The horizontal line and shaded section mark, respectively, the measured flux value and the combined statistical and systematic errors, $(5.119 \pm 0.094) \times 10^{-9} \text{ cm}^{-2}\text{s}^{-1}$.

6 Conclusion

References

- [1] D. S. Akerib et al. “The LUX-ZEPLIN (LZ) experiment”. In: *Nucl. Instrum. Meth.* A953 (2020), p. 163047. ISSN: 0168-9002. DOI: 10.1016/j.nima.2019.163047. arXiv: 1910.09124.
- [2] D. S. Akerib et al. “Identification of Radiopure Titanium for the LZ Dark Matter Experiment and Future Rare Event Searches”. In: *Astropart. Phys.* 96 (2017), pp. 1–10. DOI: 10.1016/j.astropartphys.2017.09.002. arXiv: 1702.02646 [physics.ins-det].
- [3] B.J. Mount et al. *LUX-ZEPLIN (LZ) Technical Design Report*. 2017. eprint: arXiv: 1703.09144.
- [4] D. S. Akerib et al. “Simulations of events for the LUX-ZEPLIN (LZ) dark matter experiment”. In: *Astroparticle Physics* 125 (Feb. 2021), p. 102480. ISSN: 0927-6505. DOI: 10.1016/j.astropartphys.2020.102480. arXiv: 2001.09363.
- [5] M. L. Cherry et al. “Multiple muons in the Homestake underground detector”. In: *Phys. Rev., D* 27 (7 Apr. 1983), pp. 1444–1447. DOI: 10.1103/PhysRevD.27.1444.
- [6] N. Abgrall et al. (Majorana). “Muon flux measurements at the Davis campus of the Sanford Underground Research Facility with the Majorana Demonstrator veto system”. In: *Astropart. Phys.* 93 (2017), pp. 70–75. DOI: 10.1016/j.astropartphys.2017.01.013.
- [7] B. Abi et al. “Deep Underground Neutrino Experiment (DUNE), far detector technical design report, Volume II: DUNE physics.” In: *arXiv* (Feb. 2020). DOI: 10.48550/arXiv.1512.06148. arXiv: 2002.03005 [hep-ex].
- [8] P. Antonioli et al. “A three-dimensional code for muon propagation through the rock: MUSIC”. In: *Astropart. Phys.* 7 (1997), p. 357. DOI: 10.1016/s0927-6505(97)00035-2.
- [9] V. A. Kudryavtsev. “Muon simulation codes MUSIC and MUSUN for underground physics”. In: *Computer Physics Communications* 180 (2009), p. 339. DOI: 10.1016/j.cpc.2008.10.013.
- [10] D.M. Mei et al. “Early results on radioactive background characterization for Sanford Laboratory and DUSEL experiments”. In: *Astropart. Phys.* 34 (2010), p. 33. DOI: 10.1016/j.astropartphys.2010.04.003.
- [11] C. Zhang. private communication. 2014.
- [12] J. Heise. Preprint: arXiv:1401.0861v1. 2014.
- [13] M. Richardson. PhD thesis (University of Sheffield). 2015.
- [14] T. Gaisser. *Cosmic rays and particle physics*. Cambridge University Press, 1990.
- [15] V. A. Kudryavtsev, E. V. Korolkova, and D. Woodward. *Muon simulations for LZ using MUSIC and MUSUN*. LZ note. 2015.
- [16] LUX-ZEPLIN Collaboration. *SR1 Core Cuts in ALPACA*. <https://luxzeplin.gitlab.io/docs/softwaredocs/analysis/analysiswithrqs/alpaca/tools/SR1CoreCuts.html#muon-hold-off>. Accessed: 2023-02-13.
- [17] Matthew Szydagis et al. “A Review of Basic Energy Reconstruction Techniques in Liquid Xenon and Argon Detectors for Dark Matter and Neutrino Physics Using NEST”. In: *Instruments* 5.1 (2021), p. 13. ISSN: 2410-390X. DOI: 10.3390/instruments5010013. URL: <http://dx.doi.org/10.3390/instruments5010013>.
- [18] Particle data group. *Muons in xenon liquid (Xe)*. data table.

[19] E. Fraser. *OD Commissioning Energy Calibration*. 2021.