

University of Sheffield

Robert. J. Neal

Searches for Dark Matter Annual Modulation With COSINE-100

Department of Physics & Astronomy
The University of Sheffield

Supervisor: Prof. Neil Spooner

This dissertation is submitted for the degree of
Doctor of Philosophy

2023

Searches for Dark Matter Annual Modulation With COSINE-100

Robert. J. Neal

Supervisor: Prof. Neil Spooner

Department of Physics & Astronomy
The University of Sheffield

This dissertation is submitted for the degree of
Doctor of Philosophy

September 2023



**University of
Sheffield**

Abstract

The COSINE-100 dark matter detection experiment is an NaI-based dark matter detection experiment with the primary goal of providing a model-independent test of the long-standing but controversial result of a dark matter-induced annual modulation signal, reported by the DAMA collaboration. As a model-independent test of the DAMA result, the target material and search method at COSINE-100 is the same as that in the DAMA experiments, that is searching for an annual modulation signal in the rate of signal-like events in NaI crystals. Results of searches with the first 2.8 years of COSINE-100 data have already been published, but with an insufficiently large exposure to conclusively rule on the DAMA result.

Presented in this thesis are the results of a new search for a dark matter-induced annual modulation signal at COSINE-100, using the first 4.7 years of data collected. This is the first time these data have been analysed. An improved analysis procedure is used in this search and this procedure is also described. These improvements from the method used in previous searches include an updated method of accounting for detector dead-time, folding this parameter into the fit model, and also from a updated background model in light of issues found with the treatment of surface ^{210}Pb in previous analyses. Using this new procedure, best-fit annual modulation amplitudes of 0.0040 ± 0.0034 dru (0.0047 ± 0.0041 dru) in the 1-6 keV (2-6 keV) energy regions in a fixed-phase analysis are found. These values are smaller than those reported in previous COSINE-100 analyses, but still statistically consistent with both the DAMA and no-modulation hypotheses. Similar results are found in the floated-phase case. Also presented are improvements in the process of validating the analysis procedure using simulated data, as well as a description of the long-term monitoring and resulting corrections of the liquid scintillator veto at COSINE-100, and an investigation into an intermittent type of PMT noise in the detector.

Author's Contribution

The work in this thesis is original and has been carried out by the author, except where specific references are made. All plots and values presented are the work of the author unless otherwise credited. The analysis work presented has not been submitted for prior assessment for any other degree or qualification at any other university. A small amount of the introductory content in Chapter 1 is based on extracts from *Dark Matter – Evidence, Main Candidates, and Cross Section of Annihilation from Relic Abundance*, submitted in the 2017/18 academic year as part of the third year undergraduate Physics Research Project module, supervised by Prof. Leszek Rozskowski. The author also worked on the COSINE-100 experiment in the academic year prior to the beginning of the PhD programme as a master's student, supervised by Prof. Neil Spooner, studying the rate of ^{40}K side-band events and monitoring the liquid scintillator (LS) gain using framework developed by JH Jo. During this master's project (titled *Analysis of the annual modulation of ^{40}K decay events in the COSINE-100 dark matter detector*), no corrections were made to the LS and hence all the corrections to the LS from July 2018 onward, described in Chapter 5, are original work.

The work described in this thesis exists as part of the wider body of work carried out by the COSINE-100 collaboration, large parts of which the author was not involved in. The parts where the author had a direct contribution are as follows:

The author was responsible for the monitoring of the behaviour of the LS veto system at COSINE-100 from September 2018 onwards, described in Chapter 5. The monitoring and gain correction applied to the LS prior to July 2018 was carried out by JH. Jo. This includes the development of the phenomenological function Eq. 4.2 used to characterise the location of the 1460 keV Compton edge. The initial energy calibration of the LS was carried out by multiple collaboration members prior to the author's involvement. All further monitoring (with the exception of the standard weekly monitoring described in Chapter 4) and corrections to the LS described in this work were carried out by the author using a new framework developed by the author. This includes the time-dependent gain correction, the characterisation of the behaviour of the LS PMTs, the identification of a fault in one of the PMTs, leading to its removal from the analysis, and the subsequent reworking of analysis techniques that resulted.

In Chapter 6, the BDTA variable and related parameters described were developed prior to the author's involvement in the collaboration. Using the BDTA variable, the author developed event selection for selecting Gaussian noise events that was used by the author to study the period of increased event-rate in C2 described in the Chapter This process was also used to check the rest of the data-set for Gaussian noise events. The author began a study attempting to build on the already established BDTA variable and create a parameter able to sufficiently reject Gaussian noise events such that regions of

high Gaussian noise could be used in dark matter searches.

The annual modulation fitting procedure described in Chapters. 7, 8, and 9 has been developed iteratively over the various annual modulation searches at COSINE-100, most recently by WT Thompson with help from EB de Souza for the second annual modulation search [1] and by JH Jo in the first [2]. In all cases this uses the analysis on the *COSINE Analysis Toolkit* package developed by YJ Ko. The need for an updated treatment of the background components in the fit model was uncovered by the author as part of the process of submitting the SET3 annual modulation search results for publication [1]. This occurred after a referee for the paper requested a comparison between the best-fit values for background activities in the annual modulation search and the previously published model-dependent WIMP search [3] which was carried out by the author. Also during this refereeing process the concern about the method of scaling the measured event-rate to assume 100% live-time was raised. In the latter case, the author developed a new annual modulation search procedure where the detector live-time was included in the fit model. This procedure was applied to the SET3 data and was shown to correctly reproduce the expected results in line with the referee’s request.

The study of surface alpha events used to estimate the extended effective half-life of the surface ^{210}Pb was carried out by Hyunseok Lee and were then used by the author in the development of the new annual modulation search procedure with split surface and bulk ^{210}Pb component, described in Chapter 7.

The procedure of generating simulated (“pseudo”)-data described in Chapter 8 was developed iteratively through the successive generations of annual modulation searches at COSINE-100. A new method of developing pseudo-data was developed by the author within which the measured detector live-time and the split surface and bulk ^{210}Pb components were incorporated. The tests of the fixed-phase modulation bias, pull RMS, and χ^2 values described in Chapter 8 of the fit results of pseudo-data ensembles were based on work by WT Thompson, but expanded upon by the author for all other tests described.

The check of consistent detector performance using ^{40}K events, described in Chapter 9 was solely the work of the author. This includes a similar check of detector performance described in the SET3 analysis paper [1]. The event selection and data quality checks described in Chapter 9 were predominantly not the work of the author and were mostly developed prior to the SET3 analysis. The data quality check was carried out by SM Lee, but incorporated a cross-check using the method of identifying Gaussian noise developed by the author and described in Chapter 6. This produced an initial “good run list” from which the author selected the final data-set used in the analysis.

The Boosted Decision Tree (BDT) variable used as the primary means of selecting signal events in the SET3 analysis and this analysis was developed with significant contributions from YJ Ko, based on earlier work by CH Ha and G Adhikari. A new BDT variable is being developed with the goal of lowering the COSINE-100 low-energy threshold to 0.75 keV using data from an ^{22}Na calibration run. This work is primarily being carried out by GH Yu, but was initially a joint effort with the author. The use of pseudo-data to judge the compatibility of the floated-phase annual modulation results with the no-modulation hypothesis was developed by the author, whereas the conventions for reporting the fixed-phase results were developed in prior work by WT Thompson and JH Jo.

From October 2020 to September 2022, the author was permanently based at the COSINE collaboration’s headquarters at the Centre for Underground Physics (CUP), at the Institute for Basic Science (IBS) in Daejeon, South Korea. During this time, the author was more heavily involved in the monitoring of detector performance, including multiple

investigations of high-noise regions flagged by collaboration members during standard weekly shifts. This includes regions of noise flagged by the author. The author was also jointly responsible with GH Yu for the initial studies of the data from the ^{22}Na calibration run for use in an improved analysis, including measuring selection efficiency, developing event selection for signal events, studying the behaviour of the signal events selected, and developing the early framework used in the training of the new BDT variable. Finally, the author also carried out a short study to ensure that changes in event selection resulting from increased PMT noise in some of the crystal PMTs had not had unintended effects on the tagging of single and multi-hit events in the crystal, as described in Chapter 9.

As a COSINE collaboration member, the author has appeared as an author on 14 papers released by the collaboration over the duration of the PhD programme. Aside from feedback to the main authors around wording etc, the author's main contribution to the analysis in these papers are some of the work on the LS monitoring and gain correction described in [4], and detector stability checks and fit model validation in the SET3 analysis paper [1]. The author also presented talks and posters at various conferences on behalf of the collaboration. This includes multiple talks at the bi-annual Korean Physical Society meetings for which the author won a "Best Presentation Award" in 2021 for a presentation on annual modulation searches at COSINE-100, presenting a talk on the overall status of the COSINE-100 experiment in a parallel session of the ICHEP 2022 conference and presenting a poster on an induced modulation signal in COSINE-100 data using the DAMA method [5] on behalf of the primary author, H Prihtiadi who was unable to attend, and presenting a talk about the overall status of the COSINE-100 experiment at a DMUK meeting in 2022.

Acknowledgements

Firstly, I must thank my supervisor, Neil, both for all his sage guidance and support over the last five years, going back to the start of my master's project, and also for creating the opportunity for me to undertake this work in the first place. The chance to not only do a PhD, but spend two years living and working in South Korea at CUP on an experiment like COSINE-100 is something I am very grateful for, and something I doubt would have been possible without Neil's experience and vision. Likewise, I have many things to thank Hyunsu for, from agreeing to my working at CUP and helping facilitate my time there, to the generous supervisory role he provided while I was at CUP and beyond, and more generally, for his leadership and co-ordination of the COSINE-100 experiment. My thanks also go to Mark for the wisdom and advice imparted in his role as my personal tutor.

I am also extremely grateful to all the other people who so kindly gave their time, energy, and patience to helping me all the skills and knowledge I would need to carry out this work. Without their support, this work would almost certainly have been impossible. My deepest thanks to Will for leaving the modulation analysis in such a good place for me to take over and supporting me through that handover process, and also for humouring and helping answer my various questions about all things COSINE throughout my time on the experiment, from how to carry various analysis tasks to life in Korea. To Young Ju for creating so many of the tools that would end up being crucial in my work, and no doubt the work of most of the analysers at COSINE-100, then for helping me actually understand and use those tools in my work. To Jay for so patiently helping me get started on COSINE-100 analysis work when I barely knew what a ROOT file was. I'm not sure how I would have ever gotten anywhere without that early support.

As soon as I met and started working with the people in the HEP group at Sheffield, it was clear what a friendly and supportive group of people it was. To that end, I would like to thank Andrew and Anthony for their wise words and guidance about all things PhD and also for help with Sheffield's annual COSINE monitoring shift work. I would also like to thank Warren for taking the time to help getting me set up at the very beginning of my research, even as he was at the very opposite, and far more stressful end of his PhD programme.

Moving to Korea for two years having never previously left Europe was certainly a daunting proposition, let alone during a pandemic, so I am very grateful for how welcome I was made to feel and to everyone who helped me navigate what was initially a very confusing new world. Everyone I worked with at CUP played a part in that positive experience and I am grateful to them all. I must specifically thank Eunju, Chang Hyon, and Doug for sharing their wealth of knowledge with me on various topics, to Kyungwon, Insoo, SungHyun, and Hafizh for the support on production and LS-related issues, to Young Sun for all her help with getting set up in Korea and with various administrative

tasks throughout my time in Korea, and to Gyunho, Byung Ju, Jae Jin, Seung Mok, and the rest of the PhD students, post-docs, and everyone else at CUP who helped create such a positive environment. I must also specifically thank Gyunho for so immediately being such a good friend and helping me find my feet in Korea, both with work and beyond, then continuing to navigate me through some of the more esoteric facets of life in Korea as a friend, guide, and translator, not to mention his excellent cooking. To everyone in Korea who helped make my time there what it was, 감사합니다.

Last but definitely not least, I feel extremely lucky and am extremely grateful for the immeasurable support I've received from my friends across the world, as well as my family throughout the PhD. Trying to list everyone who has helped me in one way or another would be impossible. In particular though I am grateful beyond words to my amazing partner, Michelle, for her endless love and support. Finally to my parents also for their unconditional love, and for being a source of guidance and inspiration my entire life.

Contents

1	Introduction	1
2	Dark matter: Evidence and Candidates	3
2.1	Astronomical observations of the effects of dark matter	4
2.1.1	Early observations	4
2.1.2	Galaxy rotation curves	5
2.1.3	Gravitational lensing	6
2.1.4	Dark matter in cosmology	8
2.2	Dark matter candidates	10
2.2.1	Macroscopic candidates	10
2.2.2	Particle candidates	12
2.3	Concluding remarks	20
3	Experimental Searches for Dark Matter and the DAMA Controversy	21
3.1	Dark matter searches with lowest model-dependent upper limits on WIMPs	22
3.1.1	Liquid Xe experiments	22
3.1.2	Lower mass WIMP searches	25
3.2	The DAMA result and tests from other NaI experiments	28
3.2.1	The DAMA experiments and results	28
3.2.2	Results from other NaI-based experiments	32
3.3	Alternative explanations for the DAMA result	34
3.3.1	DAMA modulation signal from modulating backgrounds?	35
3.3.2	An annual modulation signal induced by DAMA's analysis technique?	36
3.3.3	DAMA more sensitive than other detectors?	38
3.4	Concluding remarks	39
4	The Design and Operation of the COSINE-100 Experiment	40
4.1	The Yangyang Underground Laboratory	40
4.2	The COSINE-100 detector design	41
4.2.1	Motivating design principles of the detector	42
4.2.2	The NaI crystal array and inner passive shielding	43
4.2.3	Muon shielding	45
4.2.4	Liquid scintillator	46
4.2.5	Data acquisition system	49
4.3	Detector operation	51
4.3.1	Monitoring of environmental variables	52
4.3.2	Monitoring shift work	53
4.3.3	Detector component monitoring and gain correction	54

4.4	Key differences between the COSINE-100 and DAMA experiments' setup and analysis technique	56
4.5	Concluding remarks	58
5	Long Term Stability of the COSINE-100 Liquid Scintillator Veto System	59
5.1	Gain correction in the LS	59
5.1.1	Initial LS gain correction	60
5.1.2	LS gain correction for the first three years of data taking	61
5.1.3	Second phase of LS gain correction after October 2020	65
5.2	PMT-level studies of LS behaviour	68
5.2.1	Changing PMT event rates	68
5.2.2	Studies of unusual behaviour in PMT18	74
5.2.3	LS energy calibration and threshold without PMT18	77
5.3	Concluding remarks	79
6	Investigation and Mitigation of Intermittent PMT Noise	81
6.1	Initial identification and rejection of Gaussian noise in early data	81
6.2	Recurrence of Gaussian noise	83
6.2.1	Investigating a new period of high event rate	84
6.2.2	Properties of the new high-rate events	85
6.3	Concluding remarks	91
7	Development of the Third Generation COSINE-100 Annual Modulation Fitting Procedure	92
7.1	Overview of the second generation fitting procedure	93
7.1.1	SET3 data preparation	93
7.1.2	The background model used in the three year annual modulation search	95
7.1.3	Recovering best-fit modulation parameters	98
7.2	Development of the SET4 fitting procedure	99
7.2.1	Live-time scaling	100
7.2.2	Accounting for surface ^{210}Pb effects	102
7.3	Concluding remarks	108
8	Pseudo-Data Validation of Analysis Techniques	110
8.1	The motivation for pseudo-data analysis	110
8.2	Production of the pseudo-data	111
8.3	Testing the fixed-phase modulation analysis procedure on the pseudo-data	114
8.3.1	Modulation amplitude bias check	115
8.3.2	Modulation amplitude uncertainty check	117
8.3.3	Goodness of fit check with chi-square values	118
8.4	Investigating the issue with the reproduction of background activities in the SET4 fitting method	119
8.4.1	Locating the source of the issue in the fit method	120
8.4.2	Tests of individual high- χ^2 fits	123
8.5	A novel use of pseudo-data in reporting the floated-phase annual modulation search result	125
8.5.1	Measuring the modulation amplitude bias in floated-phase analyses of low modulation pseudo-data	125

8.5.2	Accounting for the floated-phase low-amplitude bias	127
8.6	Concluding remarks	128
9	Searches for Annual Modulation in the SET4 COSINE-100 Physics Data	129
9.1	Detector stability	129
9.1.1	Ensuring consistent detector performance	129
9.1.2	Environmental monitoring	130
9.2	Event selection	131
9.2.1	Event selection and noise rejection	131
9.2.2	SET4 data selection and quality checks	142
9.3	Fixed-phase annual modulation search results	143
9.3.1	Calculation of the best-fit modulation amplitudes	144
9.3.2	Interpretations of the results	148
9.4	Floated-phase annual modulation search results	149
9.5	Concluding remarks	151
10	Conclusion and Outlook	153
A	Further tests of describing background components using the split-²¹⁰Pb fit method	167
A.1	Investigating the performance of the fit of the surface ²¹⁰ Pb half-life	167
A.1.1	Interpretations of these results	171

List of Figures

2.1	Rotation curves for the M31 galaxy	6
2.2	The Bullet Cluster	7
2.3	CMB power spectrum	9
2.4	Mass scale of dark matter candidates	11
2.5	Candidate PBH lensing event	12
2.6	Constraints on sterile neutrinos	15
2.7	WIMP freeze-out	17
2.8	WIMP parameter space	19
3.1	Dark matter experiments	22
3.2	LZ TPC	23
3.3	LZ spin-independent WIMP limits	25
3.4	WIMP limits summary	26
3.5	CRESST-III spin-independent WIMP limits	27
3.6	WIMP wind	29
3.7	Expected annual modulation event-rate	30
3.8	DAMA experimental setup	31
3.9	DAMA annual modulation result	32
3.10	COSINE-100 model-dependent WIMP limits	33
3.11	Simulated DAMA residuals	36
3.12	COSINE-100 results with DAMA residuals method	37
4.1	The Yangyang Underground Laboratory	41
4.2	The COSINE-100 experiment	42
4.3	COSINE-100 crystals schematic	44
4.4	COSINE-100 plastic scintillator panels	46
4.5	COSINE-100 LS calibration energy spectrum	48
4.6	COSINE-100 LS trigger rate	48
4.7	COSINE-100 DAQ	50
4.8	Temperature over time at COSINE-100	53
4.9	FADC rate in COSINE-100 muon event	54
4.10	COSINE-100 crystal gain tracking fits	55
4.11	Pulse shape discrimination parameter	56
4.12	COSINE-100 muon rate over time	58
5.1	Early LS gain	61
5.2	Early LS energy spectra	61
5.3	^{40}K Compton edge fit	63
5.4	LS gain before correction July 2018 to October 2020	64
5.5	LS gain after correction July 2018 to October 2020	65

5.6	LS gain from October 2016 to June 2022 without PMT18	67
5.7	LS gain without PMT18 before final gain correction	67
5.8	Final LS gain without PMT18 after final gain correction	68
5.9	Single LS PMT energy spectrum	69
5.10	LS PMT energy spectra at different times	70
5.11	LS PMT event-rate	71
5.12	LS PMT noise-rate	73
5.13	Total LS event-rate no PMT-18	74
5.14	Changing LS PMT18 spectra over time	75
5.15	Changing LS PMT18 spectra single plot	76
5.16	LS PMT18 event rate	76
5.17	^{60}Co calibration run LS energy spectrum	78
6.1	Typical signal and noise waveforms	82
6.2	Increased counts in C2 from Gaussian noise	84
6.3	Increased signal-rate in C2 from Gaussian noise	85
6.4	BDT/BDTA distribution of Gaussian noise events	86
6.5	Energy/BDTA distribution of Gaussian noise events	87
6.6	C2 Gaussian noise event-rate	88
6.7	Gaussian noise waveforms	89
6.8	Borderline Gaussian noise waveforms	89
6.9	CAT95 parameter distribution	90
6.10	nvt parameter distribution	90
7.1	SET3 event-rate	95
7.2	Annual modulation model background activities	97
7.3	Background energy spectrum	98
7.4	SET3 annual modulation fit	101
7.5	Over-estimated ^{210}Pb activity in SET3 data	103
7.6	SET3 background activities vs input with split ^{210}Pb	106
7.7	Split- ^{210}Pb activities comparison with SET3 data	107
8.1	Single pseudo-experiment event-rate	113
8.2	Pseudo-experiment fitted	115
8.3	Pseudo ensemble modulation amplitude distribution	116
8.4	Pseudo ensembles modulation amplitude biases	116
8.5	Pseudo ensembles pull RMS	118
8.6	Single pseudo ensemble χ^2 distribution split- ^{210}Pb	119
8.7	Single pseudo-ensemble χ^2 distribution single- ^{210}Pb	120
8.8	Modulation amplitude bias with single and split ^{210}Pb	121
8.9	single/split- ^{210}Pb modulation amplitude difference distribution	122
8.10	High χ^2 fit pseudo-experiment	123
8.11	High χ^2 fit pseudo-experiment background activities	124
8.12	Floated-phase low modulation amplitude bias	126
8.13	Floated phase best-fit phases pseudo distributions	127
8.14	Floated-phase amplitude-phase 2D pseudo histogram	128
9.1	NaI energy spectrum with ^{40}K event selection	131
9.2	^{40}K peak position over time	132
9.3	X1/X2 parameter distribution	135

9.4	2 keV BDT signal efficiency	135
9.5	PMT noise templates	136
9.6	Mean-time parameter distribution	137
9.7	Pulse-shape/mean-time parameter distribution	138
9.8	Pulse-shape/mean-time parameter cuts	139
9.9	Likelihood parameter distribution	140
9.10	1 keV BDT selection efficiency	141
9.11	1 keV BDT variable distribution	141
9.12	Counts per subrun distributions	143
9.13	SET4 event rate and modulation fit	145
9.14	SET4 1-6 keV fixed-phase modulation amplitude posterior distributions . .	146
9.15	Fixed-phase modulation amplitude comparison	147
9.16	1 keV region modulation amplitudes	148
9.17	SET4 1-6 keV floated-phase posterior distribution	150
9.18	SET4 1-6 keV floated-phase posterior distribution with pseudo-data com- parison	152
A.1	Surface ^{210}Pb best-fit half-life pseudo distribution	168
A.2	Best-fit surface ^{210}Pb half life and χ^2 relationship	169
A.3	Best-fit surface ^{210}Pb half life and activity relationship	170

List of Tables

4.1	Properties of the COSINE-100 crystals	44
4.2	Radiogenic backgrounds at COSINE-100	45
5.1	LS energy calibration peak locations	78
7.1	Annual modulation fit model background components	96
7.2	The best-fit values for the effective half-life of surface ^{210}Pb when the updated annual modulation model for split bulk and surface ^{210}Pb is fit to SET3 data.	108
9.1	Fixed-phase best-fit modulation amplitudes	147
9.2	Floated-phase best fit modulation amplitudes and phases	149

Chapter 1

Introduction

The unknown nature of dark matter is one of the biggest unanswered questions in modern particle physics. A brief historical overview of the understanding of dark matter is presented in Chapter 2, through descriptions of some of the major astronomical and cosmological observations now attributed to dark matter. Also described in Chapter 2 are some of the main theoretical candidates postulated as explanations for the observations attributed to dark matter. Here there is a focus on particle candidates, particularly weakly interacting massive particles (WIMPs), which have been the focus of the majority of dark matter detection efforts at the time of writing.

Many of these experimental searches and their results are then described in Chapter 3, initially highlighting results from liquid Xenon (LXe) time projection chambers (TPCs) which currently set the most stringent limits on WIMP dark matter for the most-studied WIMP masses. Experiments setting the most stringent limits on lower mass WIMPs are then described before the DAMA experiments are introduced. The DAMA experiments are unique in that they are the only direct dark matter detection effort reporting a signal attributed to dark matter. Specifically DAMA report an annual modulation in signal events in NaI crystals over many years with an extremely high confidence level. Despite scepticism around the DAMA signal in light of the lack of a corroborating signals from other, theoretically more sensitive experiments, tension remains due to the lack of a model-dependent test of the DAMA result, searching for the same signal in the same target material.

The COSINE-100 experiment is also an NaI-based dark matter detection experiment searching for an annual modulation signal. Therefore COSINE-100 aims to finally end the long-standing tension brought about by the DAMA signal, either by replicating the DAMA result or by excluding it by conclusively measuring no modulation. The design and detector operation of COSINE-100 is presented in Chapter 4.

A major difference between the COSINE-100 and DAMA experiments is the presence of active shielding at COSINE-100, designed to measure and hence veto certain background events, whereas only passive shielding is used at the DAMA experiments. A key element of the COSINE-100 active shielding is the liquid scintillator (LS) veto that the crystal encapsulation is immersed in. This LS veto is designed to record events from radioactive decays from within the detector and the surrounding environment, and then veto any crystal events that occur in coincidence with the LS events. The performance of the LS over the entire COSINE-100 data-taking period is described in Chapter 5, focusing specifically on the process of gain correction, then the identification of an issue in one of the LS PMTs and the resulting changes to the LS operation that resulted.

The largest source of background events in the energy region used in dark matter

searches at COSINE-100 are from PMT noise. Most of these noise events can be rejected using multi-variate analysis parameters designed to discriminate between signal and PMT noise. However, there is a class of intermittent PMT noise at COSINE-100 with different characteristics to other types of noise measured in the detector, and during periods where this noise is present, the data taken cannot be used in dark matter searches. A study into this intermittent noise is described in Chapter 6 where the development of a new method of identifying the noise events is described.

The primary method of obtaining the result used to test the DAMA signal at COSINE-100 is to fit the event-rate over time of signal-like data to a model designed to describe the event-rate. This model is therefore designed to describe two features of the measured data-set, the background event-rate, and any annual modulation present. The model is updated and improved for every successive annual modulation search at COSINE-100 and in Chapter 7, the model used to analyse the previously untested first 4.7 years of COSINE-100 data is described. Specifically, the new additions to the model are an improved method of dealing with detector dead-time in the model, and a revised treatment of the ^{210}Pb component. This involves splitting the component into two separate bulk and surface components, in line with observations of the surface ^{210}Pb being replenished over time and hence having an increased effective half-life.

Simulated (“pseudo”)-data is used at COSINE-100 to test and validate the fitting procedure used in annual modulation searches. In Chapter 8, the process generating this pseudo-data and the subsequent of testing and validating the new fitting procedure is described. The fit model’s ability to correctly identify the modulation and background components used to generate the pseudo-data is presented for fixed-phase searches. Then follows a description of a known issue regarding correctly reporting annual modulation amplitudes in the case where no true modulation is present, and a novel use of pseudo-data to account for this issue and correctly compare results of physics data to the no-modulation case.

Finally in Chapter 9, the results of applying this new fit model to the first 4.7 years of COSINE-100 data are presented, first for the fixed, then floated-phase case. This is the first time the 4.7 years data-set has been analysed. These results are compared to the DAMA and no modulation hypothesis to determine the level of agreement with both. A small annual modulation, compatible with both hypotheses was found in all cases. This necessitates further collection and analysis of data at COSINE-100, these future efforts are briefly described in Chapter 10.

Chapter 2

Dark matter: Evidence and Candidates

The advancements in astronomical and cosmological observing techniques in conjunction with ever-improving theoretical models have allowed for an increasingly detailed and accurate understanding of the nature of the universe to be built over the last century. However, as our view of the universe becomes clearer, so too does it become clear that the majority of the mass content of the universe is unknown. This mysterious, invisible material is “dark matter”, and its nature is one of the biggest unanswered questions in modern physics.

Measurements of the anisotropies of the cosmic microwave background (CMB) radiation from the Planck Satellite [6] suggest, using the Λ CDM model, that this dark matter makes up about 26% of the mass-energy density of the universe, more than five times as much as baryonic matter [7]. The remainder comes from the equally mysterious “dark energy”, the name for the force driving the ever-accelerating expansion of the universe. Despite this apparent abundance of dark matter and a multitude of experimental efforts dedicated to detecting it, there have been no (widely accepted) successful attempts to detect dark matter, only observations of its astronomical affects. As such, the nature of dark matter is unknown. This lack of knowledge about something believed to make up such a large part of the universe is one of the largest unanswered questions in physics.

Since the information on the nature of dark matter itself is limited, potentially viable dark matter candidates that may explain the available astronomical observations exist across an extremely large number of orders of magnitude in mass and cross-section of interaction, making search efforts challenging and leading to a vast array of different experimental techniques. At the time of writing, the lower limit of dark matter mass is reported at around 10^{-22} eV/ c^2 [7], whereas well motivated particle dark matter candidates have been proposed with masses up to the Planck scale at around 10^{19} GeV/ c^2 [8], and beyond the particle dark matter paradigm, macroscopic dark matter candidates are currently reported to have an upper mass limit of around $10^5 M_{\odot}$ [7]. The most likely dark matter candidate seems to be some new particle, beyond the Standard Model, but many other candidates also exist.

Full reviews of the evidence for dark matter and potential dark matter candidates from theory matter would all individually dwarf the scope of this thesis chapter, such is the volume of each. As such, only a brief overview of these topics are given below, starting with the evidence for dark matter, both astronomical in Sections. 2.1.1, 2.1.2, and 2.1.3, and cosmological in Section 2.1.4 followed by a description of some dark matter candidates, both macroscopic in Section 2.2.1, and particle in 2.2.2.

2.1 Astronomical observations of the effects of dark matter

Observations of phenomena now attributed to dark matter long precede the development of the now widely accepted concept of dark matter, and for many decades unexplained gravitational effects were reported by astronomers without any significant speculation that these effects might point to a fundamental gap in our understanding of the nature of the universe [9]. It was not until the 1970s when observations of galactic rotation curves [10] indicated a large amount of non-luminous matter, beyond that which was already understood, in nearby galaxies that the concept of “dark matter” as it is now known gained widespread recognition in the physics community. After this discovery, an ever increasing number of observations pointing to a large amount of unseen, unexplained matter in astronomical objects were reported, building the overall picture of dark matter that exists today.

In this section some of the early observations of unseen matter are discussed, before a description of some of the major pieces of evidence that motivate the existence of dark matter in the universe.

2.1.1 Early observations

The concept that more material exists in the universe than can immediately be observed has existed in science and philosophy since antiquity [9], albeit without necessarily containing the notion that this unseen material might fundamentally differ from that which is more readily observed. Indeed, as the field of astronomy progressed, observations that required the existence of unseen bodies became more common, the discovery of Neptune in 1846 through its effects on the orbit of Uranus [11, 12] being a famous example. This however is far from the only example, with perturbing effects of unseen stars being recorded in the 19th century, and speculation about the nature of “dark nebulae” in the Milky Way being a topic of astronomical interest in the late 19th and early 20th centuries [9].

A well-known example of early observations of unseen matter we now attribute to dark matter is the observation by Fritz Zwicky in 1933 that the large differences in velocities relative to one another (corresponding to twice the peculiar velocity of the cluster) of galaxies in the Coma Cluster as measured by their redshifts. This indicated a large amount of unseen material in the cluster [13]. Specifically, Zwicky estimated the total mass of the cluster by counting the number of visible galaxies and estimating their average mass, as well as estimating the radius of the cluster. Assuming the cluster is in stable equilibrium, the virial theorem (Eq. 2.1) can be applied to the cluster and assuming an average density across the cluster, the gravitational potential energy of the cluster can be estimated (Eq. 2.2)

$$2\langle T \rangle = -\langle V_{GPE} \rangle, \quad (2.1)$$

$$V_{GPE} = -\frac{5}{3} \frac{GM^2}{R}. \quad (2.2)$$

Here $\langle T \rangle$ is the average kinetic energy (KE) of the cluster, V_{GPE} is the total gravitational potential energy (GPE) of the cluster, $G \approx 6.67 \times 10^{-11} \text{m}^3 \text{kg}^{-1} \text{s}^{-2}$ is the gravitational constant, and R is the radius of the cluster. Then using the resulting expected kinetic energy of the cluster, the expected peculiar velocity can be estimated. This estimated

value was around 80 times smaller than the measured value, indicating a large amount of unseen material in the cluster. Zwicky continued to study galaxy clusters and continued to observe large amounts of “dark matter”, although he attributed this dark matter to “cool and cold stars, macroscopic and microscopic solid bodies, and gases”, rather than some new material, or what would now be considered to be dark matter. [9].

Zwicky’s observation of unseen material in galaxy clusters was joined by others through the early and mid 20thth century [9], though crucially both these and Zwicky’s methods used various as yet unconfirmed assumptions about the nature of the clusters, as well as having a strong dependence on the assumed value of Hubble’s constant to calculate the cluster’s radius. A value that was neither accurately nor precisely known at the time. As such, the conclusion of large amounts of unseen dark matter was often rejected on these grounds with other explanations put forward.

2.1.2 Galaxy rotation curves

This position of uncertainty about the cause of the disparity between the estimates of masses of galaxy clusters as measured by the visible material versus the dynamics of the clusters remained until in the 1970s a second compelling piece of evidence for dark matter came in the form of galaxy rotation curves.

Advancements in radio astronomy in the late 1960s allowed for Vera Rubin and Kent Ford Jr to measure rotation curves of the M31 (Andromeda) galaxy more accurately than ever before [10]. They consistently found that the rotation velocity of the material in the galaxy as a function of distance to the galactic centre (R) reached a maximum value then decreased gradually as R increased (as shown in Fig. 2.1). This is in stark contrast to the rotation curve expected if the mass distribution of the galaxy followed that expected from the location of the luminous matter, i.e. a flat disc around a central bulge where a much larger decrease in rotational velocity as a function of R would be expected. The difference in the expected and observed rotation curves is then discussed by Ken Freeman in a 1970 paper [14], published shortly after Rubin and Ford’s results where the crucial inference that “*there must be in these galaxies additional matter which is undetected*” is made.

It should be mentioned that acceptance of a dark matter-based explanation to the newly observed flat galactic rotation curves was by no means uniformly accepted. A well-known alternative explanation to the observation of flat galactic rotation curves arose from the idea that gravity may behave differently at very large scales than it does at more familiar scales. The most well-known of these is the Modified Newtonian Dynamics (MOND) theory developed by Mordehai Milgrom in the early 1980s [15] which in its simplest form suggests that Newton’s second law of motion should in fact include an additional term which in the limit of high acceleration (for example of the surface of the Earth) reduces to one, but in the limit of low acceleration (for example the very small accelerations of stars orbiting a galactic centre) this term is present as a constant and acts to slightly decrease the force compared to what would be expected from the standard formulation of the second law. Using the MOND formulation, predictions of flat galactic rotation curves can easily be recovered without the need for additional dark matter.

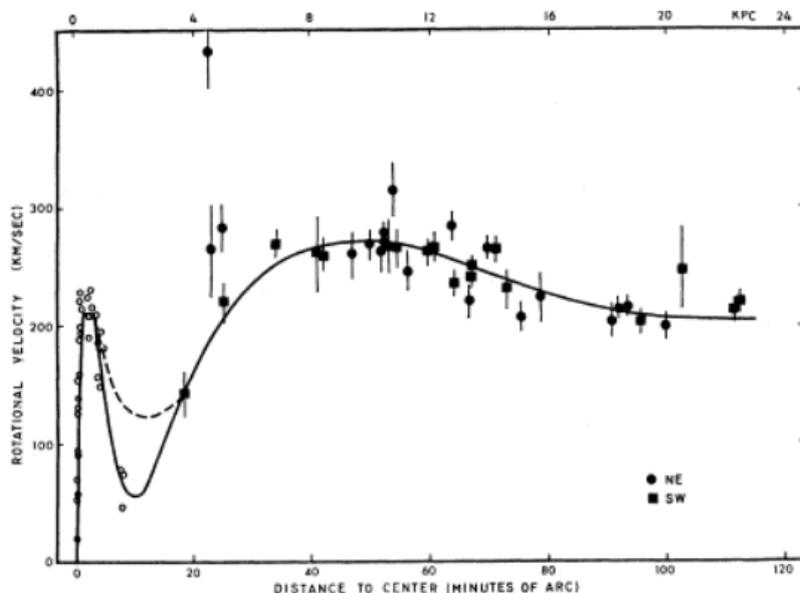


Figure 2.1: The rotational velocity of gas (data-points) as a function of radius in the M31 galaxy fit to a fifth order polynomial (solid line) as reported in the 1970 Rubin paper on galaxy rotation curves [10].

2.1.3 Gravitational lensing

After the discovery of flat galactic rotation curves in the 1970s, effort was put into studies attempting to further understand the nature of dark matter. The lack of direct observations of dark matter up to this point leads to the conclusion that dark matter must not interact through the electromagnetic force in a way that is observable on Earth, hence the lack of observations. Given this inherently invisible nature of dark matter to standard astronomical search techniques, attempting to further study its distribution and behaviour was a non-trivial task since rather than observing the dark matter itself, astronomers are only able to observe dark matter through its gravitational effects on luminous matter. One such method of observing dark matter is to utilise the effects of gravitational lensing. Dark matter, like all massive objects, bends (lenses) light to some degree. Therefore observing a large amount of gravitational lensing of background sources in a region of sky with an insufficient number of luminous sources to explain the observed lensing implies the existence of dark matter in the ‘empty’ region.

Perhaps the most famous and influential example of gravitational lensing due to dark matter comes from observations of the Bullet Cluster [16]. The Bullet Cluster consists of two initially separate clusters of galaxies which have collided and passed through one another. The majority of the baryonic mass in galaxy clusters comes not from the stars that make up the part of the cluster that is visible in the visible spectrum, but instead from the intergalactic plasma which is visible in the X-ray region of the electromagnetic spectrum [16]. During a collision and merger of two galaxy clusters, as in the Bullet Cluster, it is observed that this plasma (which is relatively evenly distributed throughout the clusters) behaves like a fluid, colliding and interacting with the plasma in the other cluster, ultimately slowing it down. The galaxies however are sufficiently far apart to behave as collisionless particles during the merger and consequently carry on unperturbed.

The distributions of the galaxies and plasma in the Bullet Cluster are illustrated in Fig. 2.2a and Fig. 2.2b from [16] respectively which show the same region of space in

different frequencies. The bright regions of the left plot are the galaxies of the cluster, and the shaded regions of the right plot represent the distribution of the plasma as measured by X-ray emissions. The prominent bow shock of the plasma on the right hand side of the cluster - further indication the plasma has interacted and slowed - is also clearly visible and shows the direction of travel of the original cluster (left to right as shown). The green contours on both plots represent the areas of greatest gravitational lensing of sources background to the cluster. Crucially in the context of dark matter, this lensing is centred not on the plasma, the dominant baryonic mass component of the cluster, but on the galaxies, thus indicating that the majority of the mass of the two original clusters has passed through one another in a collisionless manner. This combination of behaviours (collisionless and sufficiently massive to be the dominant contributor to the mass of the cluster) is not consistent with any known baryonic matter or astronomical structure and therefore strongly hints at the existence of dark matter within these contours.

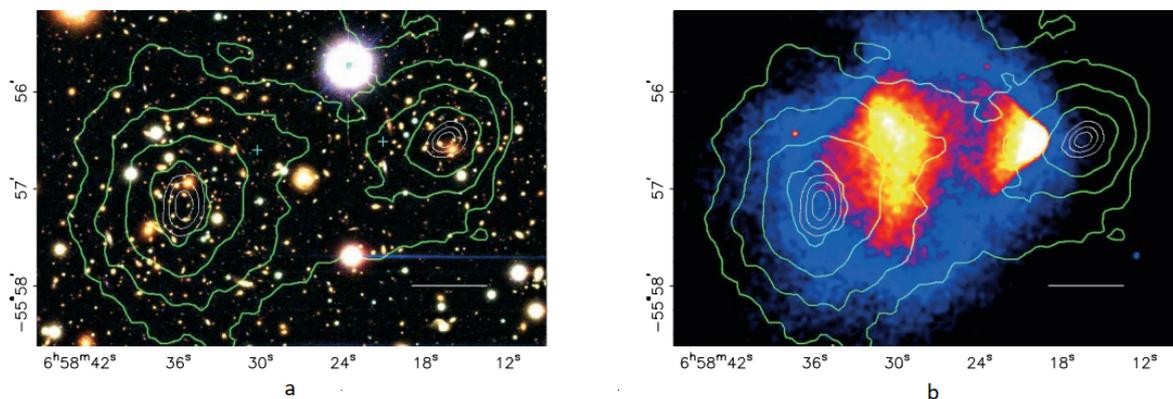


Figure 2.2: The Bullet Cluster viewed at two different wavelength ranges, visible on the left plot showing the location of the constituent galaxies of the cluster, and X-ray on the right showing the location of the plasma in the cluster, the dominant component of the baryonic mass in the cluster. The green contours show the areas of greatest gravitational lensing and hence the areas with the most mass. These do not follow the distribution of the plasma, indicating a large amount of unseen material. The scale is given by the 200 kpc-long white bar in both figures. From [16].

Unlike the flat galactic rotation curves described in Section 2.1.2, this disconnect between the location of the distribution of matter inferred from gravitational lensing and the location of the majority of the known baryonic mass of the cluster is much harder to reconcile with modified theories of gravity than with unseen material. As such, the Bullet Cluster result and others like it are often cited as sufficient evidence to rule out these theories as viable explanations to the observed phenomena. Nevertheless, modifications to these theories such as MOND, were developed, particularly in the context of non-spherically symmetric systems such as the Bullet Cluster, for example as in [17]. The Bullet Cluster is by no means the only galaxy cluster that exhibits the behaviour described above, though it is the most famous example. A review of such examples can be found in [18].

2.1.4 Dark matter in cosmology

Through a steadily increasing volume of astronomical evidence for the existence of dark matter over the last half-century, beginning from the discovery of flat galactic rotation curves, the existence dark matter is now a widely accepted assumption, particularly in cosmology. Here, the importance of dark matter in some cosmological models is described.

Studies of cosmic microwave background anisotropies

One of the clearest examples of the existence of dark matter's importance in modern cosmology is the success of the Λ CDM model (referring to dark energy, Λ , and cold dark matter, CDM). One of the main applications of the Λ CDM model is to provide a theoretical framework from which the total energy densities of dark matter, baryonic matter, and dark energy in the universe and in turn the ratios of the amounts of these values can be determined from observations. The current best estimates of these parameters are based on studies of the anisotropies of CMB radiation measured by the Planck satellite [6] and are $\Omega_c = 0.2589 \pm 0.0057$ and $\Omega_b = 0.0486 \pm 0.0010$ for the fractional dark matter energy density and baryonic matter energy density of the universe respectively. The remainder of the energy density of the universe is provided by dark energy. These values from Planck are in good agreement with values from earlier studies, for example by the Wilkinson Microwave Anisotropy Probe (WMAP) [19].

The estimates of the dark and baryonic matter energy density come from measurements of the very small ($\mathcal{O}(10^{-5})$) temperature anisotropies in the CMB radiation. The CMB reveals the state of the early universe, around 380,000 years after the big bang and at the beginning of the recombination epoch; the point at which the universe was sufficiently cool for free electrons to combine with protons to form hydrogen atoms. At this point the universe became transparent to black body radiation produced earlier in the universe's history as there were fewer free electrons this radiation could scatter with. This radiation has since been red-shifted by the expansion of the universe until it is at the microwave energies observed today.

Anisotropies in the temperature of the CMB measured today originate from regions of slightly differing density of the universe at the time of the recombination epoch. These differences in density are initially caused by the massive baryonic and dark matter clumping together due to the effects of gravity. This increase in density acted to heat the baryonic matter but not the dark matter in the dense regions, resulting in outward pressure on the baryonic matter but not the dark matter. This in turn caused oscillations in the density of the early universe where the size of the oscillations were related to the strength of the attractive forces from gravity and therefore the amount of dark matter present. These oscillations were effectively recorded at the time of recombination as more dark matter-dense and consequently hotter regions cooled to a temperature where recombination was possible later than less dense regions and in so doing, became transparent to radiation at slightly higher red shifts. The measured anisotropies can be described by power spectra (for example in Fig. 2.3 from [6]) which are fit to a function with the dark matter, baryonic matter, and dark energy energy densities of the universe as floated parameters of interest. Given this fit describes the data very well, these parameters can be estimated with high precision.

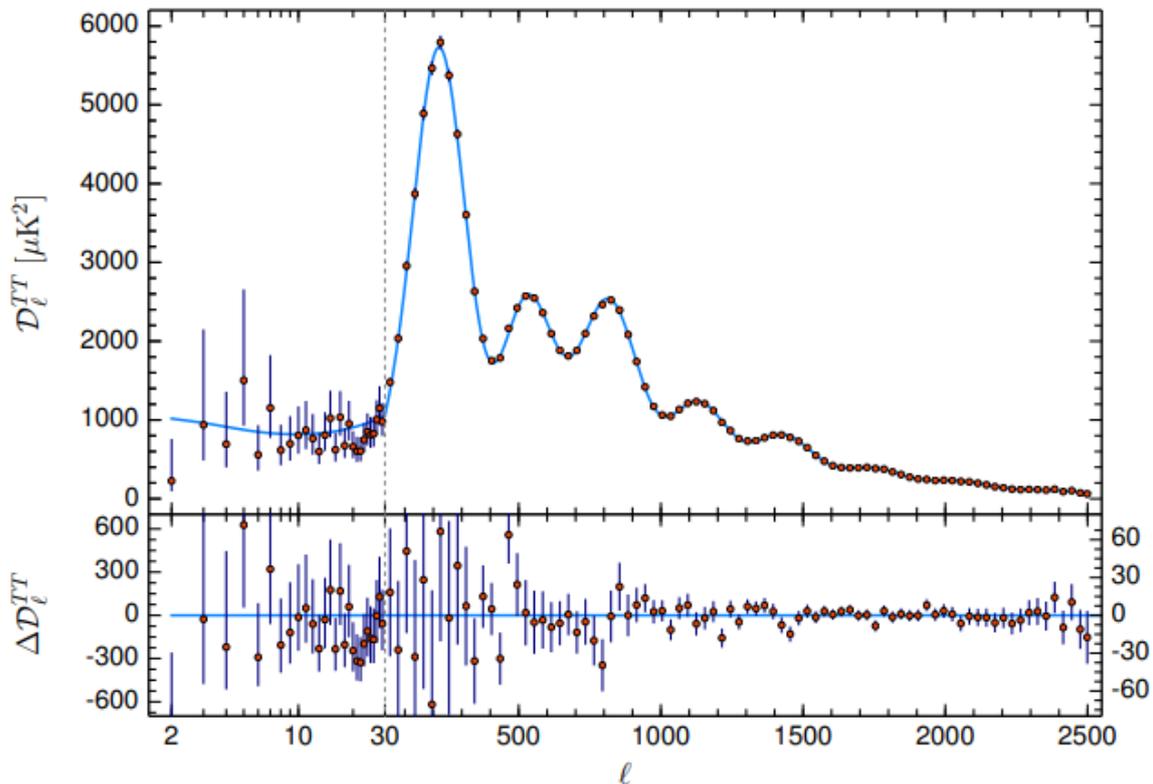


Figure 2.3: An example of a power spectrum of the CMB’s temperature (red data points) as measured by the Planck Satellite, and a fit to the data (blue line) containing parameters relating to the energy density of dark matter and baryonic matter in the universe. Shown below are the residuals between the fit and data. At the point where ℓ (multipole) = 30 (black dashed line), the horizontal axis changes from logarithmic to linear and the vertical axis also changes scale accordingly. From [6].

Big Bang nucleosynthesis

A second area of study in cosmology where the abundance of dark matter is a key parameter in successful models is the theory of Big Bang nucleosynthesis (BBN). BBN describes the production of atomic nuclei with masses above that of ^1H , in the very early universe, before the formation of stars. Specifically the formation of nuclides up to ^7Li in relatively very small amounts, but mainly ^4He with lesser amounts of ^2H (deuterium), and ^3He . This process can only occur when the universe was still hot and dense enough for fusion to occur, however, at the very high temperatures in the early universe, it was too hot for any nuclei formed to remain bound. Therefore any primordial nuclei observed today must have been formed between these two epochs of the early universe, approximately 10 seconds to 20 minutes after the birth of the universe [20].

During this period, the rate of reactions between lighter nuclei and subsequent production of heavier nuclei was dependent on the baryon-to-photon ratio (a value which remains constant during the period). Therefore measuring the relative abundance of these isotopes today will give an indication of the baryon-to-photon ratio at the time, since the majority of these nuclei in the universe today were formed during BBN. Specifically, the higher the baryon-to-photon ratio, the less deuterium is expected in the modern universe, since deuterium is produced during “incomplete” reaction chains that ultimately produce ^4He ,

leaving left over deuterium. This measurement of the relative abundance of deuterium and therefore an estimate of the baryon-to-photon ratio is related to studies of dark matter by the relationship between the amount of baryonic and dark matter. The energy density of the early universe can be estimated, and were that energy density in early universe solely provided by baryonic matter, the baryon-to-photon ratio would be significantly too high to explain the measured amounts of deuterium [7]. Thus the relatively high measurements of the present-day abundance of deuterium (summarised in Table 24.1 of [7]) imply the existence of dark matter in the early universe.

2.2 Dark matter candidates

For the remainder of this work after this chapter, the paradigm operated in is that dark matter is likely to be some previously undiscovered non-Standard Model particle. Indeed, this is the paradigm within which modern mainstream particle physics operate in. However, that is not to say that alternative explanations of the nature of dark matter do not exist, nor that there need only be one source of all the missing matter observed with multiple dark matter candidates existing in parallel. In this section, these non-particle explanations are described along with descriptions of Standard Model particles posited as dark matter, as well as particle dark matter candidates beyond the Standard Model. The potential issues of these models are also discussed. An exhaustive description just of all proposed weakly interacting massive particle (WIMP) dark matter candidates, let alone dark matter candidates of all types would far exceed the length and scope of this section, such is the volume of candidates. As such, only a small number of more mainstream candidates are discussed.

2.2.1 Macroscopic candidates

One alternative to the particle dark matter paradigm is that the discrepancy between the amounts of matter observed in astronomical systems and the amount of mass required to explain the system's dynamics is due to the presence of unseen/underestimated macroscopic astronomical objects which make up the observed hidden mass component. These objects could potentially have individual masses up to the order of Solar masses. As a result of this, potentially viable dark matter candidates have been suggested across a mass range spanning approximately 100 orders of magnitude. The spread of dark matter candidates at different mass ranges is shown in Fig. 2.4 from [21].

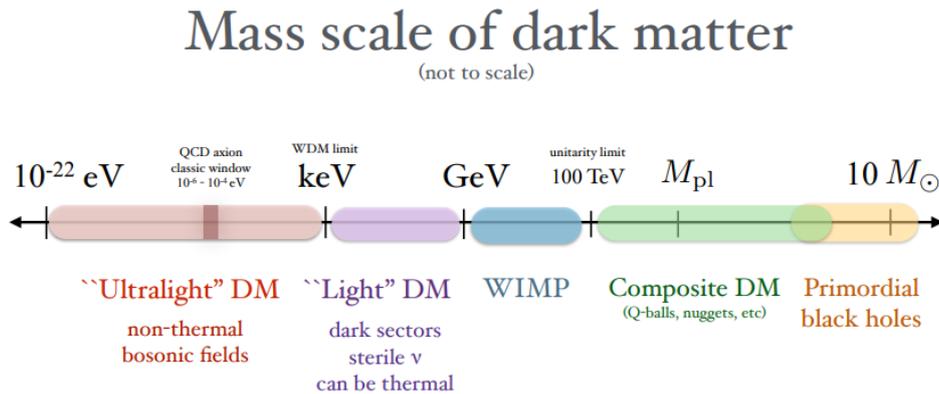


Figure 2.4: Visual representation of the approximate mass scales of various dark matter candidates. From [21].

Black holes, being non-luminous, massive, and compact satisfy the basic criteria for dark matter candidates. However, in order to explain the results of cosmological observations such as the present abundance of deuterium described in the previous section, as well as the formation of structure in the early universe, black holes that solve the dark matter question must have been formed in the very early universe, before primordial nucleosynthesis, where there is no known mechanism for the formation of black holes [22]. The black holes posited as the primary dark matter candidate are therefore of a different origin to the already observed black holes and are named “primordial black holes” (PBHs), due to the necessity of them being formed in the very early universe. This constraint is also necessary since this is the period where the universe might have been sufficiently dense for black holes to form outside of a collapsing star [22].

It is immediately apparent that PBHs as primary dark matter candidates (making up the majority of the dark matter in the universe) will have a lower limit on their initial mass as a result of the requirement that they have not lost enough of their mass as a result of Hawking radiation [23] over their lifetime to still generate the gravitational effects attributed to dark matter observed today. Additionally, a PBH finally evaporating due to Hawking radiation would produce a characteristic, measurable γ -ray signal that has not been observed [24]. PBHs with initial masses of around 10^5 kg would now be at an age where they are starting to reach the ends of their lives and would therefore be producing this characteristic γ -ray signal. Since the location of dark matter in the local universe is known, for example in the Milky Way’s galactic halo, it should be relatively easy to locate such a signal from the evaporating PBHs, however, no such signal has been observed [24], thus ruling out PBHs at and below this mass.

An upper limit on the mass of PBHs and primary dark matter candidates is also present due to the observed number of neutron stars in dark matter-dense regions. This is because were PBHs above a certain mass present in these regions, they would act to destroy neutron stars in the region through accreting their material into the PBH [25]. The upper mass limit imposed by this constraint is around 10^{15} kg [25]. A more stringent high mass limit is calculated from the lack of observations of weak gravitational lensing that could be attributed to PBHs. As described in Section 2.1.3, gravitational lensing should provide an excellent tool to observe PBHs in dark matter-dense regions when a PBH passes in front of a background source. This is because the changing brightness of a background source due to lensing would be characteristically symmetrical, of equal magnitude across all wavelengths, and only happen once per background source. An example of a light curve

with these characteristics is shown in Fig. 2.5. However, in a 2018 search observing the Andromeda galaxy with the Subaru Hyper Suprime-Cam (HSC) telescope for seven hours such events found only a single viable candidate [26] (shown in Fig. 2.5) and subsequently set limits on PBH mass of around 10^{19} - 10^{24} kg if PBH are to be the primary dark matter candidate.

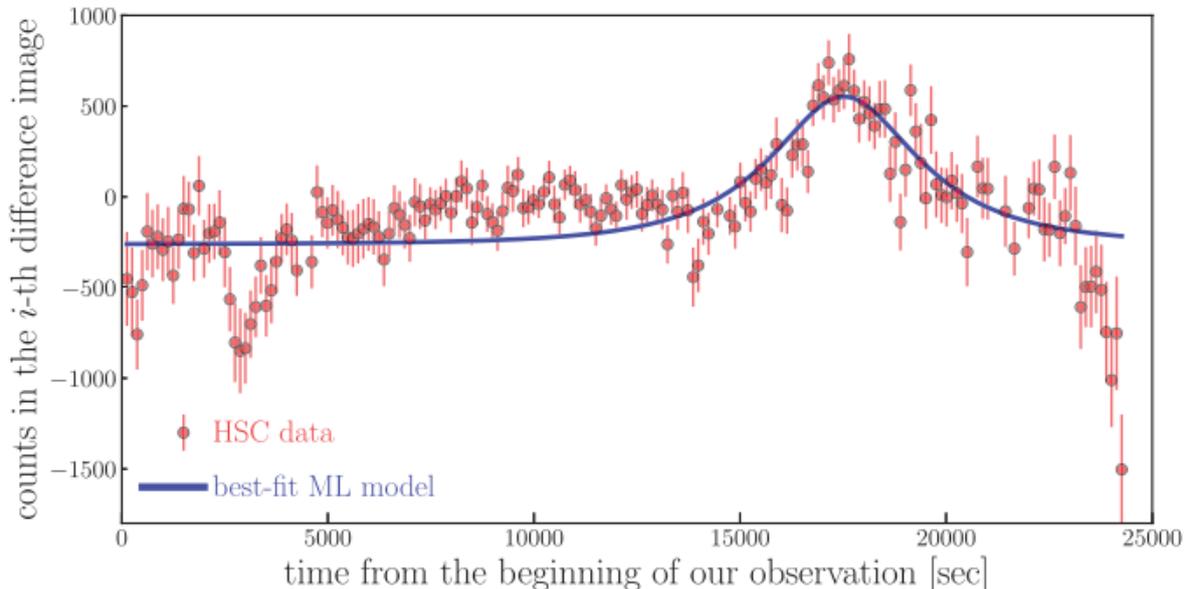


Figure 2.5: The single viable potential black hole candidate observed by the Subaru Hyper Suprime-Cam (HSC) telescope during a seven hour observation of the Andromeda galaxy. The red data points are the amount of light observed in a given time bin and the blue curve is a fit to the data using the best-fit microlensing model. Adapted from [26].

Another class of macroscopic dark matter candidates are massive compact halo objects (MACHOs), which could be any non-luminous astronomical object of sufficient mass, for example large planets or brown dwarf stars. These objects must be below a mass of around $0.08 M_{\odot}$ in the case where hydrogen is the primary component of the object, as objects above this mass would be sufficiently dense in their core to start fusing hydrogen to helium and thus no longer be dark [27]. MACHOs have been searched for utilising the same gravitational lensing techniques described for PBHs, and as with PBHs, only a very small sample of potential candidates have been identified, placing stringent limits on the mass of MACHOs as primary dark matter candidates [28].

2.2.2 Particle candidates

The tight experimental constraints on macroscopic dark matter candidates, the constraints on the abundance of baryonic matter from BBN and studies of the CMB, as well as the difficulties in reconciling adapted models of gravity to astronomical observations of dark matter have meant particle dark matter candidates have become the dominant dark matter candidates searched for. Particle dark matter candidates are an attractive proposition for theorists since the lack of knowledge about the nature of dark matter leaves open many possible guises a dark matter particle might take, and therefore a very large number of particle dark matter candidates from theory exist.

At the very simplest level, a particle dark matter candidate must be both massive in order to explain the gravitational effects described in Section 2.1, while also not inter-

acting strongly or electromagnetically (i.e. having no electrical or colour charge), since dark matter interactions with either the strong or electromagnetic forces would produce measurable signals. Dark matter particles must also be stable at time scales greater than the current age of the universe since the effects of dark matter are observed from the very early universe (for example as described in Section 2.1.4) to the present day. Beyond this, particle dark matter candidates must also behave in a way that is consistent with the observed large scale structure (LSS) of the universe. It is possible to simply imagine a new particle that fits the above criteria by definition, however, such a theory would be “poorly motivated”, inasmuch as there is no reason to believe this hypothetical particle exists, other than the fact it happens to meet the criteria required for dark matter. Much more convincing dark matter candidates appear when they arise naturally from theory designed to explain some other problem present in physics and most popular dark matter candidates are conceived in this way (effectively by accident), rather than specifically to solve the dark matter issue.

Neutrinos as dark matter

Although thought to be massless for much of their known existence, modern studies of neutrino flavour oscillations have shown that neutrinos do have some (very small) mass. Indeed, this discovery was awarded the 2015 Nobel Prize in Physics, awarded to Takaaki Kajita and Arthur B. McDonald [29]. This means that neutrinos fulfil all three of the massive, long-lived (they are stable), and only weakly interacting criteria for particle dark matter candidates. Nevertheless, Standard Model neutrinos are not a viable dark matter candidate. In order to explain why this is the case, it is useful to discuss the “hot” and “cold” classifications of dark matter.

These classifications refer to how relativistic dark matter particles were at the point at which they fell out of thermal equilibrium in the early universe, known as “freeze-out”. Cold dark matter is characterised by being non-relativistic at the time of freeze-out and could therefore clump together in the very early universe, forming gravitational potential wells which catalysed the formation of density perturbations in baryonic matter which ultimately lead to galaxies and the LSS observed today [30]. Hot dark matter is characterised by being relativistic at time of freeze-out and remaining so into the matter-dominated epoch, setting an upper mass limit on hot dark matter at around a few tens of eV [30]. Hot dark matter would have a very large mean free path due to its relativistic velocities at the time of freeze-out. This would prevent it from clustering, and therefore hot dark matter is incompatible with LSS and can be ruled out as the primary form of dark matter in the universe. Standard Model neutrinos, being extremely light, were highly relativistic with velocities very close to c at the time of freeze-out. This fact alone precludes Standard Model neutrinos as primary dark matter candidates.

Standard Model neutrinos are however not the only dark matter candidate related to the neutrino. The neutrino differs from other observed particles with non-zero spin in the Standard Model in that neutrinos have only ever been observed with left-handed chirality (and all anti-neutrinos with right-handed chirality) whereas all other spin-carrying particles are observed to exist with both left and right-handed chirality. The lack of a right-handed neutrino may be reconciled within the Standard Model if neutrinos have no mass [31], the observation of neutrino mass therefore motivates the existence of a right-handed “sterile” neutrino.

A right-handed neutrino, like all right-handed fermions, would have a weak isospin quantum number equal to zero (left-handed fermions with a weak isospin of $1/2$) and

therefore would not undergo charged weak interactions, right-handed neutrinos would also have zero weak hypercharge and as such would not feel the weak force at all [32]. Like the left-handed neutrino, the sterile neutrino also would not interact with the strong or electromagnetic forces, making the sterile neutrino a potential dark matter candidate, should it exist in a requisite mass range to explain the properties of dark matter.

In practice, there are few theoretical constraints on the possible mass range of sterile neutrinos and well-motivated neutrino mass ranges span parameter space from over 10^9 GeV in models pertaining to grand unified theory (GUT), down to eV scales in theoretical solutions to anomalies observed in data pertaining to flavour violation in neutrino experiments, and/or the cosmological origin of the baryonic matter in the universe in cosmological searches [32]. Massless sterile neutrinos are also theoretically possible, inasmuch as they are not forbidden by theory, although they are somewhat unnatural and are not usually considered [32].

Of most interest as a dark matter candidate are sterile neutrinos with a mass at the keV scale. The properties of such keV-scale sterile neutrinos are constrained by the requirement that they produce both the correct abundance of dark and baryonic matter in their interactions in the early universe (described in more detail in Section 2.1.4) as well as producing the correct LSS when tested with simulations of structure formation [32]. These constraints place viable sterile neutrino dark matter candidates in a categorisation in-between the hot and cold dark matter classes described above. This third class of dark matter, “warm dark matter”, was relativistic at the time of freeze-out, but subsequently slowed to non-relativistic speeds approximately one year after the formation of the universe, thus allowing for similar viable models of structure formation as in cold dark matter [32].

While models of sterile neutrinos can be produced with the requisite properties to obtain the observed dark and baryonic matter abundances, and the correct LSS in simulations, the many constraints now placed on sterile neutrinos as a primary dark matter candidate somewhat disfavour them as a candidate. In Fig.2.6 from [31] these constraints are shown for sterile neutrinos at different masses (M) and mixing angles (θ) with Standard Model neutrinos.

The phase space constraint on M (pink lines) arises from the fact that if the sterile neutrino mass is too low then the high number density required to have enough mass of dark matter in the early universe would be high enough to be in conflict with the Pauli exclusion principle which forbids fermions with the same quantum numbers occupying the same quantum state. Upper bounds on θ as a function of M shown by the solid blue line result from the lack of signals from different searches for X-ray emissions corresponding to decays of sterile neutrinos in dark matter-dense objects. The dashed blue line shows the sensitivity of planned future experiments. Additional upper bounds on θ as a function of M shown by the black solid line arise from constraints that not too much dark matter was produced thermally through the weak interaction in the early universe to exceed the observed dark matter density. Constraints from the amount of light nuclei produced during BBN inform the dashed green line. For brevity the other features of this plot are left undescribed since they exist in regions of parameter space covered by other constraints. For a full description of the features of this plot, see [31].

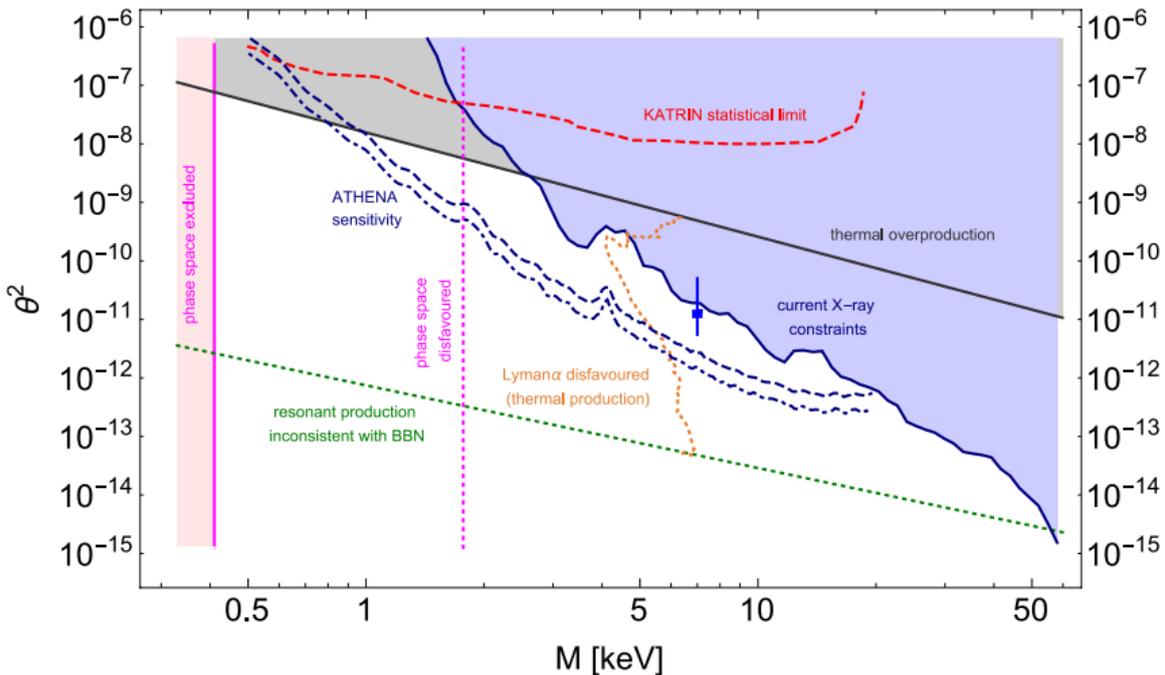


Figure 2.6: Various constraints on the mass (M) and mixing angle (θ) of sterile neutrinos as primary dark matter candidates. These constraints are described in more detail in the body of the text and in [31], the source of the plot.

WIMPs

By far the most popular dark matter candidate, at least in terms of experimental search efforts, is weakly interacting massive particles, “WIMPs”, typically with GeV scale masses. As discussed above, in addition to the massive, weakly interacting, and long-lived constraints on dark matter candidates, it seems unlikely from the lack of experimental evidence of their existence that macroscopic objects can make up the majority of dark matter. Additionally constraints from modelling the early universe suggest that dark matter should be cold, or at least not too warm. Many well motivated dark matter candidates fall under the WIMP umbrella and so WIMPs have become one of, if not the leading class of dark matter candidate.

One of the few things known about the properties of dark matter with relatively high precision (e.g. the results from studies of CMB anisotropies discussed in Section 2.1.4) is its relic density, a value determined by the mechanism of dark matter production in the early universe. The very early universe existed as an extremely hot, dense plasma and temperatures were sufficiently high that dark matter and Standard Model particles were in thermal equilibrium with one another. During this epoch, particle/antiparticle pairs could be produced then promptly annihilated back into the plasma. If WIMPs (denoted here by the χ subscript) make up all of the dark matter in the universe and the temperature of the universe is much higher than the mass of a WIMP ($T \gg m_\chi$) then WIMP pairs can be created and annihilated equally efficiently and thus WIMPs existed in large numbers and in thermal equilibrium. In this equilibrium, the number density of WIMPs (n_χ) is given by

$$n_\chi = g_\chi \left(\frac{m_\chi T}{2\pi} \right)^{3/2} \exp \frac{-m_\chi}{T}, \quad (2.3)$$

where g_χ is the number of degrees of freedom of the WIMP and m_χ is the WIMP mass [33]. This expression shows that the amount of WIMP dark matter in the universe will be suppressed as the universe expands and cools if the WIMPs stay in thermal equilibrium. Since this is evidently not the case, some mechanism to take the WIMPs out of equilibrium must exist.

Such a mechanism that gives the correct dark matter relic abundance arises from the dynamics of WIMP pair annihilation. The WIMP annihilation rate, Γ_{ann} , given by

$$\Gamma_{ann} = \langle \sigma_{ann} v \rangle n_\chi, \quad (2.4)$$

where σ_{ann} is the WIMP annihilation cross section and v is the relative velocity of the WIMPs [34]. When the expansion rate of the universe exceeds Γ_{ann} for a given WIMP any corresponding antiparticle is far enough away such that the WIMP cannot ‘find’ an antiparticle to annihilate with. The point at which this process occurs is called freeze-out and is dependant of the value of Γ_{ann} and therefore σ_{ann} . This relationship is shown in Fig. 2.7 for σ_{ann} values at the weak, electromagnetic, and strong scales.

It transpires that values of σ_{ann} of the order of magnitude of the weak force produce the observed dark matter relic abundance for GeV/TeV-scale WIMPs [30,34]. This coincidence was described as the “WIMP miracle” at the time of its discovery and placed WIMPs at the forefront of dark matter search efforts. However, further studies [35] have shown that the WIMP relic abundance may significantly differ from the value used in the above estimation, likewise, over time the range of viable WIMP masses from well-motivated theory has also increased [30]. Furthermore, in the many searches for WIMPs at this mass and cross-section, no evidence for the existence of such WIMPs has been found - the DAMA result notwithstanding - further suggesting the WIMP miracle may be a coincidence.

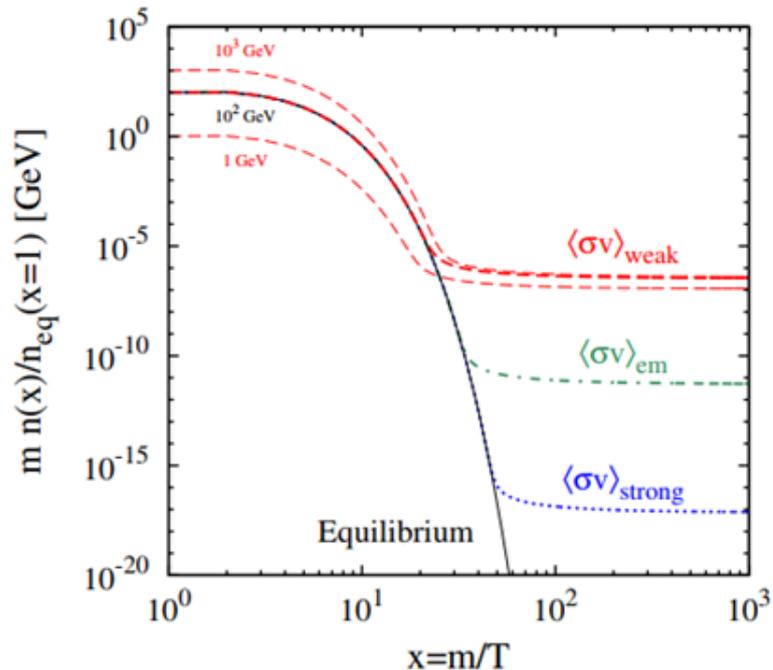


Figure 2.7: Evolution of WIMP abundance against $x = m_\chi/T$ (which is a proxy for time). The solid black curve shows the case where the WIMPs have a mass of 100 GeV and stay in thermal equilibrium and therefore become suppressed into insignificance. The thick dashed lines show the deviation from thermal equilibrium of the WIMPs if their cross section of self-annihilation is of the order of magnitude of the weak ($\langle\sigma_{ann}v\rangle = 2 \times 10^{-26} \text{cm}^3 \text{s}^{-1}$), electromagnetic ($\langle\sigma_{ann}v\rangle = 2 \times 10^{-21} \text{cm}^3 \text{s}^{-1}$), and strong ($\langle\sigma_{ann}v\rangle = 2 \times 10^{-15} \text{cm}^3 \text{s}^{-1}$) interactions for the red, blue, and green curves respectively. For the weak interaction strength case, results for WIMP masses of 1000 and 10 GeV are also shown by the upper and lower dashed red curves respectively. [33]

Many viable WIMP candidates exist from theory, where generally the most popular of these arise naturally out of versions of the supersymmetry (SUSY) theory, or more specifically, theories using the Minimal Supersymmetric Standard Model (MSSM). SUSY as a theory was not conceived to solve the dark matter problem, but rather to address other inconsistencies in the Standard Model. As such, any dark matter candidates that arise from the theory are well motivated.

At the most basic level, SUSY predicts a new supersymmetric bosonic partner for every fermion and a new fermionic partner for every boson¹. The theory is most commonly attributed to Julius Wess and Bruno Zumino in 1974 [36] although their theory built on work that already been carried out by others. For a more detailed history see [37], for a detailed primer on SUSY see [38].

SUSY in particular aims to solve is the hierarchy problem concerning the extremely large (24 orders of magnitude) discrepancy between the strengths of the weak force and the gravitational force. This issue manifests itself when considering the value of the Higgs square mass parameter. The exact meaning and definition of the Higgs square mass parameter is beyond the scope of this work, except to say that it is associated with the

¹In SUSY, particles are named by adding the prefix ‘s’ to the start of the SM particle in the case of SM fermions, e.g. quark goes to squark and by adding the suffix ‘ino’ in the case of SM boson, e.g. photon goes to photino

mass of the Higgs boson (although is not simply the rest mass squared), is affected in its value by quantum corrections result virtual corrections of every particle that couples with the Higgs field, is experimentally measurable from the value of the Higgs mass and has a value of $-(92.9 \text{ GeV})^2$ [38]. The corrections to the Higgs squared value have the potential to be extremely large, up to 30 orders of magnitude, despite this however, the Higgs squared value is measured to be (relatively) very small. The Higgs squared mass can only take this value either if there is a very large amount of fine-tuning in the Standard Model parameters to precisely cancel out all of the large corrections, a highly unnatural and disfavoured solution, or some new physics beyond the Standard Model exists which more naturally suppresses these corrections.

The corrections to the Higgs squared mass from bosons has an opposite sign to that of the corrections from fermions. The solution to the Hierarchy problem posited in SUSY exploits this and states that if every Standard Model fermion has a bosonic superpartner, the very large corrections to the Higgs squared mass from the Standard Model particles will be neatly cancelled out by corrections from their supersymmetric partners with an opposite sign.

This model of SUSY as described is simplistic compared to viable manifestations of the theory and a significant amount more complexity must be added to make a theory that works in practice. An obvious consideration is that we observe Standard Model particles yet do not observe supersymmetric particles, meaning supersymmetric particles cannot be degenerate in mass with their Standard Model partners, otherwise they would be readily observed in current experiments. Many other such issues with SUSY in its basic form also exist. An example of an extension to SUSY to make it more compatible with observations of the real world is the Minimal Supersymmetric Standard Model (MSSM), developed in 1981 [39].

In the MSSM, a new symmetry called R-parity is introduced [38] to avoid the problem of baryon and lepton number violation and to prevent squarks mediating unphysically fast proton decay. R-parity, P_R is a multiplicatively conserved quantum number where in the MSSM, all SM particles have even R-parity ($P_R = +1$) and all SUSY particles have odd R-parity ($P_R = -1$). It is this introduction of R-parity that ultimately allows for a viable dark matter candidate to arise through two key properties it gives SUSY particles. Firstly if R-parity is totally conserved then mixing between Standard Model and supersymmetric particles is forbidden, and secondly, it means that the lightest supersymmetric particle (LSP) must be stable regardless of its mass. This is because for a decay of a supersymmetric particle with negative R-parity particle to conserve R-parity, at least one other supersymmetric particle must be present after the decay. In that case the LSP has no available supersymmetric particle to decay into as decays into heavier particles are forbidden. If all SUSY particles have high masses, it would be expected that with the exception of the LSP, they all have short half-lives and would quickly decay down to the LSP in the early universe after freeze-out. This would leave a large relic abundance of LSPs and hence the LSP is an excellent dark matter candidate.

Which particle within the MSSM (or another supersymmetric extension to the Standard Model) ends up being the LSP varies across a range of candidates depending on version of the model used since the degeneracy in mass between Standard Model and supersymmetric particle can be broken in many different ways. In many versions of the MSSM, a “neutralino” is the LSP [38]. In the Standard Model, there are four neutral states arising from the four neutral Standard Model bosons, the photon, the Z^0 , and two neutral Higgs bosons [7] which mix to form four neutralinos [38]. The lightest of these is often the LSP.

In the case of the neutralino being the LSP, the mass range predicted for the particle is of the order of GeV to low TeV with a wider range of potential cross sections. As shown in Fig. 2.8, this sits roughly in the centre of the parameter space inhabited by WIMP dark matter candidates. Also shown in this figure are the masses and cross sections of other LSP candidates, Standard Model and sterile neutrinos, and the axion which is discussed in more detail in the following section.

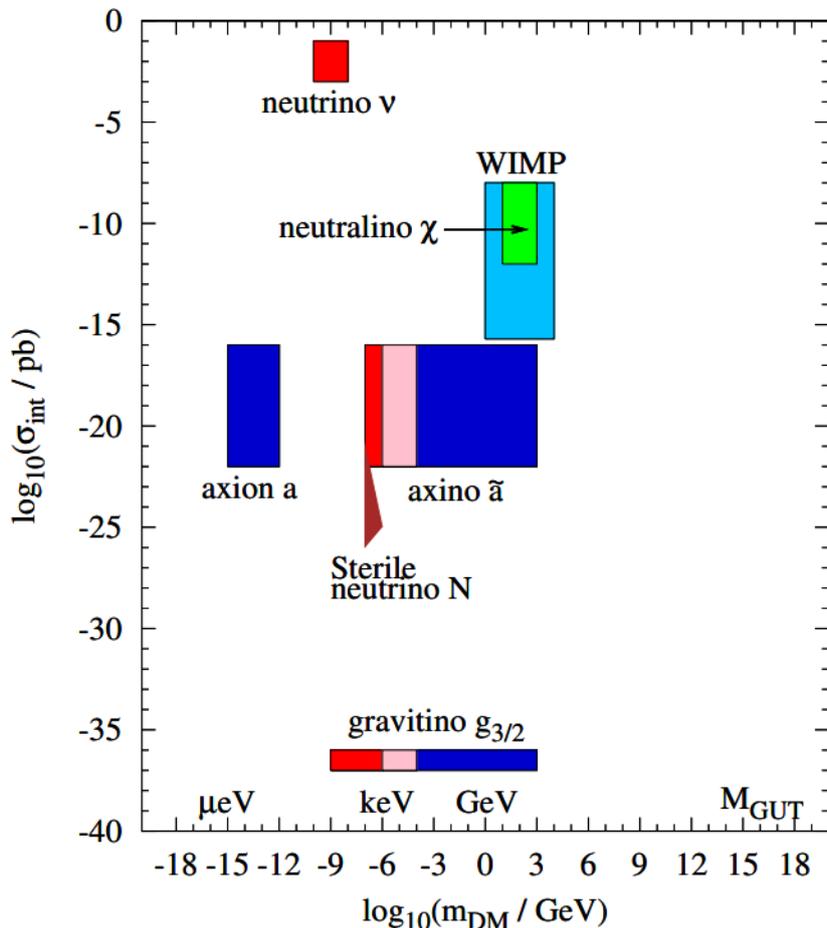


Figure 2.8: The expected masses and cross sections for a variety of dark matter candidates. Red, pink, and blue regions denote hot, warm, and cold dark matter respectively. From [30].

These regions of parameter space inform where any given dark matter detection experiment should have its efforts focused. So far, experimental searches have been concentrated around the parameter space in the green region of Fig. 2.8, but increasingly more regions of parameter space are being searched in as a positive signal from dark matter remains elusive.

Axions

There have been considerably more experimental efforts dedicated to detecting WIMPs than other dark matter candidates. There are various possible reasons for this: the link between WIMPs and SUSY given the popularity of SUSY and related searches for SUSY in particle accelerators, and the perception that WIMPs are in a region of parameter

which can more easily be probed by the experimental technology available being two major ones. However, it would not be fair to say that WIMPs have been the source of the most experimental search efforts because they are the uniquely outstanding dark matter candidate in terms of being well motivated, given the existence of an equally well motivated candidate in the axion.

The axion arises from the solution to another significant, unsolved problem within the Standard Model, the strong CP problem. The problem considered is why charge parity (CP) is conserved in quantum chromodynamics (QCD), the theory describing the interactions of the strong force in the Standard Model, when doing so requires a very fine-tuned choice of parameters [40]. In that sense the strong CP problem is similar to the hierarchy problem in that there is an area of the Standard Model that requires a high degree of fine tuning to match experimental results. An elegant solution to this issue was proposed in 1977 by Roberto Peccei and Helen Quinn [40,41] and postulates the existence of a new global symmetry which is spontaneously broken. This symmetry breaking results in a new Nambu-Goldstone boson called an axion [42] which is the theory's DM candidate.

Axions have all the necessary characteristics of DM, being stable, massive and exclusively extremely weakly interacting. Also providing they have a mass of the order of $1-100\mu\text{eV}$, they are capable of predicting the correct relic abundance of DM [43]. Since the axion mass is significantly lower than that of the WIMPs discussed, the axion would have to appear in much higher numbers than WIMPs to obtain the correct DM density. There are two main models describing axions which differ in the way they describe axion coupling to photons. These are the KSVZ (Kim-Shifman-Vainshtein-Zakharov) model [44, 45] and the DFSZ (Dine-Fischler-Srednicki-Zhitnitsky) model [46, 47]. Unlike WIMPs, the range of potential axion masses is comparatively well constrained which allows for more targeted searches to occur. No convincing signals for axions have yet been observed, although it has only been possible to search for axions in a very small region of parameter space thus far, for example [48].

2.3 Concluding remarks

As a result of the many pieces of astronomical and cosmological evidence, the existence of dark matter is widely accepted as a fact across the physics community. Attention therefore turns to experimental searches aiming to directly detect dark matter for the first time, these will be described in Chapter 3. The remainder of this thesis focuses on one specific experiment, the COSINE-100 experiment which is primarily used to search for WIMP dark matter, specifically in the context of testing the controversial claim by the DAMA collaboration of a signal which can be attributed to dark matter

Chapter 3

Experimental Searches for Dark Matter and the DAMA Controversy

As outlined in Chapter 2 the astronomical, cosmological, and theoretical case for the existence of a large amount of dark matter in the universe is now very well established. As a result, considerable effort has been put into experimental searches for a direct detection of dark matter, yet despite these various efforts, there has been no positive signal verified by other experiments that can be attributed to dark matter. This “verified by other experiments” caveat is central in the context of this work. Every experimental search for dark matter has failed to return a persistent, positive signal with one exception from the DAMA collaboration. With an extremely high confidence of over 12σ , they report a characteristic annual modulation signal with an amplitude of 0.0105 ± 0.0011 counts/(day \times kg \times keV) in the event rate of 1-6 keV signal-like events in their NaI target material [49]. The mass and cross section of dark matter implied by this result using common models of WIMP dark matter are in regions of parameter space which have since been probed by many other dark matter search experiments [50], some of which have projected sensitivity many orders of magnitude higher than the DAMA experiments. These model-dependent tests rule out the possibility that the DAMA signal is caused by dark matter from any of the mainstream models of dark matter. However, the model dependent nature of these tests and a general lack of explanation for the DAMA signal allow for the result to persist and keep the field in tension, since no model-independent test of the DAMA signal has yet been able to give a definitive conclusion. Such a model-independent test of DAMA requires an experiment using the same target material, NaI, searching for the same annual modulation signal.

In recent years, two experiments aiming to test the DAMA result have started producing results, albeit without yet reaching a definitive conclusion. These are the ANAIS and COSINE-100 experiments, the latter being the primary focus of this work. Another way to end the tension caused by the DAMA result would be to show that an alternative explanation to dark matter could explain the signal reported by DAMA. Many such explanations have been put forward, but none has received enough support to end the tension

In this chapter, first in Section 3.1 an overview of the major limit-setting dark matter detection experiments is provided, focusing on liquid Xe time projection chambers in Section 3.1.1, and others searching for low mass dark matter in 3.1.2. Then follows a description of the DAMA experiment’s results in Section 3.2 and other NaI-based experiments in Section 3.2.2, and finally potential explanations for the DAMA signal in Section 3.3.

3.1 Dark matter searches with lowest model-dependent upper limits on WIMPs

The number of dark matter detection experiments releasing results has grown from just a handful when experimental search efforts began in the 1990s to many dozens in the 2020s. Indeed, nine different experiments are listed in the Particle Data Group’s *Review of Particle Physics* [7] as having close to field-leading best constraints for dark matter candidates with different masses and from different theoretical models, and this is by no means an exhaustive list of dark matter search experiments. Fig. 3.1 shows a pictorial representation of some of the different dark matter detection efforts, capable of searching for a ‘mainstream’ 60 GeV WIMP [51] where a clear improvement in experimental sensitivity over time can be seen. It is well beyond the scope of this work to attempt to summarise all of these efforts. Instead, discussion in this chapter focuses predominantly on the classes of experiments attempting to set the tightest constraints on standard WIMP dark matter, with a view to providing the context for the DAMA result, prior to discussing the DAMA result in greater detail.

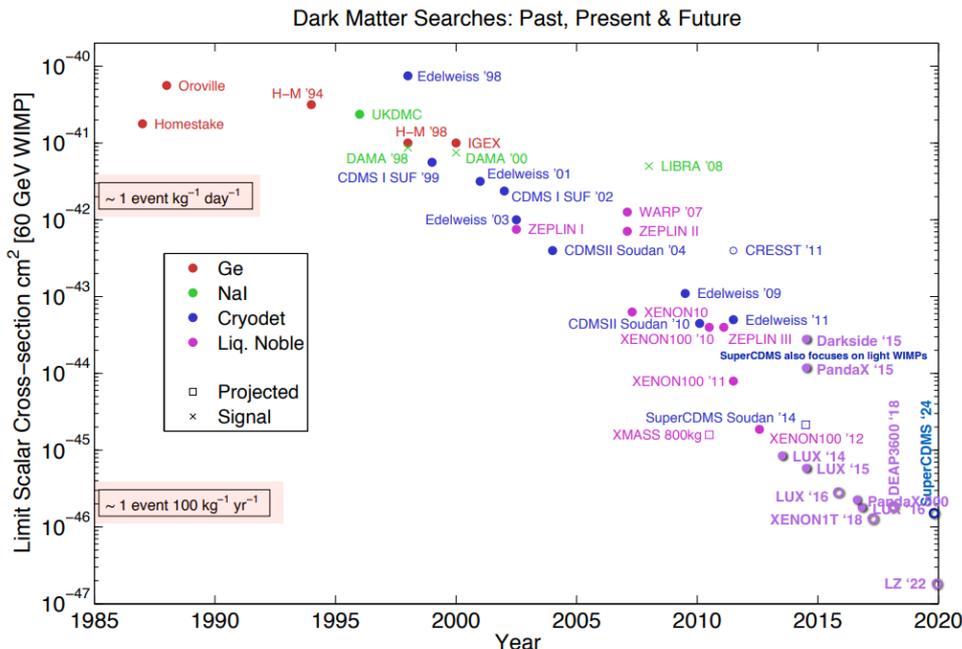


Figure 3.1: A (non-exhaustive) visual representation of dark matter detection experiments over time shown by their sensitivity to a 60 GeV WIMP. From [51].

3.1.1 Liquid Xe experiments

The class of dark matter detection experiment that has achieved the highest sensitivity and therefore set the most stringent constraints on the properties of dark matter are liquid xenon (LXe) time projection chambers (TPCs). LXe TPCs utilise a large amount (several tonnes in the case of leading experiments) of LXe as their target material since Xe has a high atomic mass, making it a good kinematic match to WIMPs of $\mathcal{O}(100 \text{ GeV})$, as well as having a very large spin-independent scattering cross section [52]. When a WIMP scatters off a Xe nucleus, it is expected to produce two distinct, measurable signals, as shown in Fig. 3.2, a schematic showing the operating principles of an LXe TPC. The first

is from scintillation photons which are detected promptly by an array of photomultiplier tubes (PMTs) at the bottom of the LXe container: this signal is referred to as “S1”. Also produced during a scattering event are ionisation electrons which once liberated from their parent Xe atom by the collision, drift towards the top of the container under the influence of an externally applied electric field where they pass through a vapour phase. In doing so they produce electroluminescence photons which are also measured by PMTs at the top of the chamber. This signal is referred to as “S2” [52]. LXe TPCs are able to achieve very high sensitivities by exploiting these dual signals.

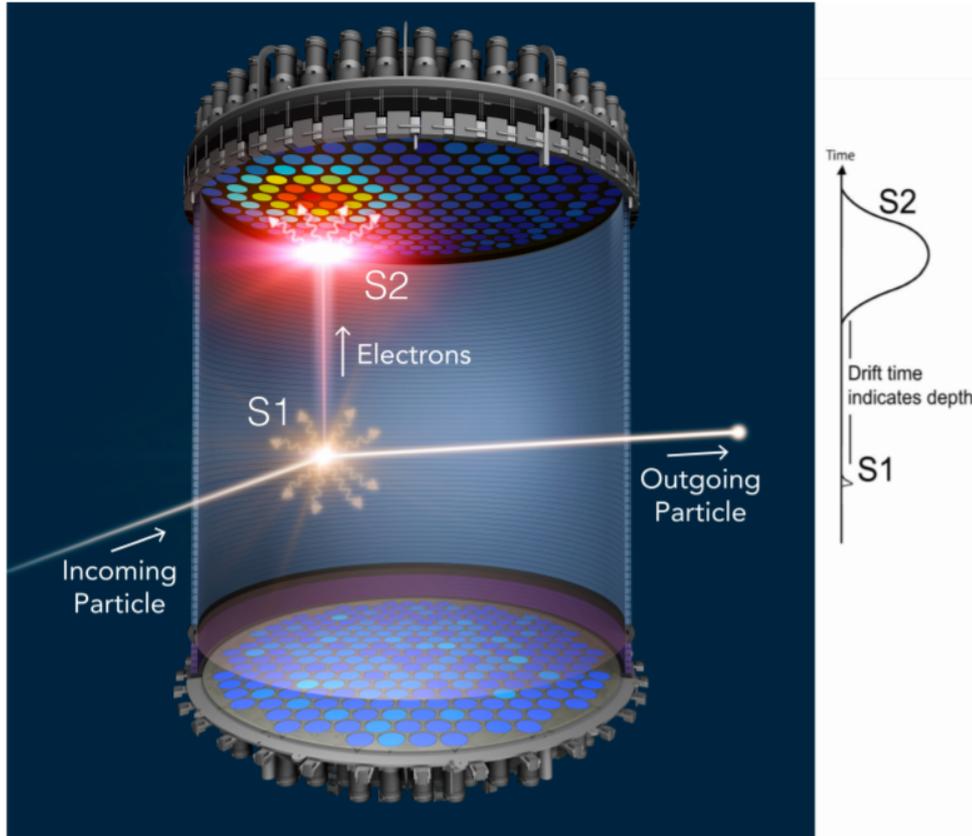


Figure 3.2: Schematic showing the operating principle of liquid Xe TPCs. An incoming particle scatters in the liquid producing photons which are measured immediately (S1), and ionisation electrons which drift in the presence of an electric field to the top of the chamber where they are recorded (S2). The depth of the event can be calculated difference in time between these two signals. From [52].

These dual signals have three very useful properties in terms of dark matter detection. Firstly, by measuring the time between the signals, the depth of the event in the chamber can be calculated (in the z direction) which combined with the location of the PMTs recording events (in the x, y plane) gives a three-dimensional event reconstruction. This is particularly useful since for experiments with sufficiently large amounts of LXe, the inner volume of the LXe is effectively shielded from background events by the volume of LXe outside the region. Consequently, a purer signal is expected in this volume, known as the “fiducial volume”. Events outside the fiducial volume are not usually considered signal events in dark matter searches, and it is this three-dimensional event reconstruction that allows for this cut to be made. The second important aspect of the dual signal of events in LXe TPCs is it allows for a powerful coincidence cut to be applied when selecting

signal-like events, since it is very unlikely for a noise event not caused by a genuine scintillation to produce both an S1 and S2-like signal simultaneously. Additionally, the ratio of scintillation photons in S1 to S2 allows for nuclear recoil events (which may be caused by $\mathcal{O}(100 \text{ GeV})$ WIMPs) caused by the incident particle scattering off a Xe nucleus to be differentiated from electron recoil events (which are unlikely to be caused by $\mathcal{O}(100 \text{ GeV})$ WIMPs [50]) where the incident particle scatters off the atomic electron cloud [52].

Another area where leading noble liquid TPCs are able to excel over other types of dark matter detection experiments in terms of sensitivity is their scalability in the mass of the target material. Experiments using crystalline scintillation materials are limited by the size of crystal that can be grown. In the case of NaI experiments there have been difficulties of finding ovens big enough to grow crystals larger than a few tens of kg for example. This is in addition to the challenges of growing large crystals with the requisite composition and structure. An array of smaller crystals can be used, but at the cost of increasing complexity when modelling the expected measured signal. On the other hand, it is relatively much easier to create a large volume of homogeneous liquid noble gas, once the scintillation material has been obtained.

As a result, the three largest LXe TPCs currently in operation, LZ [53], XENONnT [54], and PandaX-4T [55] have fiducial masses of 5.5, 4.4 and 3.7 tonnes of LXe respectively¹: this is at least an order of magnitude larger than the largest NaI, Ge, Si, CaWO₄, etc. dark matter detection experiments. The combined effects of these factors allow for these large-scale LXe TPCs to search for a dark matter signal characterised by an excess of signal events above the expected (very low) background. NaI-based experiments have a relatively higher level of background events so such an excess of signal events is not expected, hence the search for annual modulation-like signals to achieve the highest sensitivity in this class of experiment.

These three largest LXe TPCs, LZ, XENONnT, and PandaX-4T are currently competing to set the most stringent limits on spin-independent WIMP-nucleon cross section for WIMPs with masses in the 10s to 100s of GeV range. At the time of writing these best limits are set by the LZ experiment from their first 60 days (approximately 0.9 tonne \times years) of data [53]. These limits are shown in Fig. 3.3 in comparison with recent best limits from the XENON collaboration from an approximately 1.1 tonne \times year exposure [54]. Also shown are the best limits from PandaX-4T [55]. Note that the DAMA result which implies a WIMP with a WIMP-nucleon cross-section of around 10^{-41} cm^2 and these masses does not appear on this plot as it is excluded by several orders of magnitude in the WIMP-nucleon scattering models used.

¹Note that is only the fiducial mass, the total mass of LXe used is larger than these values

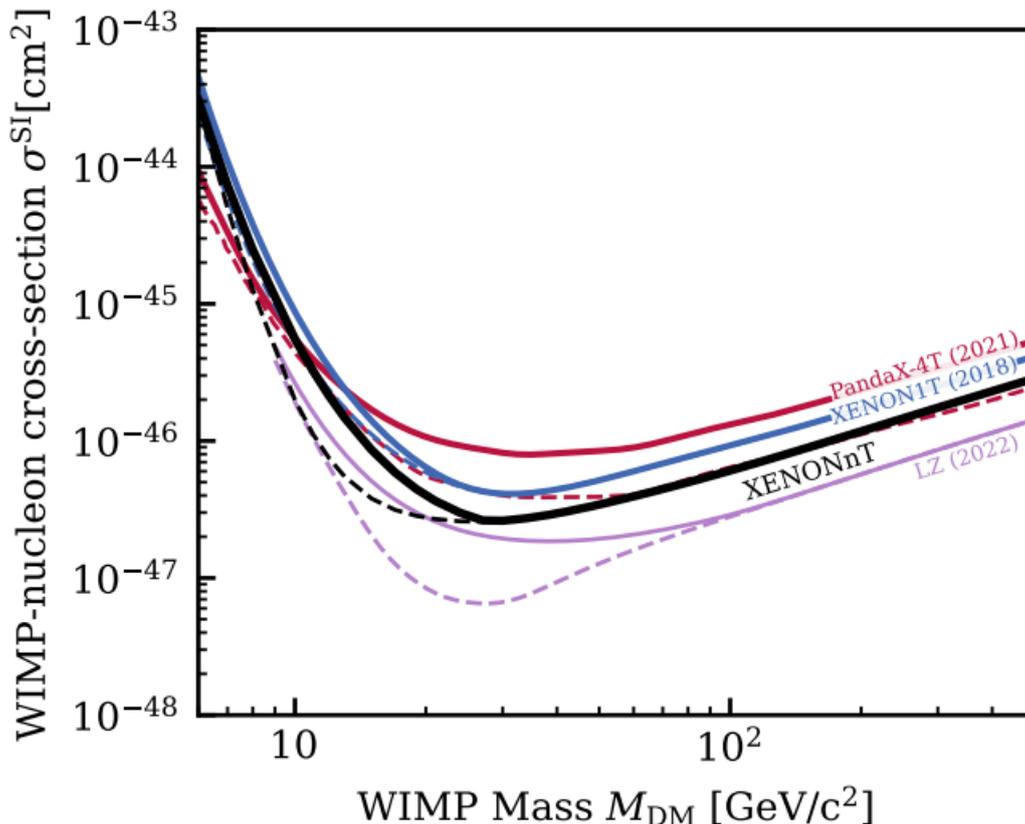


Figure 3.3: 90% confidence upper limits on spin-independent WIMP-nucleon cross section set by the PandaX-4T (red lines) [55], XENONnT (blue and black lines) [54], and LZ (purple lines) [53] LXe TPC experiments. Adapted from [54].

3.1.2 Lower mass WIMP searches

While LXe TPCs lead in terms of setting the most stringent limits on the properties of dark matter at intermediate and large WIMP mass-scales, as can be seen from Fig. 3.4 from [50]² showing the limits of dark matter search efforts for spin-independent elastic WIMP-nucleus scattering using a range of target materials and search methods, at lower WIMP masses a variety of other experiments are competing to set the tightest upper limits. Again, it is worth noting the region of parameter space containing the DAMA result (red shaded regions) and the degree to which other experiments have surpassed the requisite sensitivities to detect WIMPs of this type.

Unlike the 10-1000 GeV WIMP mass range where LXe TPCs achieve the highest sensitivity, at lower masses there is no single search technique that stands out, with LXe, liquid Ar, Ge, CaWO₄, Ne, and Si all setting competitive spin-independent limits for < 5 GeV WIMPs [7]. At these WIMP masses, a wider variety of search techniques are used, for example exploiting the Migdal effect [56], or ionisation only signals [57] for MeV-scale WIMPs.

²This figure is from a report from APPEC Scientific Advisory Committee in 2021 and shows the best limits from experiments at the time. Since then further results have been published and more stringent limits set, however it is difficult to find a suitable plot showing all these results together, such is number and frequency of experiments producing limit-setting results.

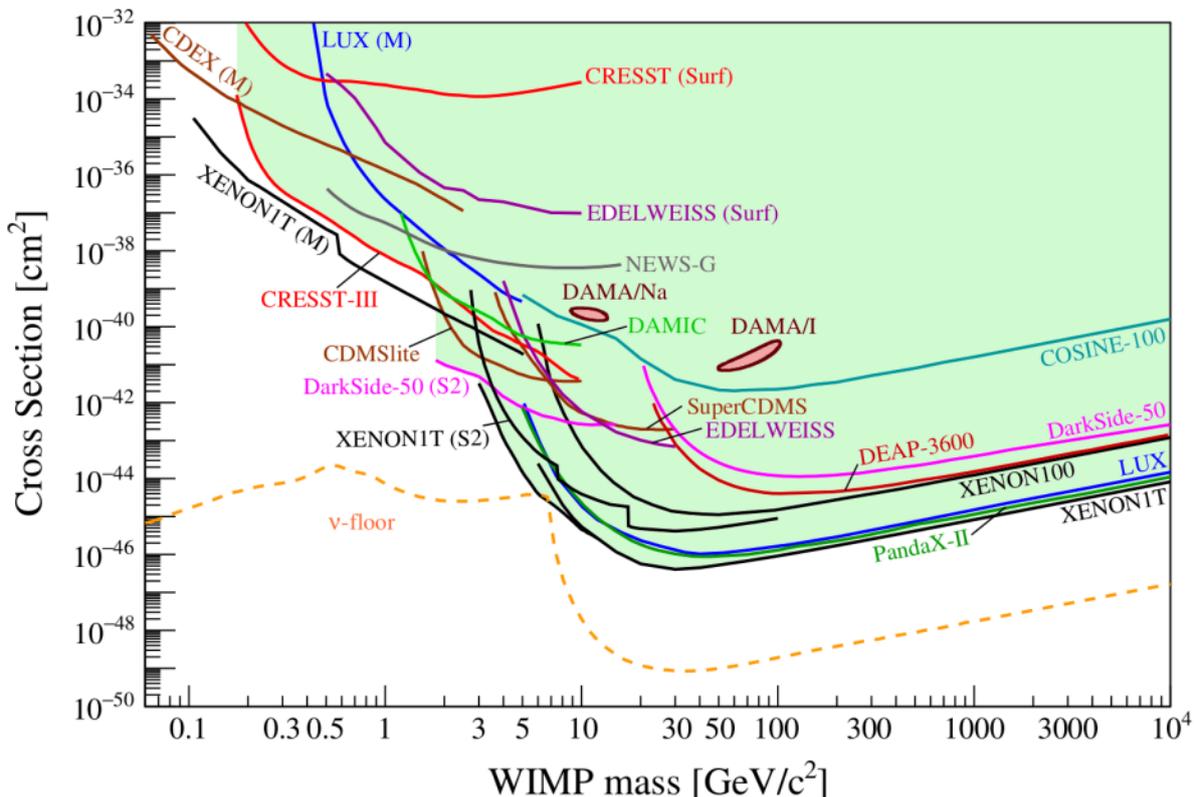


Figure 3.4: Limits from various dark matter search efforts for a spin-independent elastic WIMP-nucleus scattering assuming the standard parameters for an isothermal WIMP halo, correct as of April 2021. From [50].

Firstly, LXe TPCs are able to set competitive limits at low mass by using a technique different from that applied at the higher mass regime. Specifically, when searching for > 10 GeV-scale WIMPs, nuclear recoil events are selected as candidate signal events, whereas at keV-scales, which are not limited to WIMP searches, electronic recoil events are instead considered [58,59], the discrimination of these types of events is made possible by the different ratio of S1 to S2 signals from nuclear and electronic recoil events. An example of competitive limits from searches for electronic recoils in an LXe TPC can be found in results from XENONnT [59].

Xe is not the only noble liquid used in TPCs in leading dark matter search efforts with liquid Ar (LAr) also proving a viable target material. Since Ar has a lower atomic mass than Xe, recoils from WIMP events in LAr are more energetic than in LXe, compensating for the lower cross section in LAr than LXe [60] at WIMP masses below approximately 55 GeV. One example of an LAr TPC is the DarkSide-50 experiment with an operational principle similar to that of the LXe TPCs described in Section 3.1.1. DarkSide-50 is the experiment most sensitive to WIMP interactions for WIMP masses between 1.8 and 3.0 GeV [60] at the time of writing. DarkSide-50 is also used in searches for lower mass dark matter, down to 40 MeV, utilising the Migdal effect which predicts the production of additional electrons during dark matter-induced nuclear recoils. This results in nuclear recoil events which typically would be below the low-energy detection threshold for a given experiment [56]. Using this technique, DarkSide-50 has also set competitive limits for spin-independent interactions down to the tens of MeV scale [61]. The Migdal effect is also utilised in LXe TPCs in a similar manner, for example in XENONnT [62].

For particle candidates of mass $\mathcal{O}(0.1 \text{ GeV})$, competitive limits on spin-independent

dark matter are set by the Cryogenic Rare Event Search with Superconducting Thermometers (CRESST-III) experiment which uses as its target material individual CaWO_4 , Al_2O_3 , Si or LiAlO_2 crystals at very low temperatures (around 15 mK) acting as calorimeters [63]. Rather than in experiments utilising a scintillation material exclusively and measuring photons produced in a scattering event from a dark matter particle, calorimeter-based experiments instead search for the temperature increase as a result of phonons being produced during a scattering event, using superconducting Transition Edge Sensors (TES) devices in the case of CRESST-III.

Such experiments are able to produce competitive results with low masses of target material (less than a gramme) since they operate at very low energy thresholds where the expected flux of WIMPs is significantly higher than at higher masses. The most stringent limits from CRESST-III [63] (shown in Fig. 3.5) come from analysis of data from their Si crystal which gives leading limits for spin-independent WIMP-nucleon scattering for dark matter candidates with mass around 0.2 GeV. Note that this exclusion plot is shown over others from other experiments simply as a representative example.

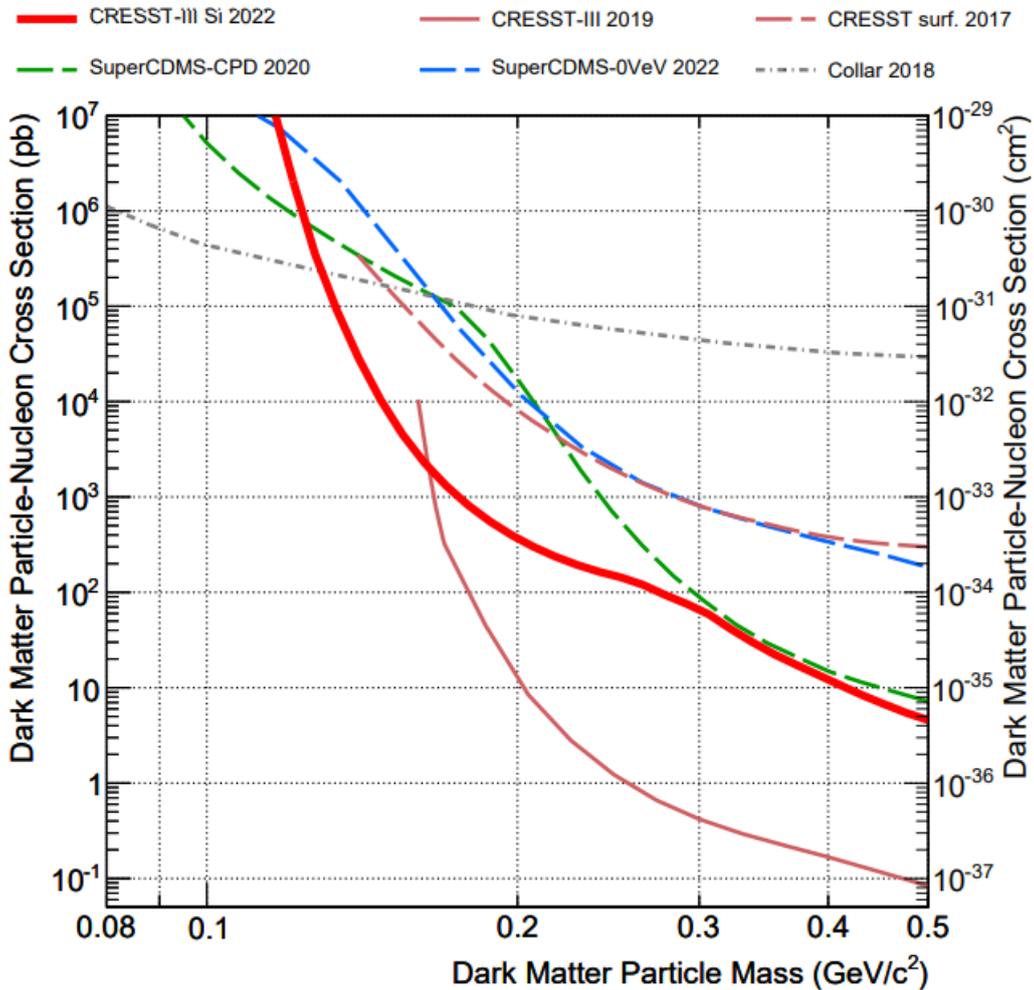


Figure 3.5: 90% confidence limits for spin-independent WIMP-nucleon scattering for $\mathcal{O}(0.1 \text{ GeV})$ WIMPs from the CRESST-III experiment using a Si crystal calorimeter. Adapted from [63].

3.2 The DAMA result and tests from other NaI experiments

As can be inferred from the lack of mention of NaI-based experiments in the discussion of the most stringent limits on WIMP properties in Section 3.1, NaI-based dark matter detection experiments are not expected to be the most sensitive experiments when searching for WIMP dark matter within the confines of most viable dark matter models. Consequently, the large-scale LXe TPC experiments mentioned are able to exclude the DAMA result by several orders of magnitude in the spin-independent WIMP/standard halo model framework. However, all of these tests of the DAMA result by experiments using a different target material and search method are inherently model-dependent, since some model of dark matter and its interactions with a given target material has to be assumed to compare the results of detection efforts with different target materials. For example, it has been suggested [64] that the specific crystalline properties NaI could raise the material’s sensitivity to dark matter above that of other target materials, as discussed in more detail in section. 3.3.3. For the DAMA result to be truly excluded, it must either be tested in a model-independent way by an experiment using the same target material and search method, or excluded if the signal measured could be satisfactorily explained by an alternative source to dark matter.

In this section, the “DAMA result” and the DAMA experiments will be described in more detail, providing the context for a discussion around some of the experimental and theoretical tests of DAMA that have been carried out thus far.

3.2.1 The DAMA experiments and results

Annual modulation signals expected from dark matter

As mentioned in Section 3.1.1, NaI-based dark matter detection experiments including DAMA are not sensitive enough to set competitive limits by searching for an excess of dark matter signal-like events above the expected background. Instead, NaI-based experiments generally exploit the relative motion of the Earth around the galactic centre (shown in Fig. 3.6) and search for the resultant annual modulation expected in their signal-like data. This expected annual modulation in the rate of dark matter-induced events results from the differing expected bulk motions of the dark matter and baryonic matter components of a galaxy. Since baryonic matter undergoes inelastic collisions, to conserve angular momentum after the loss of kinetic energy from collisions, the baryonic matter of a galaxy collapses into a disc. Conversely, the elastically scattering dark matter component of a galaxy is expected to remain roughly spherical [65]. Furthermore, since the baryonic matter component of a galaxy has collapsed towards the galactic centre more than the dark matter component, the baryonic matter component of a galaxy is expected to be rotating faster than the dark matter component [65].

This relative difference in rotation between the dark and baryonic matter components of a galaxy, including the Solar System in the Milky Way, causes the baryonic matter to experience a flux of dark matter, shown for the Solar System by the red arrows in Fig. 3.6. In the case of the Solar System, the sun orbits the galactic centre with an average velocity of around 230 km/s so is expected to experience a constant flux of dark matter proportional to this velocity. However, the orbit of the Earth is tilted at an angle of around 60° relative to the Sun’s velocity vector, changing the expected dark matter flux expected on Earth over the course of a year with the flux peaking in June. This changing

flux measured on Earth is known as the “WIMP wind”

As a result, the measured event rate at a dark matter detection experiment over time, shown in a cartoon example in Fig. 3.7 for an experiment with a linearly decreasing background over time, is expected to modulate in a sinusoidal manner throughout the year. Consequently, even if background events cannot be identified with sufficient precision to search for an excess of signal events, the annual modulation in overall event rate may still be measured if the modulation amplitude is sufficiently large relative to the rate of background events. This annual modulation is the signal searched for by DAMA and other NaI-based dark matter detection experiments and subsequently reported by DAMA.

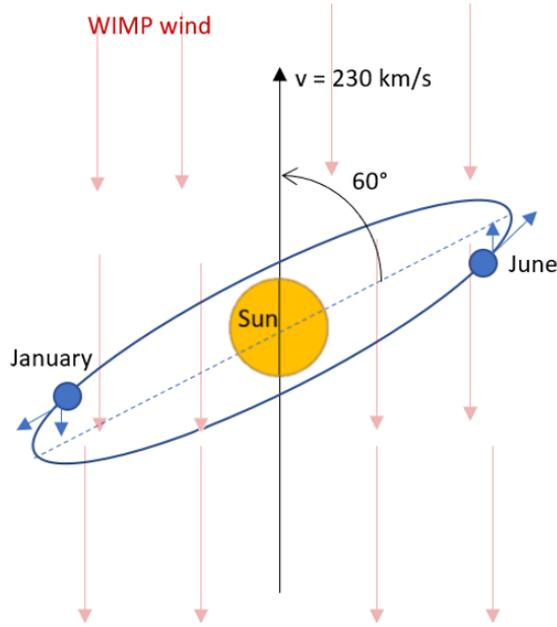


Figure 3.6: A cartoon representation showing how the Earth’s (blue circle) relative motion around the galactic centre, and consequently through the Milky Way’s dark matter halo, as a result of orbiting the Sun affects the flux of dark matter going through the Earth (the “WIMP wind”) at different times of the year. This flux peaks in June with a minimum six months earlier/later due to the 60° angle at which the Earth orbits the Sun, relative to the Sun’s path around the galactic centre/through the dark matter halo (black arrow).

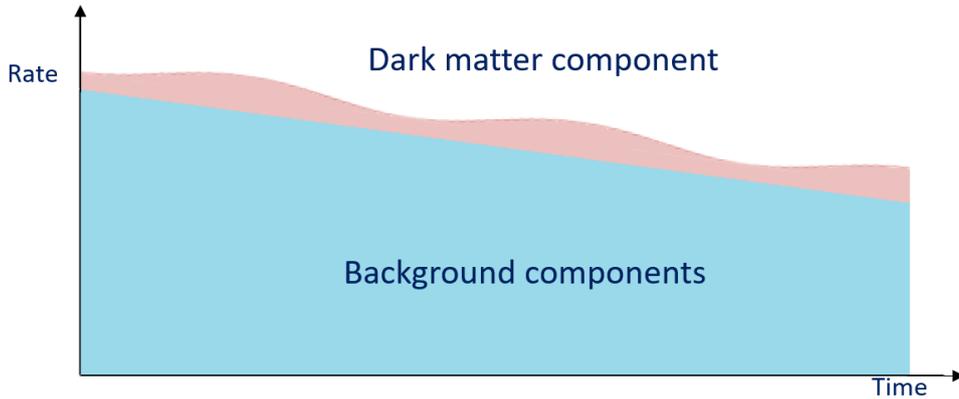


Figure 3.7: A simplified representation of the expected measured event over time for a dark matter detection experiment with a linearly decreasing background rate (blue shaded area). The event rate of dark matter-induced events (red shaded area) is expected to vary sinusoidally over time, resulting in a potentially measurable signal.

The DAMA experiments

The “DAMA experiments” are in fact two generations of NaI-based experiments, located at the Laboratori Nazionali del Gran Sasso in Italy. The first of these two generations is the DAMA/NaI experiment which used nine 9.70 kg high purity Thallium-doped NaI crystals (NaI(Tl)) in a 3×3 array coupled to a PMT at each end, and collected data over seven annual modulation cycles between 1995 and 2002 [66]. The DAMA/NaI experiment was decommissioned and upgraded in 2002 to a second generation experiment known as “DAMA/LIBRA” (Large sodium Iodide Bulk for RAre processes), mainly to allow for the original nine NaI(Tl) crystals to be replaced with 25 new NaI(Tl) crystals (in a 5×5 array), all with masses of 9.7 kg [66]. All of the NaI(Tl) crystals used in the DAMA experiments were grown by Saint Gobain Crystals and Detectors company, which has since stopped producing high purity NaI(Tl) crystals, meaning future NaI-based experiments had to find different manufacturers for their crystals. The DAMA/LIBRA apparatus received two further upgrades, one in 2008 to replace a broken PMT and implement other minor hardware improvements [67], then in 2010 to replace all of the PMTs in the experiment with ones with higher quantum efficiencies, allowing for the software low-energy threshold to be reduced from 2 keV as in DAMA/NaI and DAMA/LIBRA-phase 1, to 1 keV [68]. The DAMA/LIBRA apparatus is described as “DAMA/LIBRA-phase-1” before this second upgrade and “DAMA/LIBRA-phase 2” thereafter. DAMA/LIBRA-phase 2 continues to take data into 2023.

The shielding structure is broadly the same for all the variations of the DAMA experiments and is shown in Fig. 3.8. The crystals and their associated PMTs and light guides are housed within a low-radioactivity Cu box that is filled as much as possible with low-radioactivity Cu bricks and a high purity N_2 atmosphere is maintained. The copper box is then sealed, and surrounded by > 10 cm of low-radioactivity Cu shielding, 15 cm of low-radioactivity Pb shielding, and 1.5 mm of Cd shielding. This arrangement is housed within a 1m concrete box where the gaps between the concrete and Cd (between 10 and 40 cm) are filled with polyethylene and paraffin. The passive shielding arrangement is sealed within a plexiglass box [69].

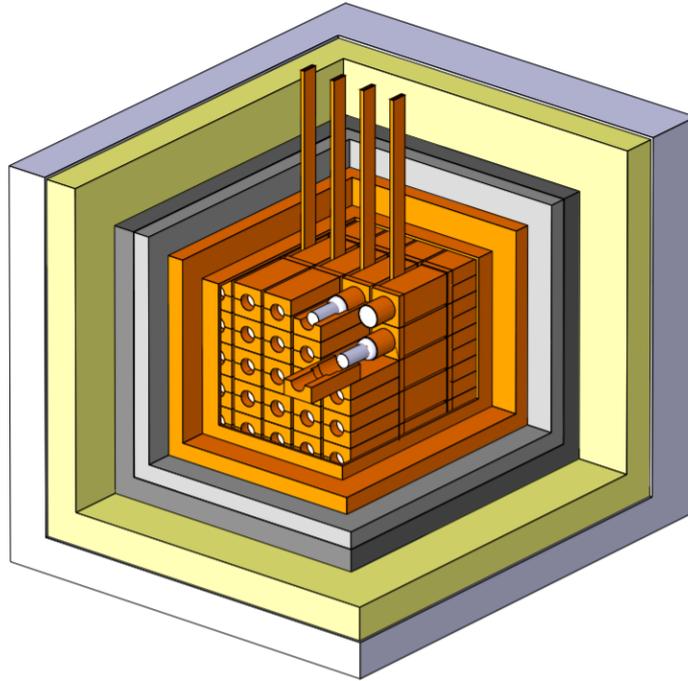


Figure 3.8: Diagram of DAMA setup showing the 25 NaI(Tl) crystals (inner orange matrix) shielded by copper bricks contained within a sealed copper box (outer orange layer), a Cu/Pb/Cd-foils/polyethylene/paraffin shield (two inner grey layers, yellow layer, and outer grey layer), and a 1 m concrete shield (not shown in image). The crystal matrix is kept in a high purity nitrogen atmosphere within the copper box. Adapted from [69]

As well as shielding from external backgrounds, DAMA, like other dark matter detection experiments, must reduce internal backgrounds from radioactive decays as much as possible by achieving high radio-purity in the detector material. In particular NaI experiments are at risk from background events from ^{40}K decay which produces a 3.2 keV electron directly in the middle of the dark matter signal region (1-6 keV). DAMA claims a very low ^{40}K contamination of < 20 ppb [69], a lower level of contamination than has been achieved subsequently by other NaI based experiments. Since no active shielding is used in the DAMA experiments, event selection consists of pulse shape discrimination and coincidence cuts where coincident events are rejected due to the low probability of two dark matter-induced events happening in a short time-frame [69].

The DAMA result and analysis technique

The data from these two experiments are combined and so far include 22 yearly cycles at a 2 keV low-energy threshold, shown for DAMA/LIBRA in Fig. 3.9, along with results from eight yearly cycles of DAMA/LIBRA-phase 2 with a 1 keV low-energy threshold [70]. The resultant exposures in these two data-sets are 2.86 tonne \times year and 1.53 tonne \times year respectively [70]. Residual data-sets are then calculated by subtracting the expected background events in roughly yearly periods, and these data-sets are fit to sinusoidal function with a fixed period and phase of one year and 152.5 days (in line with predictions from the standard halo model) and the best-fit modulation amplitude found. From this, DAMA reports a positive annual modulation with an extremely high confidence level (C.L.) in both these data-sets, quoting a 13.7σ C.L. in the 2-6 keV region of interest which they attribute to the effects of dark matter [70].

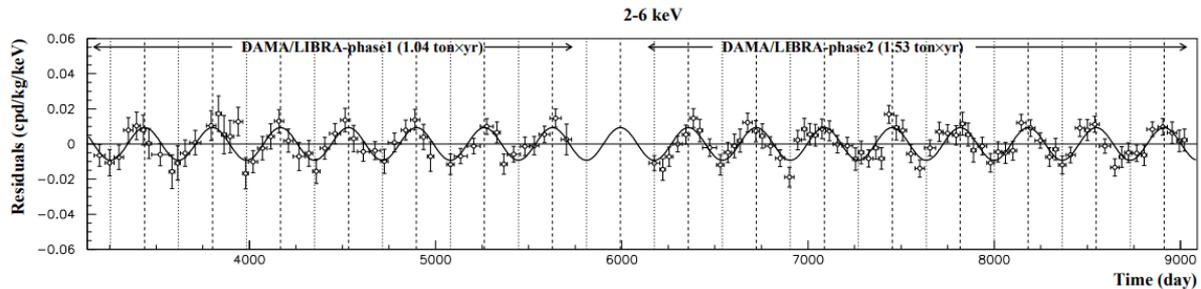


Figure 3.9: The residual event rate for single-hit events in the 2-6 keV region of interest, measured in DAMA/LIBRA-phase 1 and 2 (data points) and an annual modulation fit to these data (solid line) with a fixed period and phase of one year and 152.5 days (in line with predictions from the standard halo model). From [70].

The calculation of these residual data-sets which are subsequently fit to an annual modulation function has been investigated as one of the potential non-dark matter sources of the annual modulation observed in the DAMA data as will be discussed in more detail in Section 3.3.1. The expected background rate used in the calculation residuals is generated based on the measured contamination of radiogenic components in the detector materials [69]. This approach is different to that taken by other experiments that search for a dark matter-induced annual modulation signal, for example COSINE-100 [1], where rather than calculating then fitting residuals, the time-dependent background activities are included in the fit model.

3.2.2 Results from other NaI-based experiments

The gold standard in terms of a test of the DAMA result is an annual modulation search from another NaI-based dark matter detection experiment which has sufficiently high statistics to conclusively confirm or refute the DAMA signal, by either measuring a similar annual modulation or finding no modulation. The two experiments closest to achieving this are the COSINE-100 and ANAIS experiments which have both released results of annual modulation searches from three years of data [1, 71], followed by the SABRE North and South twin experiments which are still in the development phase [72]. These experiments and their searches for dark matter, both from an annual modulation signal and otherwise, are described below.

The COSINE-100 experiment

A full description of the operation of the COSINE-100 experiment is saved for a more detailed discussion in Chapter 4. For the purposes of this discussion of the experiment's results it is sufficient to say that the operational procedure of COSINE-100 is broadly similar to that of DAMA with some key differences. These are the use of liquid scintillator and plastic scintillator active shielding, designed to tag and veto background events from radiogenic decays and cosmic ray muons respectively, a more complex process of identifying and rejecting PMT-noise events, utilising machine learning techniques, and the inclusion of time dependent backgrounds in the annual modulation fit model described above. These features are described in more detail in [73], [74], and [1] respectively, and in Chapters. 4 and 9. COSINE-100 also has less active NaI(Tl) target material than the DAMA experiments, with 61.3 kg of active target material.

The COSINE collaboration has released results from annual modulation searches using 1.7 and 2.8 years of data [1, 75], the latter having a total exposure of $173.5 \text{ kg}\times\text{year}$. However, in both cases, there were not sufficient data to conclusively confirm or refute the DAMA signal for both the fixed and floated-phase case. Updated annual modulation searches with 4.8 years of COSINE-100 data are described in Chapters. 7, 8, 9.

The COSINE-100 experiment has also been used to test the DAMA result in a model-dependent manner and was the first NaI experiment to do so with a comparable sensitivity to DAMA. These model-dependent tests were done using 59.5 days, then 1.7 years of data [3, 75] and in the case of the latter result, excluded the DAMA result by over an order of magnitude across various WIMP hypotheses. Exclusion limits compared to the DAMA signal region from both these data-sets are shown in Fig. 3.10 for both scattering off both Na and I nuclei for WIMP-proton spin-independent interactions [3]. In these cases, rather than searching for an annual modulation in the single-hit event rate, an excess of events above the expected background is searched for. This expected background in the latter analysis comes in the form of a time-dependent GEANT4 simulation of the activity in the NaI crystals and was developed by studying the first 1.7 years of COSINE-100 data, but excluding the 1-6 keV ROI. More details of this time-dependent background model can be found in [74].

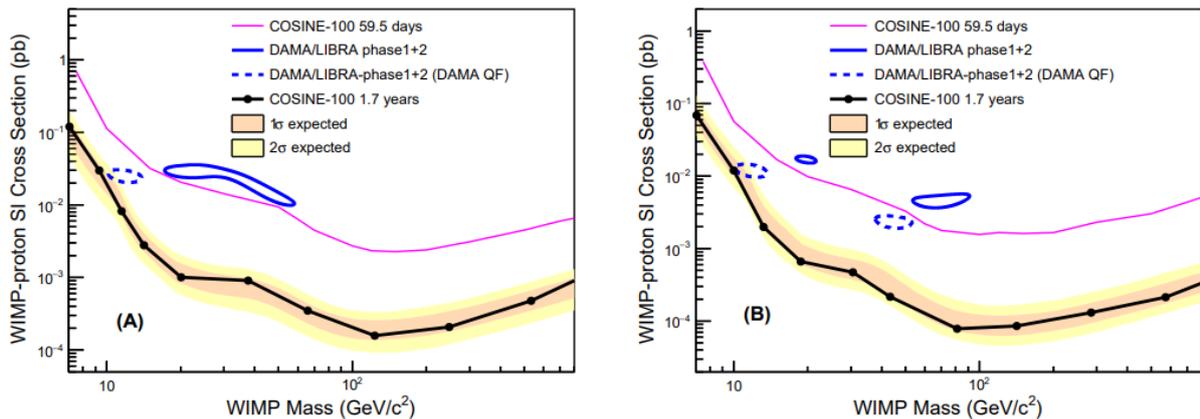


Figure 3.10: Model-dependent exclusion limits for WIMP-proton spin-independent interactions on Na nuclei (left plot) and I nuclei (right plot) in the COSINE-100 NaI(Tl) crystals for both the first 59.5 days and 1.7 years of COSINE-100 data (pink and black curves). The parameter space implied by various interpretations of the DAMA signal (blue contours) are clearly excluded. The 1σ and 2σ confidence regions for the 1.7 year COSINE-100 data-set are shown by the yellow and peach shaded regions. The DAMA results are presented with two different quenching factors, the original value used by DAMA (dashed contour), and an updated value (solid contour) From [3].

One important consideration in interpreting the measured signals in NaI-based experiments is the quenching factors of the detector material; the multiplicative factor used to convert electromagnetic recoil energies measured in the crystals to total recoil energies. The quenching factor used by DAMA in their analysis is significantly smaller than the quenching factor that has subsequently been used in most of the tests of DAMA [3]. As such, when testing the DAMA result it is useful to use test quenching factors to ensure a result that is comparable with both DAMA and other NaI-based experiments. As can be seen in Fig. 3.10, the DAMA result is excluded for both quenching factors (the DAMA

results with different quenching factors are shown by the solid and dashed blue contours) by the results from the 1.7 years COSINE-100 data.

Other NaI-based experiments

The ANAIS (annual modulation with NaI scintillators) also operates on a similar principle to the DAMA and COSINE-100 experiments, in this case with a target mass of 112.5 kg of NaI(Tl) and a plastic but not liquid scintillator active veto system [71], operating at the Canfranc Underground Laboratory in Spain. ANAIS have most recently released annual modulation search results for analysis on three years of data with a total exposure of 314 kg×year [71] where they are also not able to conclusively confirm or refute the DAMA signal. These results were subsequently updated through an improved method of removing PMT noise at low energies using machine learning techniques where similarly the DAMA result was neither excluded nor corroborated [76]. Unlike the COSINE-100 result though, the ANAIS search was able to mildly exclude the DAMA signal with a confidence level of around 3σ in the fixed-phase case. This is due to measuring a small negative modulation amplitude for both the 1-6 and 2-6 keV ROIs.

The twin SABRE experiments [72] are a third NaI-based test of DAMA with similar properties to the COSINE-100 and ANAIS experiments that was designed and began preliminary testing at around the same time as COSINE-100 and ANAIS. However, SABRE has not yet reached a stage where it is possible to produce physics data in the context testing the annual modulation signal from DAMA. A unique feature of the SABRE experiments is that it was proposed that two identical detectors would be built, one at the Laboratori Nazionali del Gran Sasso (LNGS) in Italy, the same location as DAMA, and the other at the the Stawell Underground Physics Laboratory in Australia. In having experiments in both the Northern and Southern hemispheres of the Earth, any annual modulation induced by seasonal effects rather than from dark matter would be apparent since these effects (one suggested source of the DAMA signal) would be out of phase in the two experiments by around six months, due to their latitude.

The COSINUS experiment is another NaI-based experiment still in the developmental phase [77]. COSINUS will also be based at LNGS but will use NaI at cryogenic temperatures, outfitted with TES sensors allowing for electron and nuclear recoil events to be differentiated at an event-by-event basis, and also having a lower low energy threshold for nuclear recoil events. The first dark matter search results from the COSINUS experiment are expected in 2024. Two other NaI-based experiments that should be mentioned are the DM-Ice experiment and the PICOLON experiment. DM-Ice is based in the Antarctic at the same site as the IceCube Neutrino Observatory to exploit the same phase shift in seasonal effects as SABRE. DM-Ice collected data with 17 kg of NaI(Tl) for 3.6 years up to 2015, but was not sufficiently sensitive to test the DAMA result [78]. PICOLON is based at the Kamioka Underground Laboratory in Japan and started initial testing in 2021 [79], but has not yet released any dark matter search results.

3.3 Alternative explanations for the DAMA result

Unsurprisingly given the numerous experiments that contest the DAMA result, various explanations and theories have been put forward to explain why DAMA should measure and apparently dark matter-induced signal when other experiments do not. These explanations can be split into three broad categories: those which suggest the modulation measured comes from some annually modulating background in the DAMA experiment

that has not been correctly accounted for, those which suggest that DAMA’s method of analysis has artificially introduced some modulation in the data that does not really exist, and finally those that suggest DAMA (and by extension other NaI experiments) in fact measures a region of parameter space not accessible to other classes of experiments. These hypotheses will be discussed in more detail in this section.

3.3.1 DAMA modulation signal from modulating backgrounds?

One explanation of why the DAMA signal has not been corroborated by other experiments is simply that the annual modulation measured comes from a hitherto unaccounted for annually modulating background measured by DAMA. One such modulating background is cosmogenic muon-induced events [80, 81]. High energy muons generated in showers from cosmic rays are sufficiently penetrating to be able to reach underground laboratories such as Gran Sasso where they can produce a signal in detectors which can mimic a dark matter signal. This is either by directly interacting with the target material, or by producing neutrons in interactions with other material in the detector which then go on to generate a signal [81, 82].

The muon flux on Earth and hence in underground detectors is a function of atmospheric temperature, since at lower temperatures the atmosphere is denser and hence attenuates more cosmic rays, preventing them from reaching the surface. Consequently cosmic ray muon-induced events are a candidate for an annually modulating background on account of the annual temperature changes on the surface of the Earth. However, this hypothesis has not been able to satisfactorily explain the DAMA signal on account of the peak muon flux being roughly 30 days out of phase with the DAMA signal [83].

An additional contribution to the muon signal from solar neutrino-induced muon events has also been suggested [80] to reconcile this difference in phase, since the flux of solar neutrinos on Earth peaks in early January due to the Earth’s eccentric orbit [80]. A combination of cosmic ray and neutrino induced muon signals better describes an annual modulation in phase with the DAMA signal, however more accurate simulations of the muon and neutrino induced signals at Gran Sasso have shown that these backgrounds are unable to account for the DAMA signal [82].

A second source of backgrounds in DAMA that has been postulated as a driver of the annual modulation is the direct decay of ^{40}K to the ground state of ^{40}Ar by electron capture, which is the rarest of the ^{40}K decays to ^{40}Ar with the production of a 3 keV Auger electron [84]. This is an irreducible background for DAMA in their signal region [84]. In DAMA’s calculation of their event rate, the single-hit background rate is calculated based on measurements of concentrations of said backgrounds in the detector. These backgrounds are then subtracted from each annual cycle to give the residual recoil rate [69]. In [84], it is suggested that the contribution of direct decay of ^{40}K to the ^{40}Ar ground state might be under-accounted for in this background subtraction process and as such the conclusion of an annual modulation with a low modulation fraction as reported by DAMA may be challenged.

In answer to this paper, DAMA released two responses [85, 86], criticising [84], suggesting that the claim that backgrounds are not properly accounted for or reported on is false. They cited papers and conference proceedings where further details about the direct ^{40}K decay were reported, and described the discussion of direct ^{40}K decay as “captious” since its contribution to the ^{40}K background is small (around 10%), and generally criticising assumptions made and approach taken in the paper. Further in response to this response, the authors of [84] published an addendum [87] to their original paper, rejecting DAMA’s

criticism and reaffirming their original conclusions, as well as demonstrating their numerical findings pertaining to modulation fractions still held in light of new values provided or referenced by DAMA in their response. Despite this back and forth, however, this theory did not find widespread support as a conclusive solution to the tension in the field caused by the DAMA results.

3.3.2 An annual modulation signal induced by DAMA’s analysis technique?

Recently a different explanation for the DAMA signal has been postulated whereby a slowly varying rate of background events can simulate an annual modulation signal in the case of a signal data-set calculated through the subtraction of yearly averaged background activities, as in DAMA’s analysis [88]. This procedure has the effect of generating a sawtooth-like event rate which can mimic an annual modulation. This effect was demonstrated in a bibliometric example where the rate of both citations and references in high energy physics papers since 1995 was given an apparent sinusoidal annual modulation after the above process of subtracting yearly averages from the data to generate a residual event rate was applied. The modulation is then shown to reduce when calculating the residual event rate by subtracting a continuous function over the whole data-set [88].

This analysis was extended to the DAMA residuals and again it was shown a sinusoidal annual modulation can be generated in simulated DAMA data if the above process is followed, even when no annual modulation is assumed when creating the simulated data as shown in Figure 3.11. Likewise, it is shown that a sawtooth modulation generated from the process of calculating residuals described above is “statistically comparable” to the sinusoidal modulation reported by DAMA [88]. However DAMA has not published the time dependence of their backgrounds so it is unclear exactly how well this analysis fits the true picture. In [88] several sources of potentially increasing backgrounds in DAMA are suggested, for example increases in the activity of certain backgrounds due to out-of-equilibrium radiogenic decays, or instrumental effects where malfunctions in detector equipment cause the measured event-rate to spuriously increase.

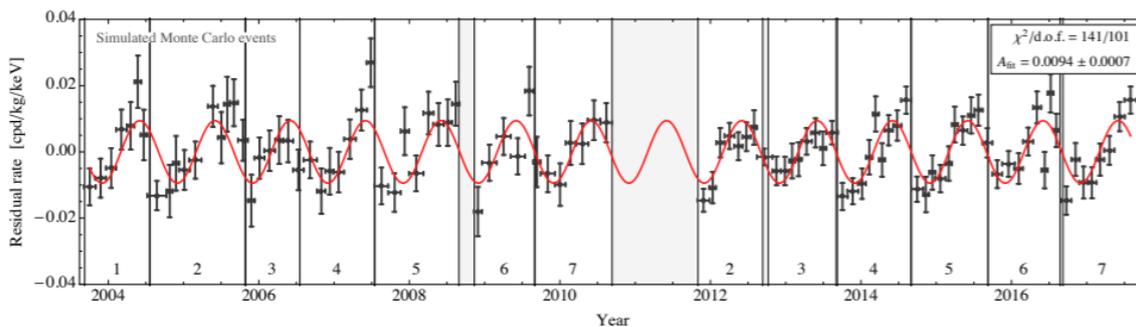


Figure 3.11: Residuals of simulated DAMA data (black points) from subtracting yearly averages of backgrounds. A sinusoidal function with DAMA-like characteristics can be fit to the data (red curve) which excludes the no modulation hypothesis up to 12.7σ , even though no modulation was included when creating the simulated data. From [88].

To test the hypothesis that DAMA’s analysis method might induce an annual modulation signal in an NaI-based dark matter detection experiment, the first three years of COSINE-100 data were re-analysed, aiming to follow the DAMA method as closely as

possible [5]. Since the DAMA experiments do not use any active shielding and use a different method of rejecting PMT noise to the machine learning techniques employed in the COSINE-100 experiment, the event selection procedures at COSINE-100 were updated for this analysis to follow the DAMA method. Two different methods were used in the calculation of the background event rates from which residual data-sets were calculated. In one, a continuous single exponential is used to describe the background event rate over the entire data-set, in line with the method used in the first COSINE-100 annual modulation analysis [75]. This method is used as a point of comparison for the second method which is to calculate the yearly averaged background rate, the same method as used by DAMA. The residual data-sets were then fit to a sinusoidal function with a fixed phase and period of the second of June and 152.5 days, as per the standard halo model. These data-sets, residuals, and fit results are shown in Fig. 3.12.

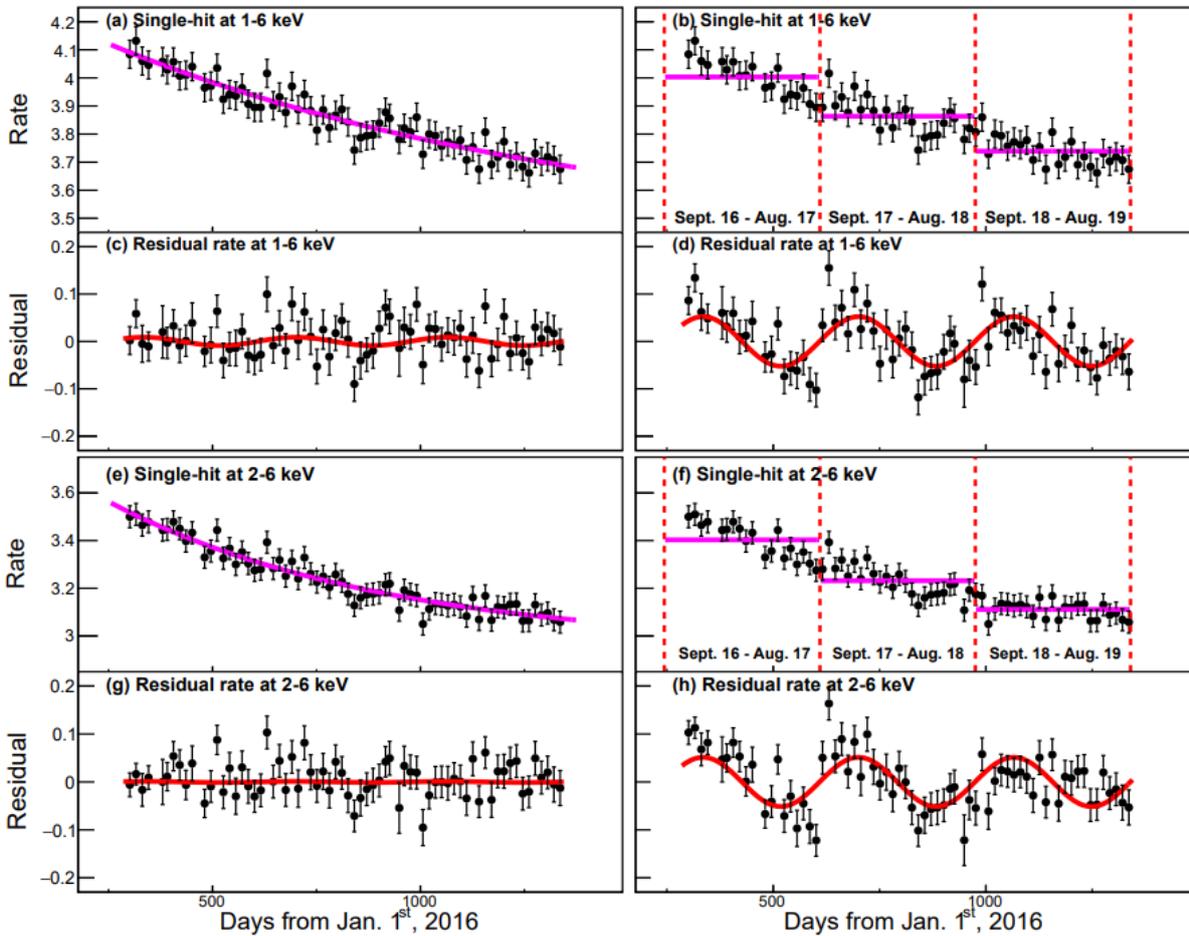


Figure 3.12: The calculation of residuals and subsequent annual modulation fitting of COSINE-100 data using two different methods for both 1-6 keV and 2-6 keV ROIs. In the left plots, the background event rate is modelled (purple curve) as a continuous single exponential, in the right plots three yearly averages are used to model the background event rate. These expected background event rates are subtracted from the single hit data (black data points), then fit with a simple sinusoidal function with a fixed phase and period in line with the standard halo model (red curve). From [5].

As can clearly be seen in Fig. 3.12, these two different methods create very different residual data-sets and this is reflected in the modulation fit. In the exponential background

case, the best-fit modulation amplitude is consistent with zero modulation, whereas in the yearly averaged background case a large, negative best-fit modulation amplitude of -0.044 ± 0.006 (-0.046 ± 0.006) counts/day/keV/kg for the 1-6 keV (2-6 keV) ROI is reported with a very high confidence level of around 12σ [5]. The negative sign on the best fit modulation amplitude implies a modulation with an opposite phase to that reported by DAMA. The relative negative on the induced modulation amplitude measured in the COSINE-100 data using DAMA's method means the choice of analysis method is not immediately obvious as the explanation for the DAMA signal, however, the fact this choice of analysis method does seem to induce an annual modulation in data where modulation is not otherwise observed may indicate an issue with the DAMA method.

3.3.3 DAMA more sensitive than other detectors?

In the absence of any widely accepted non-WIMP explanation of the DAMA signal, efforts have been made to reconcile the result with the null results from other experiments through providing theoretical reasons as to why DAMA might detect a WIMP signal while other experiments do not. These explanations rely on the fact that experiments using target materials different to DAMA - despite publishing the more stringent limits in WIMP searches - cannot provide a model independent test of DAMA due to their differing target material. This is true of any experiment that does not both use NaI as a target material and search for an annual modulation. By considering certain models of WIMP interactions with nuclei, beyond the most standard assumptions, it is possible to reconcile the DAMA signal with limits from other experiments in certain theoretical frameworks.

For example, it has been posited [64] that channelling in the NaI crystals, whereby fast moving ions created in a scattering process involving a WIMP and a crystal nucleus are preferentially guided along an axis in the crystal lattice, could increase DAMA's sensitivity such that it is the most sensitive dark matter detection experiment at low recoil energies. This is possible since in most dark matter detection experiments only part of the recoil energy of a nucleus is measured. In the case of NaI experiments only the electromagnetic fraction of recoils is measured and the full recoil energy is calculated from quenching factors which are assumed to be independent of energy by DAMA [64, 89]. However, it has been suggested [64] for recoils at low energies (i.e. the DAMA signal region) that due to channelling, nuclear recoils can travel along the crystal axis more freely than previously expected, thus causing the recoils to be measured at a higher energy than their true energy, hence causing DAMA to have a higher sensitivity in low energy regions than initially expected.

With certain choices of input values for the dynamic quenching factor, WIMP mass, etc, the tension with DAMA and other non NaI dark matter detection experiments can be removed, however this conclusion is sensitive to the input values used [64]. The effect of dynamic quenching factors on the DAMA signal compared to other results has been tested by the COSINE-100 experiment and it is reported that dynamic quenching factors are not sufficient to explain the DAMA signal in terms of spin independent WIMPs in light of results from COSINE-100 [89], [3], as shown in Fig. 3.10.

Further to channelling-based explanations, it has also been suggested [90] that the tension between DAMA and LXe experiments can be attributed to the different origins of nuclear spin in Xe, where the majority of a nucleus's spin originates from an unpaired neutron, compared to Na and I where the majority of nuclear spin is from an unpaired proton [90]. Since in spin-dependent interactions, coupling is suppressed for neutrons, if

this is the main channel by which WIMPs interact with normal matter then the bounds set by LXe experiments may be less stringent than originally reported and the DAMA result may not be in tension for some models of dark matter [90]. Once again however this conclusion is sensitive to the input parameters of any WIMP-nucleus interaction model. Furthermore, the above analysis can only reconcile the DAMA result prior to the lowering of the DAMA threshold to 1 keV in DAMA/Libra-phase2 [91], and if a non-Maxwellian velocity distribution of WIMPs is assumed [90].

With many of these suggestions there is no definitive conclusion as to whether or not the DAMA signal is adequately explained by the theory posited. Instead the theories presented in Sections 3.3.1, 3.3.2, as well as 3.3.3 have somewhat been put out to pasture, neither being conclusively confirming or refuting the DAMA signal. In this case it can be assumed that the scientific community has broadly decided the explanation is not sufficiently convincing to conclusively refute DAMA, even if a large deal of scepticism still exists around the result. There is therefore the need for new NaI experiments to provide a definitive model independent test of DAMA.

3.4 Concluding remarks

Over the last three decades, there has been considerable effort put into experiments using many different target materials and methods attempting to directly detect dark matter for the first time. All of these experiments have so far been unable to produce a signal attributed to dark matter, with the sole exception of the DAMA experiments which report a dark matter-induced annual modulation in NaI target material. This positive signal has yet to either be conclusively explained through non-dark matter-based sources or tested in a model-independent way with a different NaI-based experiment also searching for an annual modulation signal. There is therefore a requirement for an NaI-based experiment to search for an annual modulation signal and hence test the DAMA result. The COSINE-100 experiment is a leading example of such an experiment and the experiment itself, analysis techniques, and results of dark matter searches will now be described in the remainder of this work.

Chapter 4

The Design and Operation of the COSINE-100 Experiment

Based at the Yangyang underground laboratory near the Northeast coast of South Korea, COSINE-100 is a dark matter detection experiment that uses an array of high purity thallium doped sodium iodide (NaI(Tl)) crystals to search for an annually modulating dark matter signal caused by the modulating flux of dark matter particles expected on Earth. This expected annual modulation signal is the result of the Earth's orbit around the Sun and the Sun's orbit of the galactic centre, as described in Section 3.2.1. COSINE-100 was designed with the same target material and search method as the DAMA/NaI and DAMA/LIBRA experiments [49, 92] in order to test the longstanding and controversial DAMA signal in a model-independent manner.

The COSINE collaboration was formed as an amalgamation of the DM-Ice and KIMS (Korean Invisible Mass Search) collaborations, and the COSINE-100 experiment itself finished construction and started data taking in October 2016. The detector then operated almost continuously until March 2023, at which point the detector was taken offline to allow for upgrades to be carried out. The upgraded COSINE-100U experiment is planned to come online in 2024.

In this chapter, the design and operation of COSINE-100 detector will be described. First the detector site is described in Section 4.1, followed by a description of the detector design in Section 4.2 and operating methods in Section 4.3, followed finally by a comparison of the differences between the COSINE-100 and DAMA experiments in Section 4.4.

4.1 The Yangyang Underground Laboratory

As with DAMA and other high sensitivity dark matter detection efforts, it is necessary for the COSINE-100 experiment to be housed in an underground facility. This is to reduce the rate of potentially problematic background events from cosmic ray muons which are attenuated by the rock. Additionally, shielding from cosmic rays prevents the cosmogenic activation of radionuclides within the detector and surrounding materials which would increase the rate of background neutrons in the detector upon their decay; neutrons are a dangerous background in dark matter searches as they can mimic the signal from a WIMP [81]. This is because neutrons, as slow moving neutral hadrons, typically produce nuclear recoils which is the same signal expected from WIMPs.

The Yangyang Underground Laboratory (also known as Y2l) is situated beneath the

1424 m high Jeombongsan mountain in the Gangwon Province of South Korea, approximately 9 km Southwest of the coastal town of Yangyang and provides an excellent site for the COSINE-100 experiment. The location of the Y2l site in relation to the main offices of the Institute for Basic Science (IBS) in Daejeon, which can be thought of as the collaboration’s headquarters, as well as a schematic of the Y2l site, can be seen in Fig. 4.1. The site was originally excavated to house the 1000 megawatt Yangyang Pumped Storage Power Station, a pumped storage hydroelectric power station which uses water from the Namdae-Chun River to generate electricity, but now alongside the power station facilities, areas for underground science have been developed. The 700m granite overburden above the facility provides 1800 metre-water-equivalent of shielding [73].



Figure 4.1: A schematic of the Yangyang Pumped Storage Power Station (right) and its location in Korea relative to the IBS headquarters, located in Daejeon (left). The Underground Laboratory is located in the area marked “Power Plant” on the right image.

In addition to housing the COSINE-100 experiment, Y2l was the site of the original KIMS dark matter search experiment, KIMS-CsI [93], and the laboratory also now hosts the AMoRE experiment which uses molybdate crystal scintillators to search for neutrinoless double-beta decay [94]. After the KIMS-CsI experiment stopped taking physics data prior to the start of the COSINE-100 experiment, the space (known as the A6 tunnel) and equipment used for the KIMS-CsI experiment was turned into a facility for measuring the natural backgrounds at the Y2l facility.

The high purity germanium (HPGe) detector system from the KIMS-CsI experiment was repurposed to measure the rates of gamma rays emanating from the surrounding rock at the facility. Subsequently a new HPGe detector system was also added to study rare decays and further increase the sensitivity of measurements of background assays. Due to the low background rate in the A6 tunnel, the area is also a useful site to test any other components of COSINE-100 before installation, or any components of planned future detectors [95].

4.2 The COSINE-100 detector design

The COSINE-100 detector, shown in Fig.4.2, uses eight NaI(Tl) crystals with a total mass of 106 kg [73] (referred to as C1, C2, etc) to search for scintillation signals resulting from nuclear recoils of dark matter particles on the crystals. As can be seen in Fig. 4.2, the COSINE-100 detector has three main physics data taking components: the NaI

crystals (primarily used for dark matter searches), the liquid scintillator (predominantly for tagging and vetoing radiogenic decays), and plastic scintillator panels (primarily for muon tagging). Surrounding these components are several layers of passive shielding for further background rejection. In this section the motivation for the design of the detector will briefly be discussed in Section 4.2.1 before the detector components are described in more detail, starting with the NaI crystals in Section 4.2.2, followed by the muon veto in Section 4.2.3, the liquid scintillator in Section 4.2.4, and the data acquisition system in Section 4.2.5.

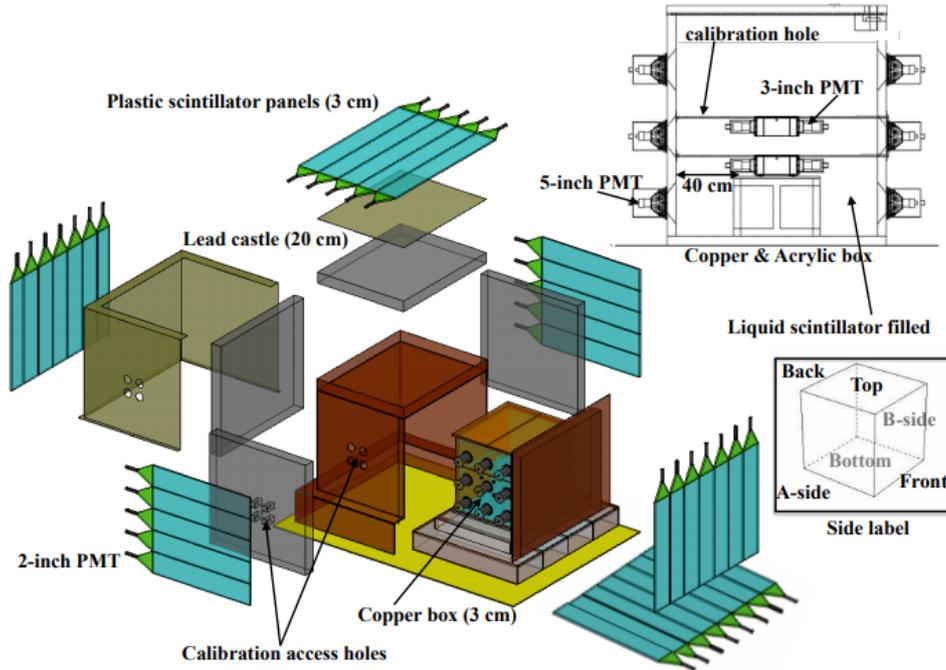


Figure 4.2: Diagram of the COSINE-100 detector. Show from outside to in are: plastic scintillators for muon tagging, a 20 cm thick lead castle, a 3 cm thick copper box, and 1 cm thick acrylic box. Immersed in 2200 L of Linear Alkyl-Benzene liquid scintillator contained by the acrylic box are eight highly radio-pure NaI(Tl) crystals with a combined mass of 106 kg which are the main target material for the detector. The detector is further shielded for cosmic rays by being housed underground at the Yangyang Underground Observatory. From [73].

4.2.1 Motivating design principles of the detector

As described in Section 3.2.1, if the Milky Way has a spherically symmetric halo of WIMPs (or other dark matter particles) that is not co-rotating with the baryonic material in the galaxy, a given star orbiting the galactic centre will experience a constant flux of WIMPs. If the orbit of a planet orbiting said star is not perfectly perpendicular to the star's path through the WIMP halo, the WIMP flux as experienced on the planet will vary in a sinusoidal manner throughout the year with the general form

$$R(t) = R_{av} + R_{max} \cos\left(\frac{2\pi(t - t_0)}{T}\right), \quad (4.1)$$

where $R(t)$ is the flux at a given time in the year, t , R_{av} is the yearly averaged flux experienced by the planet, R_{max} is the maximum flux above R_{av} , i.e. when the planet's

velocity vector is most in the direction of its host star’s around the galactic centre which happens at t_0 , the phase, and T is the period or the length of one year on the planet. As such, it can be expected that the rate of scintillation events from WIMPs as measured on Earth would annually modulate, peaking in early June, this is known as the “WIMP wind”. This is the signal reported from the DAMA/NaI and DAMA/LIBRA experiments and hence the signal searched for at COSINE-100.

In order to detect a faint signal from rare events, such as the annual modulation expected from dark matter, the background event-rate must be minimised as much as possible. This is primarily achieved in two ways, purifying detector materials to reduce internal backgrounds from radioactive decay, and shielding the detector to reduce the rate of external background events reaching the crystals and being recorded. Despite these efforts, some background events will inevitably be measured in the detector and in this case it is important to tag these events in order to separate them from potential signal events as much as possible. The rareness of an expected dark matter signature also naturally lends itself to a powerful discrimination tool to remove noise events in that since dark matter events are expected to be very rare and only produce one measurable signal per event, any potential signal-like events occurring in short time coincidence with any other event in the detector are unlikely to be from dark matter, assuming a sufficiently low background event-rate. This leads to the concept of “single-hit” and “multi-hit” events. Single-hit events being potential dark matter events, and multi-hit events being sideband events or noise.

4.2.2 The NaI crystal array and inner passive shielding

As shown in the schematic of the crystal layout shown in Fig. 4.3, the crystals are arranged in a 4×2 array atop an acrylic table. Each crystal is wrapped in PTFE reflective sheets and hermetically sealed in a nitrogen gas environment in a fibre-optic tube with 12.0 mm thick quartz windows optically coupled to the crystal at each end with 1.5 mm thick optical pads. The quartz windows are then light-coupled via high viscosity optical gel to 3-inch Hamamatsu R12669SEL photomultiplier tubes (PMTs) with quantum efficiencies of around 35% [73]. The crystals are also each encapsulated in a high-purity copper cylinder.

The crystals at COSINE-100 were grown using powders made by Alpha Spectra Inc. (AS). This is a different manufacturer to that used by DAMA as that company no longer produces high-purity NaI(Tl) crystals. Four different powders were used with some powders being used to grow multiple, cylindrical crystals. From measurements of the crystals after purchase by the COSINE-100 collaboration, the contamination of each crystal by various background materials was calculated. This was reported along with crystal mass, size, and light yield before WIMP search results were released in [73]. These results are shown in Table 4.1.

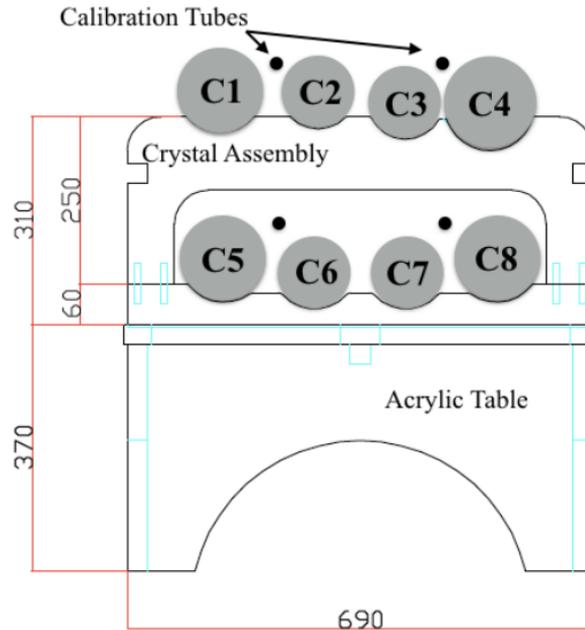


Figure 4.3: A schematic showing the layout of the eight COSINE-100 crystals on the acrylic table. Four calibration holes allow calibration sources to easily be inserted into the detector. From [73].

Crystal	Mass (kg)	Size (inches) (diameter×length)	Powder	α Rate (mBq/kg)	^{40}K (ppb)	^{238}U (ppt)	^{232}Th (ppt)	Light Yield (PEs/keV)
C1	8.3	5.0×7.0	AS-B	3.20±0.08	34.7±4.7	<0.02	1.3±0.4	14.9±1.5
C2	9.2	4.2×11.0	AS-C	2.06±0.06	60.6±4.7	<0.12	<0.6	14.6±1.5
C3	9.2	4.2×11.0	AS-WSII	0.76±0.02	34.3±3.1	<0.04	0.4±0.2	15.5±1.6
C4	18.0	5.0×15.3	AS-WSII	0.74±0.02	33.3±3.5		<0.3	14.9±1.5
C5	18.3	5.0×15.5	AS-C	2.06±0.05	82.3±5.5		2.4±0.3	7.3±0.7
C6	12.5	4.8×11.8	AS-WIII	1.52±0.04	16.8±2.5	<0.02	0.6±0.2	14.6±1.5
C7	12.5	4.8×11.8	AS-WIII	1.54±0.04	18.7±2.8		<0.6	14.0±1.4
C8	18.3	5.0×15.5	AS-C	2.05±0.05	54.3±3.8		<1.4	3.5±0.5

Table 4.1: Properties of the COSINE-100 NaI(Tl) crystals, from [73]

As reported in Table 4.1, C5 and C8 have significantly lower light yields than the other crystals; this is suspected to be due to a degradation of their optical couplings soon after installation. This ultimately results in C5 and C8 not having the requisite performance levels to be used in the single-hit modulation searches. Additionally, a high rate of PMT noise in C1’s PMTs which cannot be sufficiently removed by the noise rejection process excludes C1 from single-hit modulation searches. Consequently, the effective mass of NaI searching for a DAMA-like modulation signal at COSINE-100 is 61.4 kg. It should be noted that coincident signals in C1, C5, and C8 with another crystal can still be used to veto multi-hit events.

As can be seen in Fig. 4.2, the crystal array is housed inside a 1 cm thick acrylic box, the inner surface of which is coated in reflective foil to improve light collection. The acrylic box also serves to contain 2,200 L of Linear Alkyl-Benzene (LAB)-based liquid scintillator for active rejection of backgrounds from radiogenic decays internal to the

detector. There is at least 40 cm between the crystal PMTs and the inner wall of the acrylic box. A 9 cm gap is left between the level of the LS and the top of the box as a safety measure in the case of the LS increasing in volume due to thermal expansion from an unexpected rise in temperature, however thanks to the high heat capacity of the liquid, a constant temperature of around 24.2°C can be reliably maintained. Nitrogen gas is continually pumped into this gap to protect the liquid from airborne contaminants. The LS is described in more detail in Section 4.2.4.

Further shielding the crystals and LS is a 3 cm thick copper box with dimensions of 152 cm by 142 cm by 142 cm, wall thickness 3 cm, and mass 6.4 tonnes. Surrounding the copper box is a 20 cm thick lead castle made up of two layers of lead bricks, each with dimensions of 20 cm by 10 cm by 5 cm; the outer layer is made of normal lead with 99.99% purity, and the inner layer of low-contamination lead. Finally, the outside of the detector is clad in plastic scintillator panels designed to tag muons. Since muons can produce neutrons from spallation events that can mimic a dark matter signal [81], it is crucial to be able to distinguish muon-induced events from signal events.

The crystals are accessible through four calibration holes in the shielding into which calibration sources can be placed, thus allowing for additional calibration to be carried out after installation. The concentrations of ^{238}U and ^{232}Th in various parts of the COSINE-100 shielding as reported in [73] are shown in Table. 4.2.

Detector Component	^{238}U (ppt)	^{232}Th (ppt)
High purity lead castle layer	6.9	3.8
Copper box	27	51
Liquid scintillator	7	4

Table 4.2: Radiogenic backgrounds from ^{238}U and ^{232}Th in different parts of the COSINE-100 shielding, from [73]

4.2.3 Muon shielding

The muon veto at COSINE-100 consists of 37 3-cm thick Eljen EJ-200 plastic scintillator panels (the outer blue panels in Fig. 4.2) which provide 4π solid angle coverage for muon tagging. Events elsewhere in the detector occurring in coincidence with muon events can be attributed to the muon, rather than some other source. The panels are either 40 cm or 33 cm in width, depending on their location, in order to properly fit within the shielding structure. Panels on the top of the detector are 282 cm long and are outfitted with a two-inch H7195 Hamamatsu PMT at each end, whereas panels on the side of the detector are around 200 cm long and are outfitted with a single PMT of the same type as the top panels. The orientation of the plastic scintillator panels is shown in Fig. 4.2 and a schematic showing one of the side panels can be seen in Fig. 4.4.

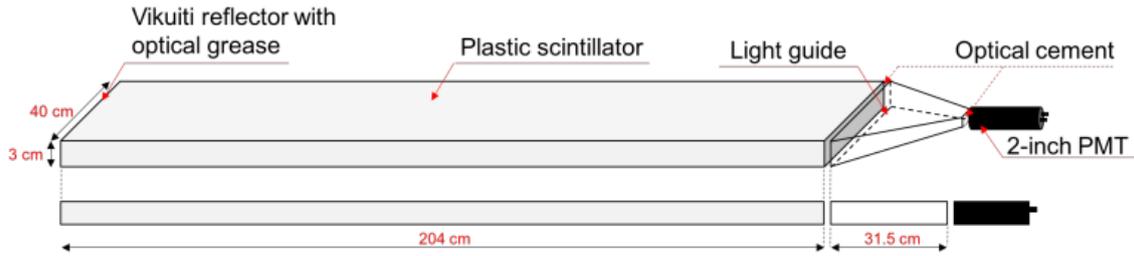


Figure 4.4: A schematic of one of the side plastic scintillator panels for muon tagging. From [96].

For a muon to be tagged at least two panels must have a signal above a threshold of 4,000 analogue-to-digital conversion (ADC) counts within a 400 ns time window. This threshold is at least three times less than the integrated charge of most muon events. When this criterion is met, the readouts from all the muon panels are recorded in a 4 μ s window centred around the event. The minimum energy for a muon to be detected is around 5 MeV and the measured muon flux at the COSINE-100 experiment is 328 ± 11 muons/m²/day; this is around 2.8×10^{-5} of the rate measured on the surface above the rock overburden. More information about the muon veto system at COSINE-100 can be found in [96].

4.2.4 Liquid scintillator

Since it is not possible to totally remove radiogenic contaminants in detector materials, nor total insulate the detector from external radiation, passive shielding alone cannot prevent background events from being measured in the COSINE-100 crystals. Indeed, such contaminants are the main source of background events in the COSINE-100 experiment [4]. As such, when searching for signals with low expected event rates such as WIMPs, it is desirable to employ active shielding where background events are tagged, along with passive shielding, to more easily identify background events and exclude them from physics searches. This makes use of the fact that since dark matter events measured in the detector are likely to be rare, crystal events in coincidence with events in active shielding are unlikely to be dark matter events. An additional benefit of actively measuring backgrounds in this way is it easily allows many more sideband studies than otherwise would have been possible with main target material alone, thus allowing for a more robust test of dark matter searches carried out using the main target material. One such example of active shielding at COSINE-100 is the liquid scintillator veto which surrounds the crystal encapsulation and is described in this section.

Design of the liquid scintillator

Surrounding the crystal encapsulation at COSINE-100 is 2200 litres of Linear Alkyl-Benzene (LAB) liquid scintillator (LS) that fills the acrylic box housing the crystals. The basic principle of detecting signals in a liquid scintillator is as follows. A particle enters the liquid and deposits some energy through the excitation solvent molecules in the liquid. This excitation energy is then transferred to fluor molecules which themselves become excited, then emit photons which are measured, in this case by PMTs.

The design and location of the copper and acrylic box holding the LS liquid, and the arrangement of the LS PMTs are shown in Fig. 4.2. Mixed into the LAB (the solvent) is 3

g/L of 2,5-diphenyloxazole and 30 mg/L of 1,4-bis (2-methylstyryl) benzene which, along with being the fluor molecules, act as wavelength shifters to match the wavelength of scintillation light to the wavelength that optimises PMT efficiency [4] (around 340 nm), as well as improving the optical transparency of the LS. The LS is filtered through 1 μm pore size Meissner filters and purified using a water extraction method to minimise contamination from radio-isotopes. Despite this, a non-trivial amount of radio-isotopes remain in the LS as reported in Table. 4.2.

To estimate the level of background contamination in the LS, a prototype detector with the same solvent material as in the main COSINE-100 detector was constructed and installed at the Yangyang Underground Laboratory. This detector used pulse shape discrimination to distinguish alpha and gamma events, and by utilising the time coincidence of events in their decay chains, events from ^{238}U and ^{232}Th could be tagged to estimate the upper limit of their chain activities. These were found to be 0.087 mBq/kg in the case of ^{238}U and 0.016 mBq/kg in the case of ^{232}Th . More information on the event selection of these events can be found in Section 4.3.3.

The walls of the acrylic box and outer surface of the crystal assembly are coated in Vikuiti-ESR specular reflective foil to increase light collection efficiency and thus better allow signals in the LS to be measured. To prevent oxygen, water, or any other contaminant from reaching and chemically interacting with the LS, a 9 cm gap is left at the top of the acrylic box, into which nitrogen gas is pumped. This gap also acts as a safety margin in the case of thermal expansion of the LS. However, the high thermal capacity of the LS helps keeps the temperature at $24.2 \pm 0.5^\circ\text{C}$.

Two 9×9 arrays of 5-inch R877 Hamamatsu PMTs are attached and optically coupled to two parallel sides of the acrylic box to measure the light from signals generated in the LS. The LS may be operated in “passive” or “active” mode. In the former, in order for signals measured in the LS to be recorded, the LS signal must fall within a 200 ns global trigger generated from an event in the crystal PMTs. In this mode LS signals are recorded every time there is a signal in the crystal. The trigger conditions in COSINE-100 are described in more detail in Section 4.2.5. In active mode, LS signal are recorded irrespective of a trigger from the crystal PMT, however this mode also requires threshold and multiplicity conditions to be met in order for a signal to be recorded. As with the crystal triggering and DAQ, more information can be found in [97]. The LS was mostly operated in passive mode and the reader should assume this to be the case in all future chapters unless specified otherwise.

Liquid scintillator calibration, threshold, and gain correction

The LS PMTs were calibrated by first adjusting the gain of each PMT such that saturation occurred at around 40,000 ADC units; the 18 separate PMT spectra were then summed to produce an uncalibrated spectrum for the whole LS. Events that saturate more than one PMT are then removed from the spectrum. In the resultant spectrum, shown in Fig. 4.5, Compton shoulders for ^{40}K and ^{208}Tl at 1.46 MeV and 2.6 MeV are visible. However, as a result of photons scattering and losing energy in the LS, there is not enough finer structure from mono-energetic features in the spectrum to achieve sufficiently accurate energy calibration from spectral features alone. Therefore to achieve a final energy calibration, the spectrum is compared to a Geant4-generated simulated spectrum considering all the known background components. Major features in the two spectra, such as the 1460 keV Compton edge from ^{40}K are compared, thus allowing the resolution to be estimated and found to be approximately 10% [73].

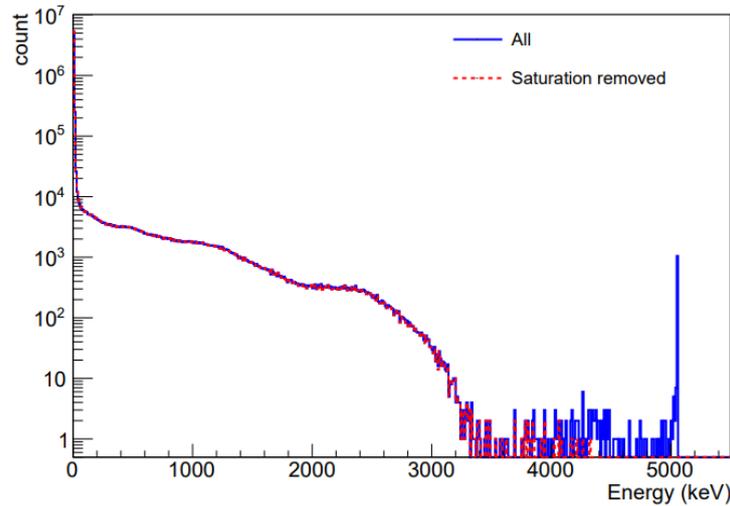


Figure 4.5: Summed spectra of all 18 LS PMTs for all events (blue solid line) and with saturated events removed (red dotted line). The Compton shoulders of ^{40}K and ^{208}Tl can be seen at approximately 1400 keV and 2600 keV. From [4]

The LS low-energy threshold was initially chosen with the goal of keeping dead time from LS veto events below 0.5% of the total live time, while sufficiently vetoing genuine multi-hit events. As can be seen in Fig. 4.6, tests of different LS thresholds showed a threshold of around 20 keV is sufficient to reach the $< 0.5\%$ dead time goal [4].

While 20 keV was initially chosen as an ideal low-energy threshold and was used up until after the 1.7 years annual modulation search, a decision was taken to raise the threshold to 80 keV for the annual modulation search with three years of data due to concerns about the stability of the LS during the analysis. These concerns were ultimately allayed, but the threshold remained at 80 keV for the subsequent modulation analysis due to uncertainty in the ADC to keV conversion after it became necessary to remove one of the LS PMTs. For more information on this process see Section 5.2.2.

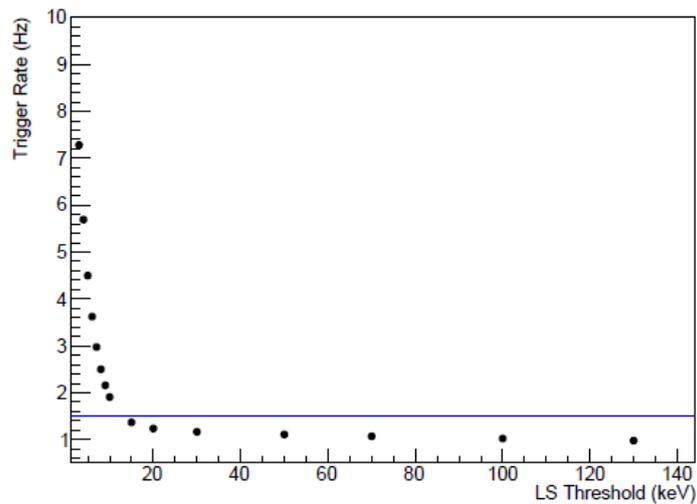


Figure 4.6: The measured LS trigger rate as a function of the low-energy threshold chosen. The blue corresponds to 1.5 Hz, the target trigger rate for 0.5% dead time in the LS. From [4]

During data taking, issues with one of the 18 LS PMTs became apparent and it

was necessary to remove its signals from the analysis. This slightly changed the shape of the measured energy spectrum and as such necessitated updating the LS calibration, threshold, and gain correction. This process is described in more detail in Section 5.2.2.

Overall the background tagging efficiency of LS, defined as the fraction of NaI events that are vetoed as a result of LS coincidence, varies between 14% and 29% between 2-6 keV depending on the crystal (as determined by the background composition of a crystal) [4]. The background tagging efficiencies are slightly lower between 6-20 keV across all crystals, suggesting the tagging efficiency of the 3.2 keV peak in the crystal and 1460 keV Compton edge in the LS from ^{40}K is higher than other components. The ^{40}K tagging efficiency is estimated to be between 65% and 75% for a given crystal [4]. This value is obtained from a Geant4 simulation. Untagged ^{40}K events are a result of either the 1460 keV γ not interacting in the LS and not being tagged, or the ^{40}K nucleus decaying into ^{40}Ca instead of ^{40}Ar and consequently not producing a 1460 keV γ at all. The background tagging efficiency above 100 keV is significantly higher than that at lower energies, mostly over 50%, as a result of there being a relatively higher concentration from radioisotopes that produce events at higher energies, e.g. ^{238}U and ^{232}Th [4].

The stability and gain of the LS is tracked using the 1460 keV emission from ^{40}K . However, these 1460 keV signals must pass through the bulk of the LS before being recorded by a PMT and so are Compton scattered significantly more than signals measured in the crystals. Consequently a Compton edge is fit rather than a Gaussian peak. The phenomenological function used to characterise this feature is

$$R(x) = \frac{p_0}{e^{(p_1(x-p_2))} + 1}, \quad (4.2)$$

where $R(x)$ is the count rate at a given ADC value, x , and p_0 , p_1 , and p_2 are fitting parameters controlling the height, width, and location of the Compton edge respectively. The location of p_2 is tracked over time using spectra from 100 hours of data (where the spectra are generated from summing the outputs from all LS PMTs) to measure the stability of the LS. Shifts in LS gain tend to be more long-term and gradual than in the crystals so gain correction is carried out using a linear time-dependent correction factor to restore the Compton edge to the reference value. The process of gain correction in the LS is described in more detail in Section 5.1.

4.2.5 Data acquisition system

As described above, the COSINE-100 experiment records and stores data from multiple detectors: the NaI crystals as well as two types of active shielding in the LS and muon vetoes. In total there are 76 different PMTs collecting data at COSINE-100 with a total of 92 channels so having efficient and effective data acquisition system (DAQ) and trigger algorithms is crucial. The basic flow of data from detector component to DAQ computer is shown in Fig.4.7. The features of this diagram will now be explained in more detail. Additional information can be found in [97].

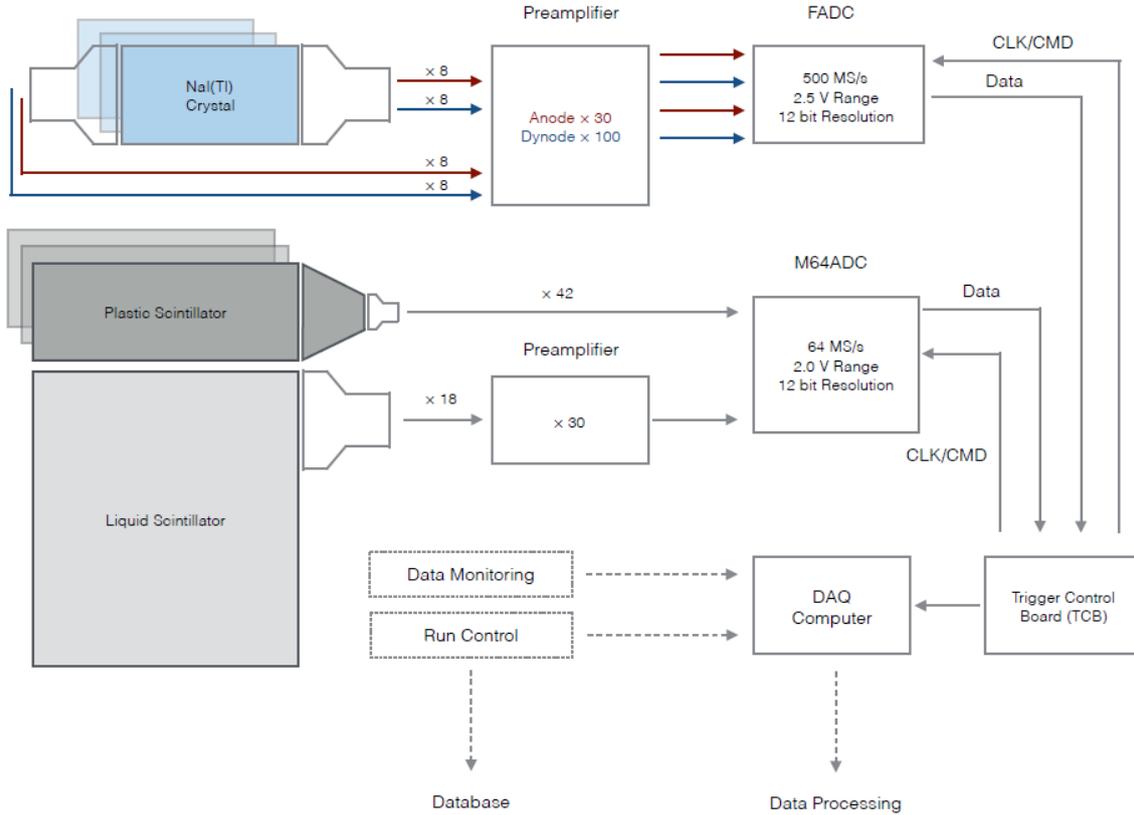


Figure 4.7: A diagram showing the flow of data at COSINE-100. Both the anode (red) and dynode (blue) signals from the PMTs of the NaI crystals are processed through a preamplifier, flash analogue-to-digital converter (FADC) module, and a trigger control and clock board (TCB) before reaching the DAQ computer for data processing. Signals from the PMTs of the LS pass through a separate preamplifier and then go into a charge-sensitive flash analogue-to-digital converter (M64ADC) module whereas the plastic scintillator PMT signals go directly to the M64ADC module. These signals then pass through the TCB before reaching the DAQ computer. Image from [97].

Overall there are eight flash analogue-to-digital converter (FADC) modules which are used to digitise signals from the NaI crystal PMTs, two charge-sensitive flash analogue-to-digital converter (M64ADC) modules which digitise the signals from the plastic scintillator panel and LS PMTs, a trigger control and clock board (TCB) for global decision making and synchronising the module clocks, four preamplifiers, and a high voltage supply. The crystal PMTs produce two separate signals: high-energy signals from the anode, and low-energy signals from the fifth stage dynode (the cutoff between the two being at around 100 keV). The anode signal is amplified by a factor of 30, and the dynode signal by a factor of 100. The LS signals are also amplified by a factor of 30, but through a different preamplifier to the NaI PMT anode signals. The signals from the plastic scintillator PMTs are sent straight to the M64ADC modules. In the FADC, anode signals show linear behaviour up to around 100 keV and the dynode signals have a linear response up to around 3 MeV [97].

Raw data sent to the DAQ computer is stored in ROOT files. Each one of these ROOT files corresponds to two hours of data known as a “subrun”. Subruns are organised into runs with a new run beginning either after 2000 hours (1000 subruns) has passed, or the DAQ has to be shut down for any reason (e.g. as a result of a power outage, being overloaded by a high-energy muon event, changing a detector setting, etc).

Trigger conditions

Signals are recorded from the FADC modules when two anode signals greater than a threshold of 6 mV (around 10 ADC counts or 0.2 photo-electrons) occur within a 200 ns time window. When such coincident events occur, a trigger output signal is sent to the TCB, triggering 8 μ s waveforms of each the two crystal PMTs to be recorded, starting 2.4 μ s before the trigger. The TCB having received the trigger output signal from the anode generates a global trigger and the dynode signals are recorded [97].

Instead of saving raw waveform data, the M64ADC modules digitise the analogue signal input every 16 ns and when certain trigger conditions are met in a preset field programmable array, the following 192 ns is digitised and the charge sum in ADC counts is returned. A 192 ns integration time is chosen as it is sufficiently longer than the decay time of the plastic scintillator and LS signals. There are two different trigger modes for the M64 module at COSINE-100. In the first known as “active mode”, the M64ADC modules send a trigger signal to the TCB if a second channel is triggered in a 400 ns coincident window after the charge of a single channel exceeds a preset value of 4000 ADC counts. The TCB then sends a signal to the FADC and M64ADC modules to record a signal. In the M64ADC, this opens a 4 μ s gate window, with the end of the 400 ns coincident window in the middle of the 4 μ s gate window. Within this gate window, in each M64ADC channel the maximum integrated charge is saved along with the corresponding time information.

The second global trigger mode differs from active mode in that it requires a signal from the FADC modules from the crystals in order for a M64ADC signal to be recorded. This mode is called “passive mode”. In this mode, two signals in a single crystal’s PMTs must occur within 200 ns for a trigger to be created and this trigger causes 8 μ s of data from both the M64ADC and FADC modules, centred around the initial event, to be recorded. Unlike active mode, where self triggered events in the plastic scintillator and LS PMTs can be recorded, passive mode inherently requires some signal from crystal PMTs for plastic scintillator and LS signals to be recorded. For the purposes of this work, it can be assumed the detector always operates in passive mode unless otherwise stated.

Muon events in the crystals sent to the FADC modules can cause trigger rates over 1 kHz which can exceed the capacity of the DAQ system. To prevent errant signals in the immediate aftermath of muon events from being included in physics searches, a 1 ms hardware dead time is applied to an FADC channel that generates a trigger of any type. This causes a total dead time of around 1.5%, as the overall DAQ rate of the FADCs is around 15 Hz. This amount of dead time is sufficiently small to not have a significant effect on dark matter searches.

4.3 Detector operation

Since the signal being searched for at COSINE-100 is a very small annual change in rate, it is crucial that the detector is as stable as possible over a time scale of many years. This is a particularly pertinent concern in an experiment looking for an annual modulation peaking in the (Northern Hemisphere) Summer, as many natural processes follow the same general trend and so if not controlled and accounted for could lead to spurious results. Likewise, other long-term changes can also induce or mask real physical signals. Indeed, many of the non-dark-matter explanations for the DAMA result described in Section 3.3 refer to incorrectly accounted for signals generated by long term shifts in detector behaviour. To that end, great attention is taken to ensure that the detector is as stable as possible and that any changes that do occur are properly measured and accounted for.

To achieve this - beyond efforts to design and build the detector to be as stable and as isolated of outside conditions as possible - great care is paid to monitoring as many pertinent variables within the detector that may be subject to change.

4.3.1 Monitoring of environmental variables

Many parameters inside the detector are continuously monitored, then displayed for convenient viewing in real time by collaboration members on the Grafana web application which allows users to view environmental data from second to multiple year-long scales. This allows for detector behaviour to be easily compared with past behaviour and any irregularities to be reported. Additionally, thresholds can be set on measured parameters such that when a certain value is exceeded, the time is recorded and users are notified. More details about the COSINE-100 environmental monitoring system can be found in [98].

One of the most important variables monitored is the temperature of the detector. Temperature monitoring is of particular importance as the scintillation response of NaI crystals is dependent on temperature [99], so the temperature of crystals must be closely monitored and kept constant. A thermometer is placed directly on the crystal encapsulation surface to monitor the temperature of the crystals and the temperature of the detector room is automatically maintained by an air conditioner. The crystal temperature is generally stable to around $\pm 0.1^\circ\text{C}$, as shown in Fig. 4.8 which shows the measured crystal temperature for a 12 month period between November 2021 and 2022.

Thermometers are also placed in various parts of the detector shielding, around the LS, around the detector room, and outside the detector room in the A6 tunnel. The tunnel is not required to be carefully temperature controlled like the detector, so an annual modulation in temperature from around $21 - 27^\circ\text{C}$ is seen as expected based on the season.

Alongside temperature, the following parameters are also measured and made available for viewing to all collaboration members on the Grafana app:

- The humidity inside the detector shielding, in the detector room, and inside the air conditioner. The humidity within the detector shielding is kept below 5%.
- The current and voltage variations as measured across a number of detector components as well as the input voltage, output voltage, and battery status of the uninterruptible power supply. Any fluctuations above either 1 V or 1% are reported.
- The levels of oxygen and radon, and the number of airborne particulates, as well as the flow-rate of nitrogen over the LS which must be maintained at over 2 litres/minute. The temperature and humidity of the RAD7 is also measured.
- The overall event rate in the FADC and M64ADC modules. Spikes over 100 Hz are reported.

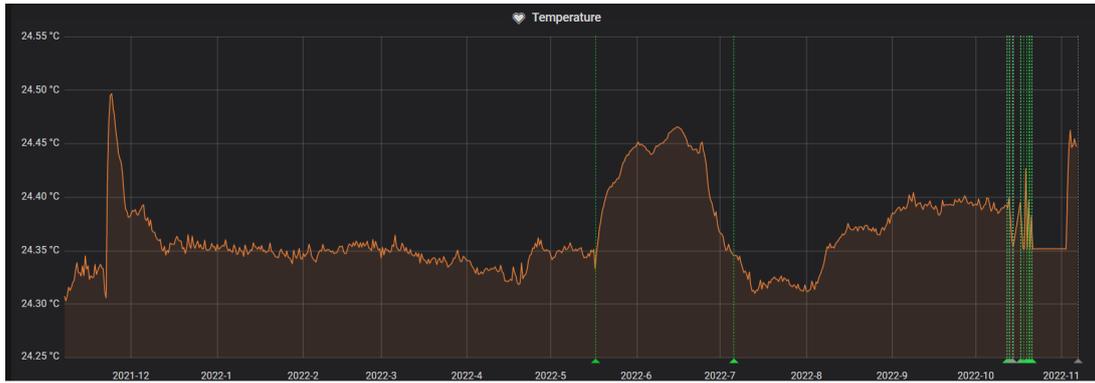


Figure 4.8: The temperature of a thermometer directly attached to the crystal encapsulation surface over the course of 12 months from November 2021 to 2022 as seen in the Grafana monitoring application. The green dashed lines represent instances where some other measured parameter has exceeded a predefined level.

4.3.2 Monitoring shift work

Members of the COSINE-100 collaboration are required to carry out a week-long monitoring shift roughly once every year and report the results internally. These shifts have two major components, firstly the monitoring of the above environmental variables and reporting of any irregularities, and secondly monitoring of plots of various physics variables produced every subrun, as well as plots of daily and weekly aggregated variables. In all cases this is to check for deviations from standard behaviour. These plots are compared to the behaviour of a previous 200 hour period of good data for ease of comparison. The most common cause of a subrun being marked as bad is a muon event in either the FADC or M64ADC modules causing the DAQ to be temporarily overloaded. An example of this is shown in comparison with a good subrun in Fig. 4.9. Typically a threshold of 100 Hz is set in either the FADC or M64ADC modules to mark a subrun as bad, although a smaller spike in rate from muon events can also cause a subrun to be marked as bad in the case that the muon causes a very high event-rate in a single crystal.

Many other parameters are also plotted for each subrun and checked by the person on shift duty. For these plots, deviations from the reference plot are reported and decisions about whether to mark a subrun good or bad are made on a case by case basis. This process can be particularly useful for observing times when additional noise is generated in the detector. Instances where some part of the detector has unexpectedly changed over time should also be detected during this process, particularly from the daily and weekly plots. Variables monitored include the crystal trigger rate (both with various types of event selection applied and no event selection), the pedestal value in each crystal, the energy spectrum in the LS and muon vetoes at both overall and PMT-levels.

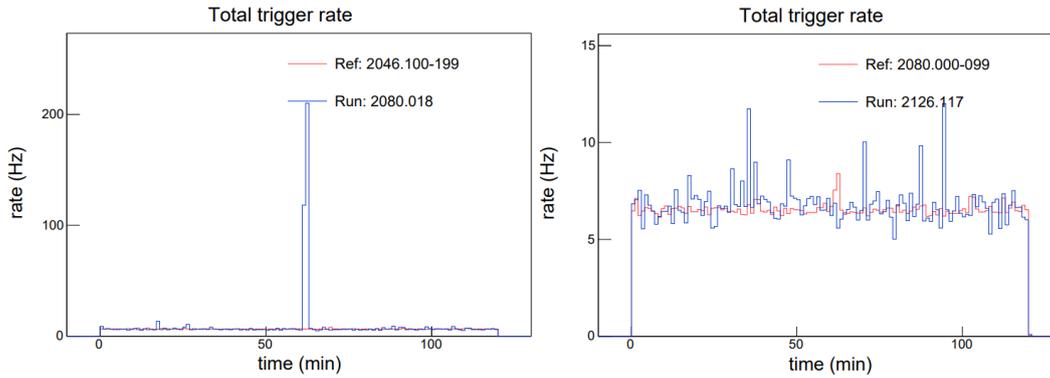


Figure 4.9: The total event rate in the FADC modules (blue line) for a two hour subrun compared to a reference 200 hours of good data (red line) of a subrun marked as bad (left) due to a muon event causing a spike in rate, and a normal subrun (right) as seen in the online monitoring plots.

4.3.3 Detector component monitoring and gain correction

The above two subsections describe efforts to ensure the COSINE-100 detector’s stability primarily over short time frames. While collaboration members on a monitoring shift should as much as possible try to observe long-term trends, given that each shift is only a week long, it can be difficult to observe such long-term behaviour. Therefore in order to ensure the stability of the detector components over longer time-frames, each section of the detector is assigned a collaboration member responsible for long-term monitoring and gain correction (if required). These efforts are described briefly below for the NaI crystals and plastic scintillator panels, and for the monitoring of neutron events measured in the detector. A brief description of this process in the LS can be found in Section 4.2.4, and a more detailed description in Section 5.1.

NaI crystals

The behaviour of the NaI crystals is monitored through the location of the 50 keV peak from ^{210}Pb decay in the anode and 1460 keV peak from ^{40}K decay in the dynode. The 50 keV ^{210}Pb peak in the anode spectrum is fit to a Gaussian plus a second order polynomial function. This is done continuously in 10 subrun-long batches where each spectrum is the sum of the two PMT outputs for a given crystal. This fit along with the fit to the dynode is shown in Fig. 4.10. In regions where the location of the peak, measured as the mean of the Gaussian, moves away from the reference value (around 235×10^3 ADC counts) for a prolonged period, a flat correction factor is applied to the anode energy spectrum in that region to shift the peak back to the correct position.

The 1460 keV ^{40}K peak in the dynode is in a more complex region of the spectrum than the anode peak and as such requires a more involved fitting procedure. Specifically, in four of the five crystals used for single-hit searches, two additional Gaussians either side of the main peak are also fit, along with a linear function and the main peak which is also fit to a Gaussian, the mean of which characterised the dynode behaviour. In the case of C2, a fourth Gaussian overlaps the central ^{40}K peak. This is also fit along with the four other functions described for the other crystals and the mean of the means of these two overlapping Gaussians are used for tracking. In the dynode, fits are applied to spectra consisting of 100 subruns (again from the sum of the two crystal PMTs) and

corrections to the spectra are applied in the same way as in the anode in the case of consistent deviation from the reference value.

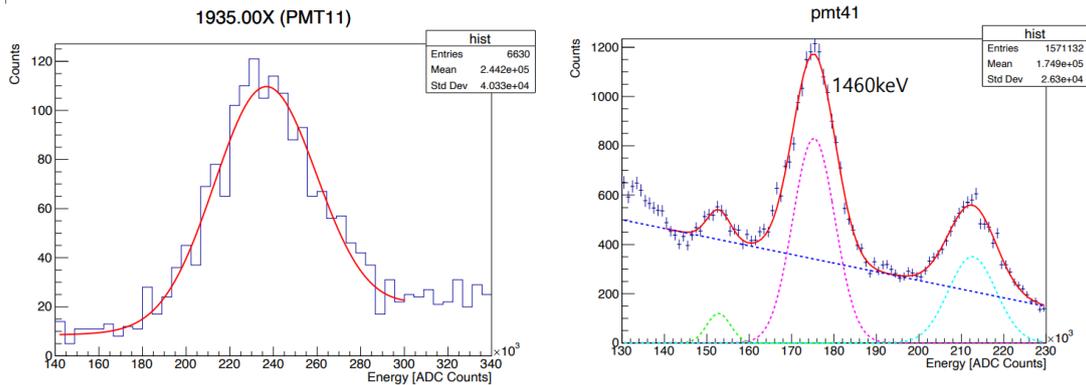


Figure 4.10: The fits for gain tracking and correction to the 50 keV peak from ^{210}Pb in the anode (left), and the 1460 keV peak from ^{40}K in the dynode (right). The anode peak is fit to a Gaussian plus a second order polynomial (shown combined in the red line), in this case the dynode peak is fit with three Gaussians (green, pink, and cyan dashed lines) plus a linear function (blue dashed line). The red solid line is the overall fit. In C2 a fourth Gaussian is also required. Plots taken from internal presentation by SH Lee.

Plastic scintillator panels

The plastic scintillator panels are monitored and corrected individually. The overall spectra from muon events of each panel is fit to either a Gaussian or Landau distribution depending on the spectral shape. The location of the means of the Gaussian distributions and the most probable value of the Landau distributions are tracked. Changes in plastic scintillator gain are observed to be linear in time at COSINE-100 and so a linear time-dependent correction factor is used to maintain the constant performance of the muon veto system.

There is also some noise in the plastic scintillator panels which is more prevalent at lower energies and decreases exponentially as energy increases. At the end of each run, the spectra of each muon event in the run are combined and fit to two Landau functions to fit muon events, and an exponential to fit the backgrounds. The tagging efficiency is then calculated as the difference between the integrals of the muon functions and the background function over the integral of the muon functions. The tagging efficiency is consistently over 99%.

Neutron monitoring

Since neutron events measured in the COSINE-100 crystals can mimic the signature from WIMPs (and indeed muon induced neutrons have been suggested as a source of the DAMA signal as described in Section 3.3), an LS-based neutron detector is installed in the detector room (as mentioned in Section 4.2.4) in order to measure the rate of environmental MeV-scale “fast” neutrons in the COSINE-100 setup.

Fast neutrons, mainly produced in the COSINE-100 detector by either the decay of ^{238}U to ^{232}Th or from spallation from cosmic ray muons incident in the detector, are selected first by using a variable comparing the tail charge of the waveform of an event in the LS neutron detector with the total change of the waveform. Neutron-generated events

have a larger value in this parameter than gamma events, the main background, as can be seen in Fig. 4.11 where the alpha events have a similar distribution to neutrons [100]. This variable is known as the Pulse Shape Discrimination (PSD) parameter. An energy dependent cut in this variable is used to select an initial sample of neutron-generated events. Secondary event selection is then applied using time coincidence.

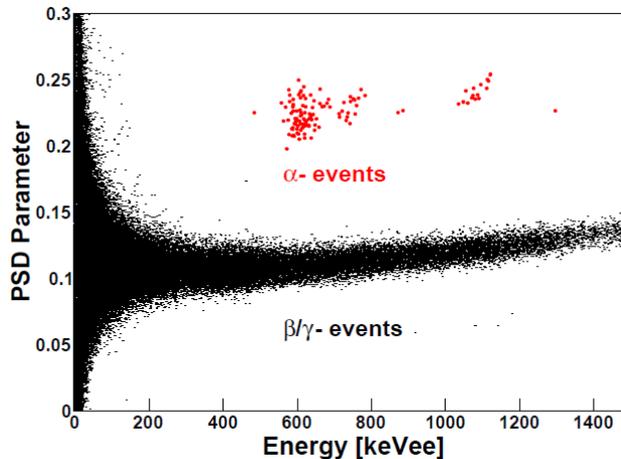


Figure 4.11: Pulse Shape Discrimination (PSD) parameter against energy for a selection of background events in the fast neutron detector at COSINE-100. α -events (red points) are separated from β and γ -events by the PSD parameter. From [100].

There are three major sources of fast neutrons generated at COSINE-100 which can be further tagged using time coincidence, all coming from the ^{238}U decay chain. These are the 5.5 MeV and 6.1 MeV neutrons arising from the α decay of ^{222}Rn to ^{218}Po followed by the decay of ^{218}Po to ^{214}Pb , and the 3.2 MeV β decay of ^{214}Bi to ^{210}Po followed by the 7.8 MeV neutron from the α decay of ^{214}Po to ^{210}Pb . Since ^{218}Po has a half-life of 3.1 minutes and ^{210}Po a half-life of 160 μs , a time window up to around 20 minutes in which two fast neutron-like events are measured can be used to tag neutrons from the ^{222}Rn to ^{218}Po to ^{214}Pb chain, and a time window up to around one minute after a measured β decay in which a fast neutron is measured is used to tag muons from ^{210}Po decay. Fast neutrons at 5.3 MeV are also produced in the decay of ^{210}Po to ^{206}Pb , however ^{210}Po has a half-life of 138 days and ^{206}Pb is stable so these neutrons cannot be tagged using a coincidence cut. The fast neutron event rate is observed to be decreasing so is fit with a linear function, but no significant annual modulation in neutron rate is observed. More information about the fast neutron detector at COSINE-100 can be found in [100].

4.4 Key differences between the COSINE-100 and DAMA experiments' setup and analysis technique

While some key aspects of the COSINE-100 detector are necessarily similar to that of the DAMA experiments to facilitate a model independent test of DAMA, certain features of the experiment, both in design and analysis technique, have been altered or added to to improve upon or further test the DAMA signal. Additionally, as a natural result of differences in construction between the two experiments, other, unintentional differences also exist between the two experiments.

A first, very simple example of differences between the COSINE-100 and DAMA experiments is the mass, distribution within the detector, and exact chemical composition of the target material. In the DAMA/LIBRA experiment, 25 NaI(Tl) crystals, each with a mass of 9.70 kg are arranged in a five by five array [69]. Since the company that made the DAMA crystals, Saint Gobain Crystals and Detectors, stopped producing similar grade NaI(Tl) before the construction of COSINE-100, a new manufacturer of NaI(Tl) crystals was required. Consequently, replicating DAMA’s 242.5 kg mass of target material was not possible, hence the lower target mass of 106 kg of NaI(Tl) in COSINE-100. As such, COSINE-100 must collect data for longer than DAMA/LIBRA to match their exposure. This however is not necessarily an issue since the DAMA result can be confirmed or refuted with significantly less exposure than the very high exposure DAMA have achieved through two decades of operation, see for example [101].

Furthermore, since the crystals for the different experiments were grown from different powders, they have different levels of background contamination and as such produce a different background spectra. The levels of background contaminants in the COSINE-100 crystals can be seen in Table. 4.1, and in [49, 69] for the DAMA experiments. Specifically the levels of ^{210}Po (in equilibrium with ^{210}Pb) in the powder used to produce the COSINE-100 crystals are elevated relative to that in DAMA/LIBRA, resulting in a higher alpha rate in COSINE-100 [73], additionally the levels of ^{40}K in the DAMA/LIBRA crystals are slightly lower than in COSINE-100 [73], on the other hand the COSINE-100 crystals have a slightly lower contamination of ^{238}U and ^{232}Th than the DAMA/LIBRA crystals [69]. The light yield of the DAMA/LIBRA crystals are around 5.5 to 7.5 PE/keV compared to around 15 PE/keV for COSINE-100.

Another fundamental difference between the experiments comes from the type of shielding used. The DAMA experiments use only passive shielding (consisting of various materials including lead, copper, cadmium, polythene, paraffin, and concrete, see Section 3.2.1 for more detail) [69] around the crystals with the goal of blocking background signals from entering the detector. Since only passive shielding is used, background events cannot be identified on an event-by-event basis. If event-by-event information were available, it could significantly strengthen DAMA’s claim, since alternative explanations posit a non-dark-matter source for the modulation, caused by DAMA’s method of subtracting expected background rates to calculate residual event-rates (see Section 3.3.2). In order to tag and more closely study the rates of background events, active shielding is used at COSINE-100 in addition to passive shielding. Namely a liquid scintillator veto for radioactive decays (described in Section 4.2.4) and plastic scintillator panels for muon tagging (described in Section 4.2.3).

Using this active shielding, studies beyond around the rate of different backgrounds have been carried out at COSINE-100. In one such example, the annual and diurnal muon flux at the COSINE-100 site was studied [102] using measurements from the plastic scintillator panels. The muon flux in four day bins with a sinusoidal annual modulation fit is shown in Fig. 4.12. A annual modulation amplitude of $(0.6 \pm 0.20)\%$ of the mean annual flux is observed with the maximum phase at 179 ± 19 days. This is in good agreement with predictions from effective temperature and muon production models. No significant diurnal modulation in the muon flux is observed.

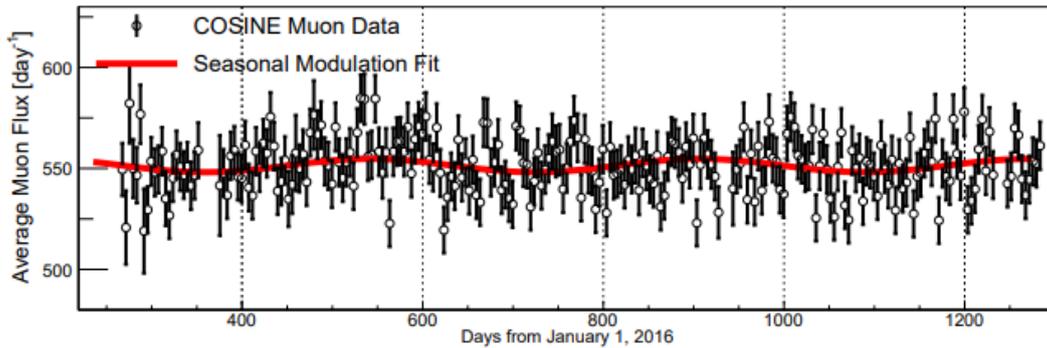


Figure 4.12: The flux over time of muons observed at the site of the COSINE-100 experiment in four day bins (black data points) with a seasonal modulation (sinusoidal) fit (red curve). From [102].

Partly as a result of these design differences, different analysis techniques are used for annual modulation searches in the two experiments. In DAMA’s case, residual event-rates are calculated by subtracting roughly yearly averaged backgrounds from the signal event rate and these residuals are fit to an annual modulation. As discussed in Section 3.3.2, this technique has also been a target of speculation as the source of the annual modulation measured by DAMA. Thanks in part to data from the LS as well as detailed simulation studies, the time-dependent background energy spectrum at COSINE-100 is very well understood (see for example [74]). Hence instead of following a process similar to the DAMA experiments and calculating residual event-rates, all the single-hit events not rejected as non-scintillation events are fit to a model designed to describe the background event-rate as well as any annual modulation component. The prior distributions of these components are determined by this time-dependent background model. Incorporating the background components into the annual modulation fit as nuisance parameters removes any possibility that a spurious annual modulation signal could be generated at COSINE-100 as a result of calculating residuals, although, as discussed in Section 3.3.2, in order to test the possibility that DAMA’s residual calculation could induce an annual modulation signal, the process was reproduced with the COSINE-100 data, separate to the main analysis [5].

4.5 Concluding remarks

The COSINE-100 dark matter detection experiment primarily aims to provide the first conclusive, model-independent test the contested claim of a positive dark matter signal from an annually modulating event-rate in NaI crystals, reported by the DAMA collaboration. To achieve this, COSINE-100 also searches for an annual modulation of signal-like events as measured in NaI crystals. Also utilised at COSINE-100 are two types of active shielding, plastic scintillator panels to tag and veto muon events, and a liquid scintillator to do the same for events from radiogenic decays. The stability of all these detector components over long time-scales is crucial to accurately observe a modulating signal with a small amplitude and long period, such as a dark matter-induced annual modulation. In the upcoming chapters, efforts to ensure the stability of the detector components will be described, then new annual modulation searches with COSINE-100 data are described.

Chapter 5

Long Term Stability of the COSINE-100 Liquid Scintillator Veto System

As made clear in previous chapters, detector stability over multiple year-long timescales is crucial for the success of COSINE-100 and other annual modulation dark matter detection experiments where the expected signal is a very small annual modulation. This is especially true in the case where unaccounted for instabilities could potentially induce a signal that mimics an annual modulation (see for example Section 3.3). Since the 2200 litres of liquid scintillator (LS) surrounding the NaI crystals at COSINE-100 is an important component in the event selection process for dark matter searches with regards to identifying and vetoing multi-hit events which are very unlikely to be from dark matter, the long-term behaviour of the LS must be well understood and any changes accounted for. Additionally, should any changes be made to the operation of the LS, the effects of these changes should be well understood and accounted for. The LS itself (described in more detail in Section 4.2.4 and shown in Fig. 4.2) consists of 2200 litres of Linear Alkyl-Benzene contained in an acrylic box and surrounding the crystal encapsulation. Signals from the LS are recorded by two 3x3 arrays of photo-multiplier tubes (PMTs) on two opposite sides of the acrylic box.

In this chapter the new gain correction applied to the LS will be described in Section 5.1, before a discussion in Section 5.2 of novel PMT-level analysis of the behaviour in the LS. This includes in Section 5.2.2 the justification for and changes resulting from removing PMT18 from the LS analysis as a result of its unstable behaviour.

5.1 Gain correction in the LS

The process of gain correction is commonplace in any large-scale particle physics detector, since if the gain of a component changes and this change isn't accounted for correctly, then that component will no longer give an accurate reading. In the case of the LS at COSINE-100 however, due to the nature of the annual modulation signal being searched for and the specific role of the LS in vetoing events, proper gain correction of the LS is of extra importance. As described in Section 4.2.1, in dark matter searches at COSINE-100, only single-hit events (events that occur in an $8 \mu\text{s}$ window with no other physics events in the detector) in a crystal are considered for dark matter searches, since the probability of two genuine dark matter signals occurring in the detector in such a short time window is

negligible. As such this single-hit requirement provides a powerful tool for removing non-dark matter-induced events. However, to avoid the incorrect tagging of single-hit events as multi-hit, a low energy threshold must be set to avoid electronic noise events in the LS (which are mainly from PMT noise [103] and far more prevalent at low energies) in coincidence with a genuine single-hit signal spuriously vetoing the signal. Consequently, an unaccounted for gain change in the LS would effectively change the low energy threshold of the LS and therefore change the single-hit event selection, thus potentially obscuring an annual modulation signal.

The gain correction carried out prior to the main body of work described in this chapter is outlined in Section 5.1.1 before a more detailed description of the LS gain correction using the Compton edge arising from ^{40}K decay in COSINE-100 from July 2018 to October 2020 in Section 5.1.2 and then across the whole data set after the removal of PMT18 in Section 5.1.3.

5.1.1 Initial LS gain correction

The gain correction described in the bulk of this section refers to work which was carried out on data taken from 1.7 years after the start of data-taking at COSINE-100 onward and so begins in July 2018. However, the gain correction was also applied in this early data taking period, albeit calculated using a different method to the rest of the LS gain correction. In this early period, particularly towards the start of the data-taking period, the gain of the LS was very volatile, as shown in Fig. 5.1 where the mean of the LS spectrum, constructed from the summed energy spectra of the 18 PMTs for 50-subrun (100 hour) bins is plotted for approximately the first 2.5 years of data taking. In around 100 days, the LS gain dropped by roughly 50% before sharply increasing to slightly above the original level, before then dropping again more gradually. The effect this initial drop had on the LS spectrum is shown in Fig. 5.2.

During this period of early volatility, gain correction was not yet optimised and was not applied with a repeatable, quantitative method. Data was most commonly, but not exclusively, split up into 100 hour chunks and corrected by hand, applying a multiplicative correction factor to restore the LS energy spectrum to the location of a reference spectrum at the start of the data taking period (similar to the red spectrum in Fig. 5.2) by eye. While this method of gain correction was acceptable for early periods of data taking, especially given the rapidly changing behaviour of the LS response, a more convenient and quantitative method was required after the early volatile behaviour of the LS had abated and the above method of gain correction was ended after 1.7 years of data taking. The cause of the sharp rise and fall in LS response shown in Fig. 5.1 is yet unknown, but no repeat of such behaviour has been observed in subsequent data taking. Upon the deconstruction of the detector in the Spring of 2023 in order to facilitate the moving of the crystals to a new site in preparation for the upgrade to COSINE-100, it was discovered that some of the LS liquid had permeated into the casing for some of the LS PMTs which may possibly have been the cause of the fluctuating behaviour.

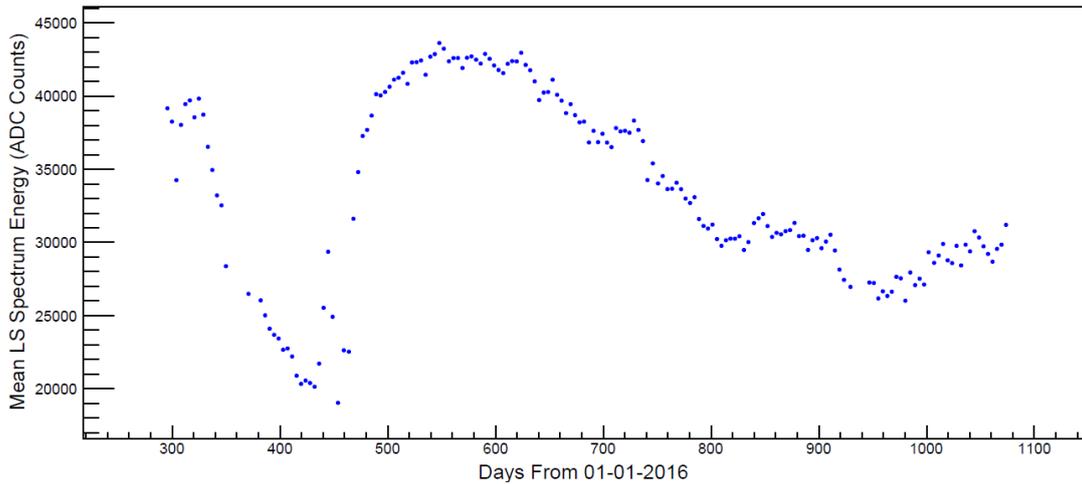


Figure 5.1: The mean energy of the liquid scintillator energy spectrum (acting as a proxy for LS gain) for approximately the first 2.5 years of data taking in 50-subrun (100 hour) bins.

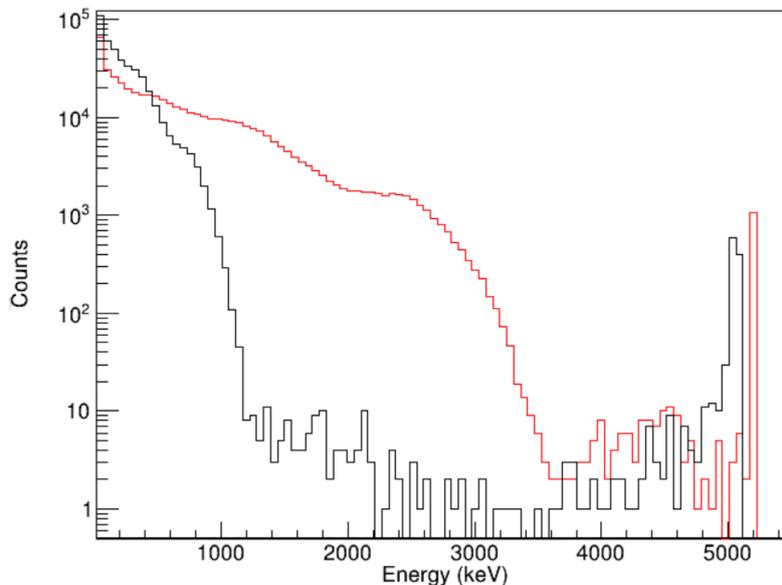


Figure 5.2: The LS energy spectra corresponding to 100 hours of data at the peak in Fig. 5.1 at around 310 days (red line) and trough at around 410 days (black line).

5.1.2 LS gain correction for the first three years of data taking

As described in Section 4.2.2, as K is in the same chemical group as Na (alkali metals), they share chemical properties and consequently it is in practice impossible to produce NaI crystals without some amount of K, and therefore ^{40}K , present since natural K will contain around 0.012% ^{40}K [104]. It is signals from the decay of ^{40}K that allow the LS gain to be quantitatively tracked.

The ^{40}K Compton edge

^{40}K is a potentially problematic source of background events for experiments searching for a signal of the order of a few keV such as COSINE-100 since one of ^{40}K 's decay

channels is to decay to ^{40}Ar in the excited state via electron capture, producing a 3.2 keV Auger electron [104] which creates background events directly in the signal region of the experiment. The excited ^{40}Ar nucleus then relaxes to the ground state via the prompt emission of a 1461 keV γ ray photon. The half-life of ^{40}K is very long (approximately 1.3 billion years [104]), but since it is present in relatively high amounts (see Table. 4.1) in the COSINE-100 crystals, a significant signal from the electron capture decay chain is still produced. This signal is effectively constant in time due to the long half-life of ^{40}K .

While the 3.2 keV electron is potentially problematic for experiments such as COSINE-100, the coincident 1461 keV γ provides a natural method for tagging 3.2 keV electrons from ^{40}K decay, provided the 1461 keV signals can be efficiently detected. As such, the LS at COSINE-100 is a very useful tool since the 1461 keV γ produces a clearly visible Compton edge (also known as a Compton shoulder) in the LS, as can be seen between 1000 and 2000 keV in Fig. 5.2. This not only allows for 3.2 keV electrons from ^{40}K decay detected in the crystals to be tagged through their coincident γ events measured in the LS, but the location of the Compton edge from the 1461 keV γ s can be tracked through fitting. Since the rate of production of these γ s is effectively constant, the feature therefore provides a useful tool in tracking the long term behaviour and changes in gain of the LS.

The spectral feature produced by the 1461 keV γ s is a Compton edge somewhere below 1461 keV rather than a more Gaussian-like peak at around 1461 keV as a result of the combination of the γ s losing some energy through scattering events in the detector before finally being deposited in the LS, as well as the kinematics of the scattering resulting in a distinct high energy cutoff, above which no additional energy can be deposited, which occurs at the maximum scattering angle of 180° . During a Compton scatter of a photon with energy E , the energy of an outgoing photon, E' , is given by

$$E' = \frac{E}{1 + \frac{E}{m_e}(1 - \cos\theta)} \quad (5.1)$$

for energy in units of keV, where m_e is the mass of the electron in units of keV/c^2 and θ is the scattering angle. When $\theta = 180^\circ$ and $E = 1461$ keV, the energy of the outgoing photon is around 217 keV, so consequently no more than around 1244 keV may be deposited into the LS from a single scatter of a 1461 keV γ .

Fitting and tracking the Compton edge

To fit this Compton edge and subsequently quantitatively monitor the gain of the LS over time, the following phenomenological function was developed:

$$y = \frac{p_0}{\exp[p_1(x - p_2)] + 1} \quad (5.2)$$

where y is the number of counts measured in the LS, x is the energy of an event, and p_0 , p_1 , and p_2 represent fit parameters that respectively describe the height, shape, and location in energy of the Compton edge. By tracking the location of the Compton edge, the LS gain can be quantified and therefore corrected. An example to this fit on 100 hours of LS data is shown in Fig. 5.3 where analogue-to-digital conversion units (ADC counts) have been used instead of keV for simplicity. In this case 1 keV corresponds to 143.8 ADC counts. A more thorough discussion of the conversion between ADC counts to keV can be found in Section 5.2.3. The two other major features on the spectrum at around 70,000 ADC counts and 340,000 ADC counts are from annihilation of positrons produced

in β^+ decay, for example in ^{22}Na decay, which produce 511 keV γ s, and 2.6 MeV γ events resulting from the decay of ^{208}Tl .

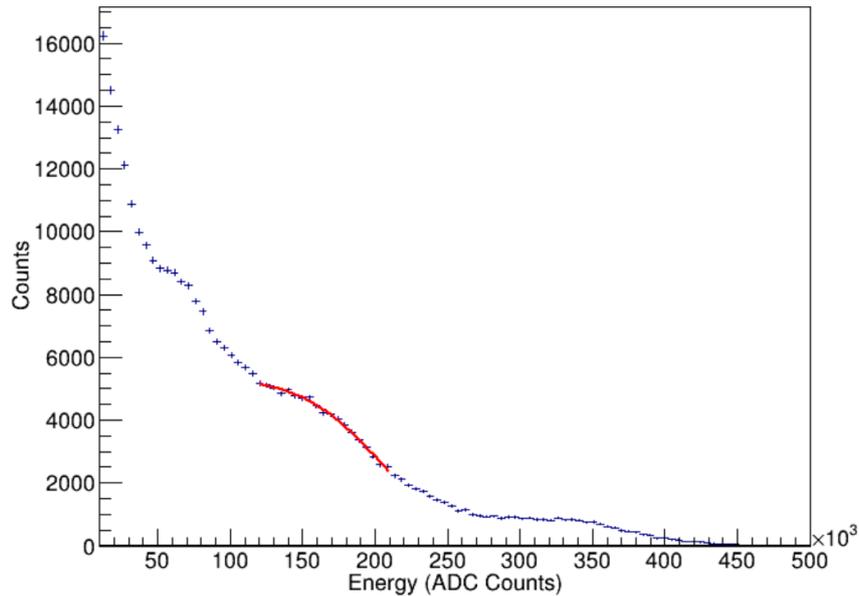


Figure 5.3: Energy spectrum for 100 hours of LS data (blue data points) with the fit described in eq. 5.2 applied to the area of the spectrum corresponding to the Compton edge from 1461 keV ^{40}K decay.

The location of the 1461 keV Compton edge as described by p_2 in eq. 5.2 was tracked from July 2018 (the time where the initial gain correction described in Section 5.1.1 ends) until November 2019 by plotting an LS energy spectrum for 50-subrun long bins of data, roughly corresponding to 100 hours each¹. The location of the 1461 keV Compton edge for each bin is then compared to the location of the Compton edge from the fit of a 50-subrun reference spectrum from the start of data taking in October 2016. The inverse ratio of the measured location of the Compton edge in each bin compared to the reference fit is shown for the above time period in Fig. 5.4. A roughly linear in time decrease in the location of the Compton edge is observed. This

This plot (where the value of the location of the reference fit is divided by that of the Compton edge) is used for the gain correction and is fit with a linear function to generate a time-dependent function with the output of a multiplicative factor that, when applied to the LS energy spectrum, should restore the location of the 1461 keV Compton edge and as such the rest of the spectrum to that of the reference. Uncertainties are calculated by propagating the uncertainties outputted by the fitter from the best fit value of the fit to the Compton edge of the bin in question and the reference fit.

¹Each subrun is two hours long so each bin contains 100 hours of data, but in the case where the detector has any down time subruns may not be continuous leading to a 50-subrun bin finishing over 100 hours after the bin started

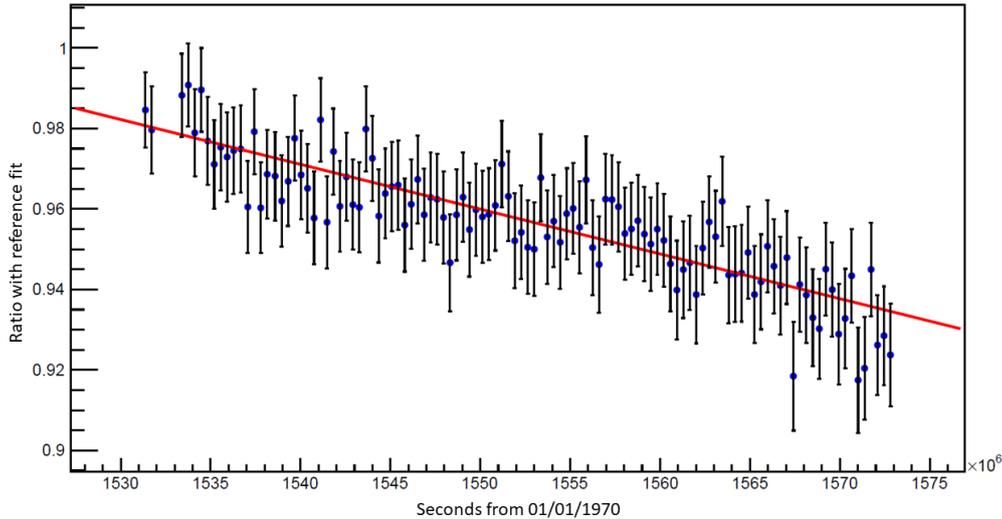


Figure 5.4: Inverse ratio (blue data points) of the location of the ^{40}K Compton edge of the LS spectrum from fitting in 50-subrun bins from July 2018 to November 2019, to the location from a reference fit from the start of data taking (October 2016). A clear decrease with time can be seen representing an increase in the gain of the LS. This increase is fitted to a linear function (red line) and this fit function is used as a time-dependent correction function.

The COSINE-100 data are then reproduced with this time dependent gain correction applied and the location of the Compton edge is again checked against the reference value. This is shown in Fig. 5.5 for data from the beginning of this phase of LS gain correction in July 2018 to the start of an approximately month long period of calibration starting in October 2020. In order for this process to continue to correctly correct the LS gain beyond the point in November 2019 when the correction function is calculated, the uncorrected location of the Compton edge must continue to change in the same linear fashion as seen in Fig. 5.4 which is not guaranteed. As such the location of the corrected ^{40}K Compton edge is periodically checked to verify the continued efficacy of the gain correction. As Fig. 5.5 shows, this assumption generally holds until October 2020. While there are still some fluctuations in the location of the Compton edge in the gain corrected data, these fluctuations are within 2% of the reference value which is deemed to be an acceptable level. This gain correction was used in COSINE-100 analysis until after the modulation search from three years of data was published in 2022.

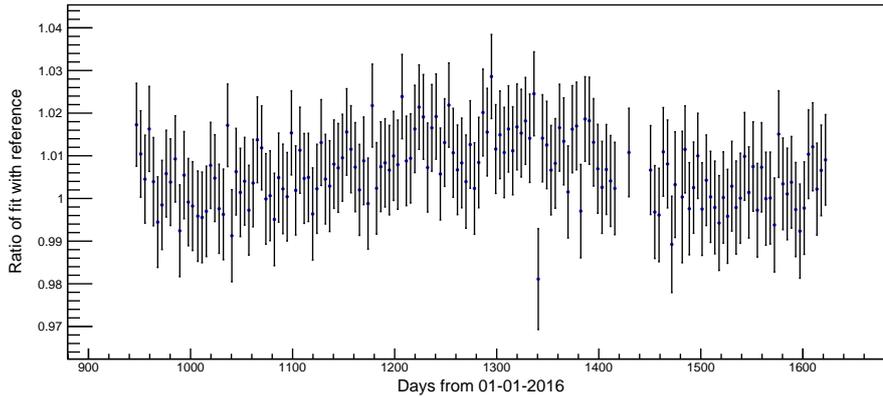


Figure 5.5: The ratio of the location of the ^{40}K Compton edge (as found by fitting Eq. 5.2 to the LS energy spectra) to the reference value after the correction described in Section 5.1.2 for 50-subrun bins from July 2018 to October 2020. The source of the anomalous data point at 1340 days is not well understood.

5.1.3 Second phase of LS gain correction after October 2020

The LS was observed to be relatively stable from July 2018 to October 2020 and in this period no changes were made to the operation of the LS. On the 5th of October 2020 normal physics data taking at COSINE-100 was halted in order to carry out a programme of calibration runs. Two short (approximately eight hour) calibration runs, first using an ^{241}Am source, then a ^{60}Co source, were initially carried out before a ^{22}Na source was used to collect calibration data for approximately 43 days until the 17th of November 2021. While this calibration run was necessary to collect a sample of signal data to be used in the training of a new boosted decision tree (BDT) variable with the goal of improving low-energy noise discrimination and hence reducing the detector low-energy threshold from 1 keV to 0.75 keV in future analyses, it had a somewhat disruptive effect on some of the detector components, including the LS. Specifically, after the period of gain correction, the location of the ^{40}K Compton edge is observed to drop relatively rapidly. After this period, the fluctuations and uncertainties in the location of the ^{40}K Compton edge between concurrent bins become larger than that observed at previous times.

While this instability is likely in part caused by the calibration run, two other significant factors also must be taken into consideration. The first is that around this time it was observed that one of the LS PMTs (PMT18) had been exhibiting increasingly unstable behaviour of an unknown origin, but possibly due to LS liquid leaking into the PMT housing. This behaviour could not be resolved through software corrections (e.g. gain correction etc) so the decision was taken to remove the signals from the PMT from the overall LS variables throughout the data taking period. Separate to the issue in the LS PMT, the noise rate of one of C8's PMTs suddenly increased in January 2021 and signals from this PMT also had to be removed from any event selection process below 20 keV as described in Section 9.2. This slightly affected the LS event selection as now signals in the LS could be recorded from an event observed in just one of C8's (the eight NaI crystals at COSINE-100 are labelled C1, C2 etc) PMTs as opposed to requiring both crystal PMTs to register a signal, as in the other crystals. Then in August 2021 (after the end of the data-taking period used in the annual modulation search described in this work), the PMT noise-rate in the other PMT for C8 and one of C5's PMTs also increased.

Therefore signals from these crystals below 20 keV were discarded in the event selection process entirely from this point.

The discarding of signals from one of the LS PMTs across the whole data-set has clear ramifications on the operation of the LS. For example, the gain correction had to be recalculated across the whole data-taking, and the energy calibration similarly re-considered. This is described in more detail in Section 5.2.2. Discarding signals from only a single PMT in C8 has a negligible impact on the LS operation, since signals from the crystal were still measured by the remaining PMT. However, when the signals from C5 and C8 were discarded completely, the event-rate in the LS was noticeably affected since LS signals require a coincident crystal signal to be recorded and after this point it any signals that previously would previously have been tagged by events in C5 or C8 were no longer recorded.

Despite the removal of one of the LS PMTs, the shape of the LS spectrum does not change significantly so the general process of calculating the LS gain using the Compton edge from ^{40}K decay remained the same with eq. 5.2 still used to fit the LS energy spectrum and locate the feature. A new reference fit value was calculated using the same method as previously (fitting an energy spectrum from the start of data taking), but with PMT18 removed from the spectrum. The entire LS data-set at the time was then split into 50-subrun bins and the location of the ^{40}K Compton edge located without PMT18. This is shown in Fig. 5.6. Note that in this plot the previous gain correction is still applied and subsequent gain corrections are applied alongside previous corrections. The rise and fall of ^{40}K Compton edge at the beginning of the data-taking period corresponds to the feature shown in Fig. 5.1 and described in Section 5.1.1.

It is clear from this figure that a single linear correction will not be sufficient to restore the LS spectrum across the entire period of data taking. Instead the data is split into three sections, following the sections described in Section 5.1.1 and Section 5.1.2, as well as a third section starting after the calibration period. It is also relevant to note that despite PMT18 in the LS being observed to have divergent behaviour from the PMTs throughout the data taking period, the trends in the movement of the ^{40}K Compton edge are broadly the same regardless of whether the PMT is included when building the LS energy spectrum, indicating the PMT is not the cause of the shifting LS gain.

While the first period of data taking (which was initially corrected by hand as described in Section 5.1.1), exhibits the greatest fluctuations in the location of the ^{40}K Compton edge, there is no clear overarching trend in behaviour. Therefore for the sake of brevity, the gain correction for this period was left mostly unchanged relative to the original correction with the exception of a small number of by-eye additional tweaks to individual runs to further restore the ratio of Compton edges to unity.

The entirety of the middle section between July 2018 and October 2020 (the start of the ^{22}Na calibration run) was split into 50-subrun bins and the location of the ^{40}K Compton edge was found for each bin. The gain change throughout this region once PMT18 had been removed was observed to still be linear and this trend was fit with a linear function as in Section 5.1.2 to generate a time dependent gain correction function. Likewise, the third section from November 2020 (the end of the ^{22}Na calibration run) until June 2021 was analysed in the same way and a time dependent linear correction function was generated which was applied to all subsequent data. The corrected location of the ^{40}K Compton edge is shown from the start of data taking to February 2023 in Fig. 5.7. The behaviour is as expected with good agreement between the fitted location of the ^{40}K Compton edge and the reference value until around 2000 days. This corresponds to the removal of signals from C5 and C8 in August 2021 due to high noise in the crystals'

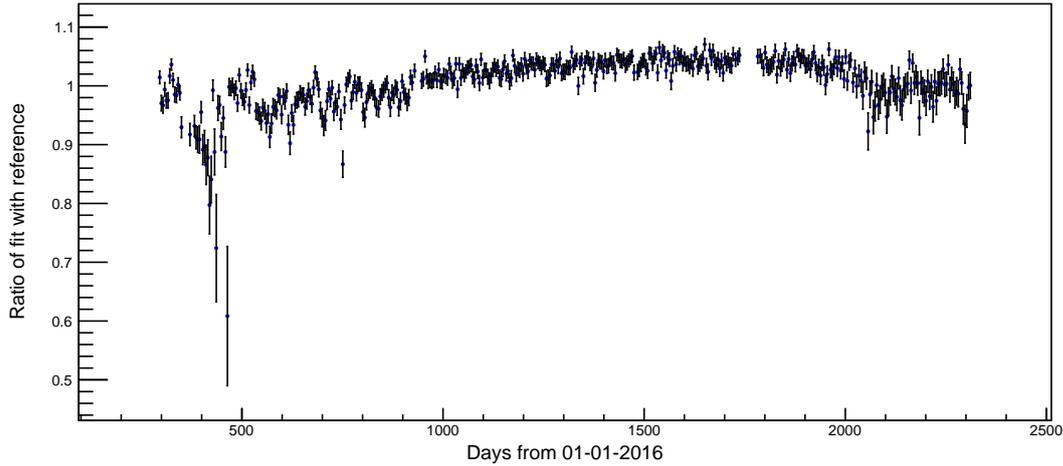


Figure 5.6: Ratio of the location of the ^{40}K Compton edge of the LS spectrum plotted without signals from PMT18 to a new reference value from fitting in 50-subrun bins from the start of data taking in October 2016 to June 2022. Three distinct periods are considered, the early period of large fluctuations (until approximately 900 days), a middle period of relative stability until a period of no data taking for a calibration period (approximately 900 to 1700 days), and later period of increased fluctuations from approximately 1750 days onward.

PMTs, described above. The subsequent change in gain from this period was also linear in time so was corrected using the same method as in previous cases. The final corrected location of the ^{40}K Compton edge in the LS is shown in Fig. 5.8 where stable behaviour is achieved as shown up till the end of the first phase of COSINE-100 operation.

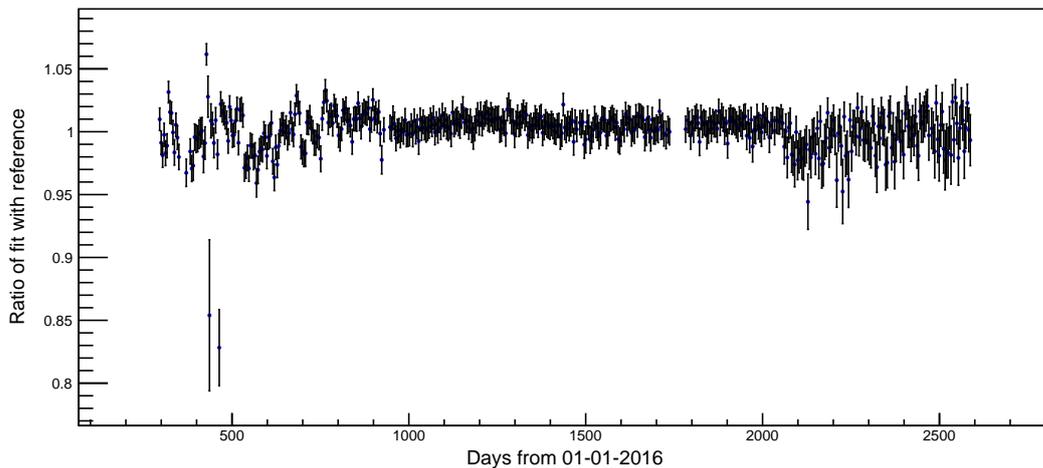


Figure 5.7: The ratio of the location of the ^{40}K Compton edge relative to the reference location after gain correction without LS PMT18 after an initial gain correction was applied based on the changing gain on a sample of data after the calibration run. A period requiring further correction remains in the final portion of the data-set due to a change in event selection as a result of high noise in C5 and C8.

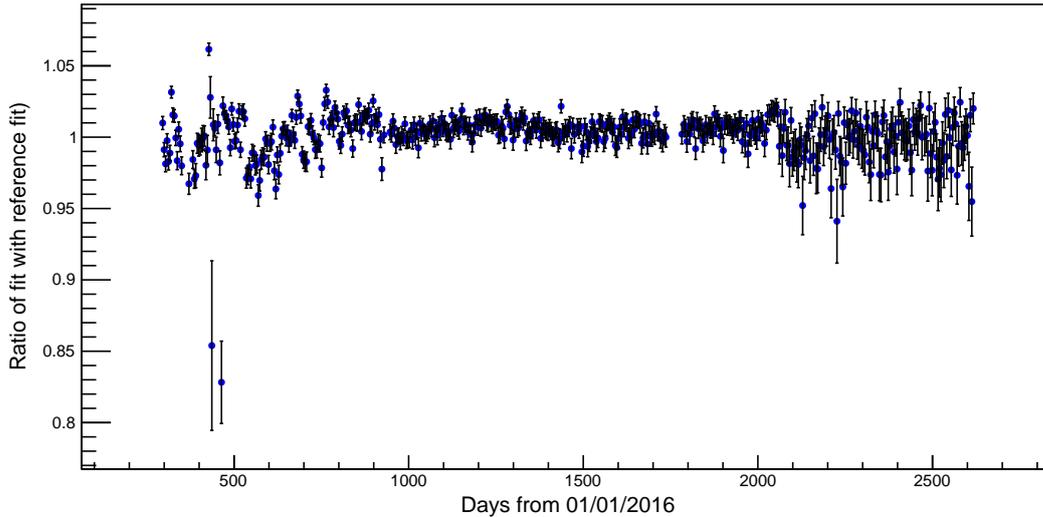


Figure 5.8: The final ratio of the location of the ^{40}K Compton relative to the reference location after gain correction have been applied. The LS gain is shown to be stable across the entire data-set.

5.2 PMT-level studies of LS behaviour

The response and behaviour of the LS is broadly dependent on the individual behaviours of two separate components: the behaviour of the liquid itself, for example responding to changes in temperature or any other environmental conditions or responding to any contamination introduced, and separately the behaviour of the 18 PMTs which measure and characterise the signals produced in the LS. Generally speaking, thus far the LS has been described as a single, homogeneous component. While this is true for the liquid, it is not true of the PMTs which may degrade and change in behaviour completely independently from one another. Therefore to fully understand the changing behaviour of the LS, the individual PMTs must be studied separately.

In this section LS behaviour at the PMT level is described, looking initially at all of the PMTs in 5.2.1, before a specific focus is applied to PMT18 in Section 5.2.2, motivating its removal from the calculation of LS variables, finally the recalculation of LS parameters after the removal of PMT18 is described in Section 5.2.3.

5.2.1 Changing PMT event rates

LS energy spectra shown in Section 5.1 are formed from the summed spectra of the 18 LS PMTs, and when combined in this way, spectral features such as the ^{40}K Compton edge can be observed and tracked. However, in the case of individual PMT spectra themselves (see for example Fig. 5.9 where 500 subruns of data are plotted for a single PMT), these spectral features are not prominent enough to be fitted, even when large sample size is used. Note that again ADC counts are used for energy for the sake of brevity as there has been no dedicated calibration between ADC units and keV for individual PMTs. An order of magnitude estimation of the ADC to keV calibration factor can be made by assuming each PMT equally contributes one eighteenth of the overall LS ADC counts, in which case 1 keV corresponds to roughly 8 ADC counts in a given PMT, using the initial calibration

value of 143.8 ADC counts per keV for the overall LS.

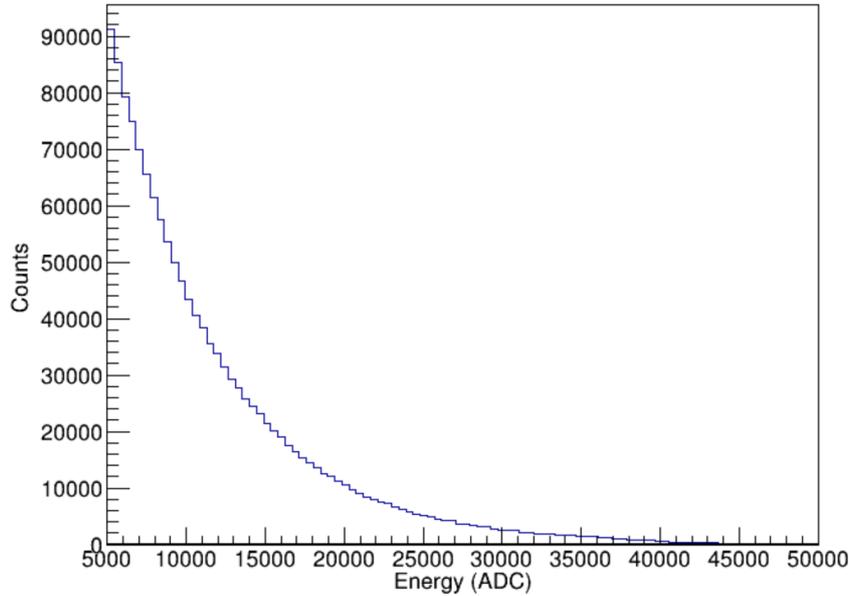


Figure 5.9: The energy spectrum of an individual LS PMT (PMT0) for 500 subruns of data. No spectral features significant enough to fit are present.

As such, the gain of the individual LS PMTs cannot be tracked through the location of the ^{40}K Compton edges or any other spectral features. It is however possible to observe that the individual PMTs do not all exhibit the same changing behaviour over time, for example by plotting the spectra of individual PMTs at different times, as in Fig. 5.10 where the spectra for each PMT is shown for 100 subruns of data at different points between July 2018 and August 2020. Note that the PMTs are arranged in two 3×3 arrays at either side of the LS tank and their variable names are assigned such that each PMT is in a pair with the opposite PMT on the other array, i.e. PMT0 is opposite to PMT10 etc. This results in the PMT9 variable being left intentionally blank in this, and all subsequent plots.

An initial method to more clearly observe the changing behaviours of the PMTs was to simply draw an energy spectrum (starting at 100 ADC counts to remove low energy noise events) for each individual PMT in 50-subrun bins and record the integral of each spectrum; i.e. the total number of events above 100 ADC counts. This is shown for data between October 2016 and June 2022 in Fig. 5.11. Many features of interest stand out in this plot which will be described in turn.

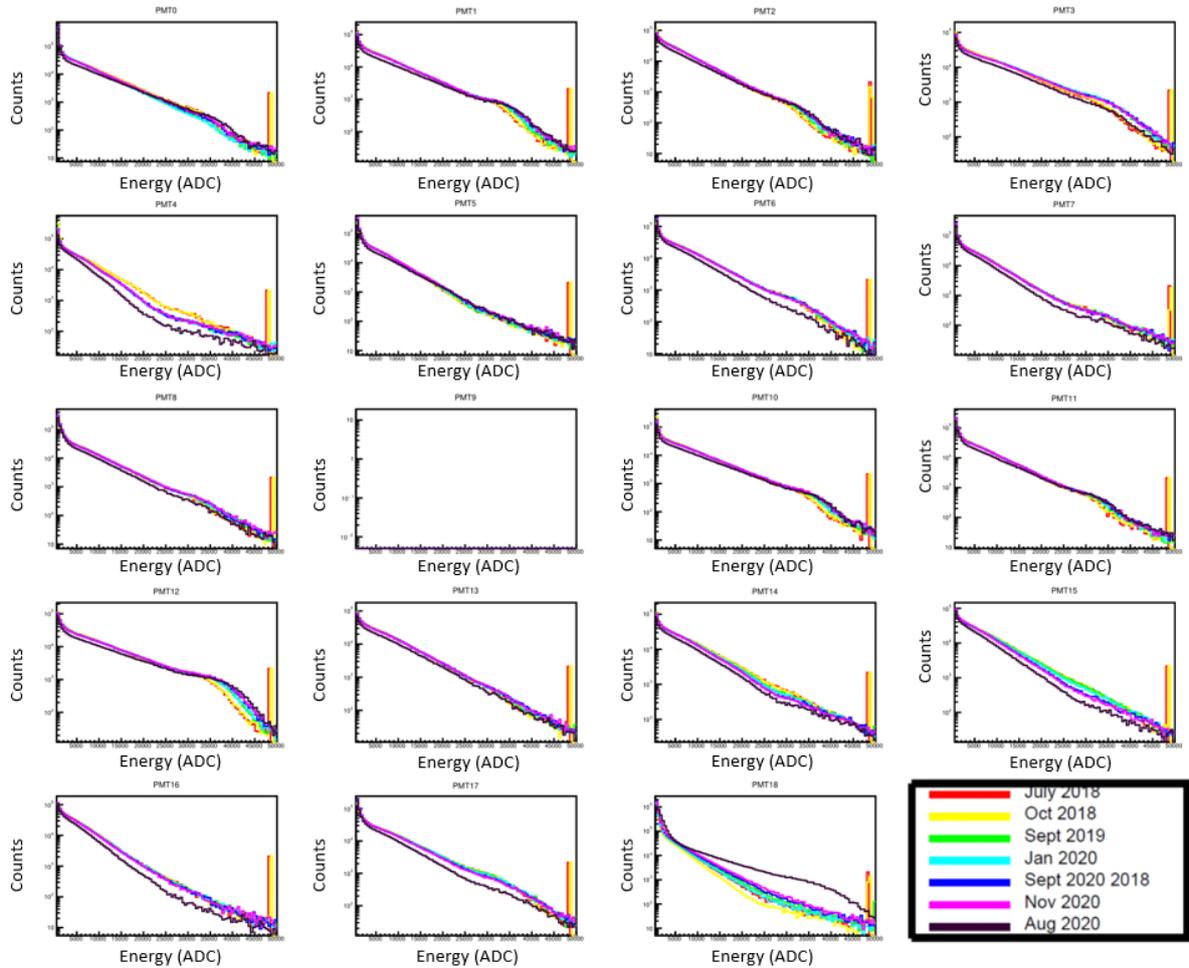


Figure 5.10: The energy spectra of all 18 LS PMTs for 100 subruns of data at different points from July 2018 to August 2020, showing that the spectra of different PMTs change in different ways over time. Note that PMT9 is a placeholder variable and is left intentionally blank.

All of the PMTs exhibit a sharp drop in the measured event rate at around 2000 days from the start of 2016 (i.e. July 2021). This is a well understood feature and arises as a result of a change in crystal event selection in C8 as described in Section 5.1.3. The PMT noise rate of one of the PMTs in C8 was observed to have significantly increased at the start of 2021 so in order to prevent these noise events from erroneously acting to tag potential single-hit/signal events as multi-hit events, the low energy threshold of this PMT was increased by a factor of 20. The effect of this change was relatively insignificant in terms of the LS event selection since the second PMT of C8 was still operating at a normal threshold. However in July of 2021, the noise rate in the second PMT of C8 also increased in a similar manner to the first PMT, with similar behaviour also observed in one of the C5 PMTs. Consequently the low energy thresholds for these crystals also had to be increased. As a result, there was no way of triggering LS events that occurred only in coincidence with C8 or C5, and so there was a noticeable decrease in the event rate across the LS. The cause of the increased noise in the PMTs is not known. This feature is the only one shared by all the PMTs on these plots.

Visible in all of the PMTs aside from PMT18 is a large spike in the number of events at around 350 days. This corresponds to the unexplained increase then decrease in event rate seen in the LS at the beginning of data taking, described in Section 5.1.1. In general

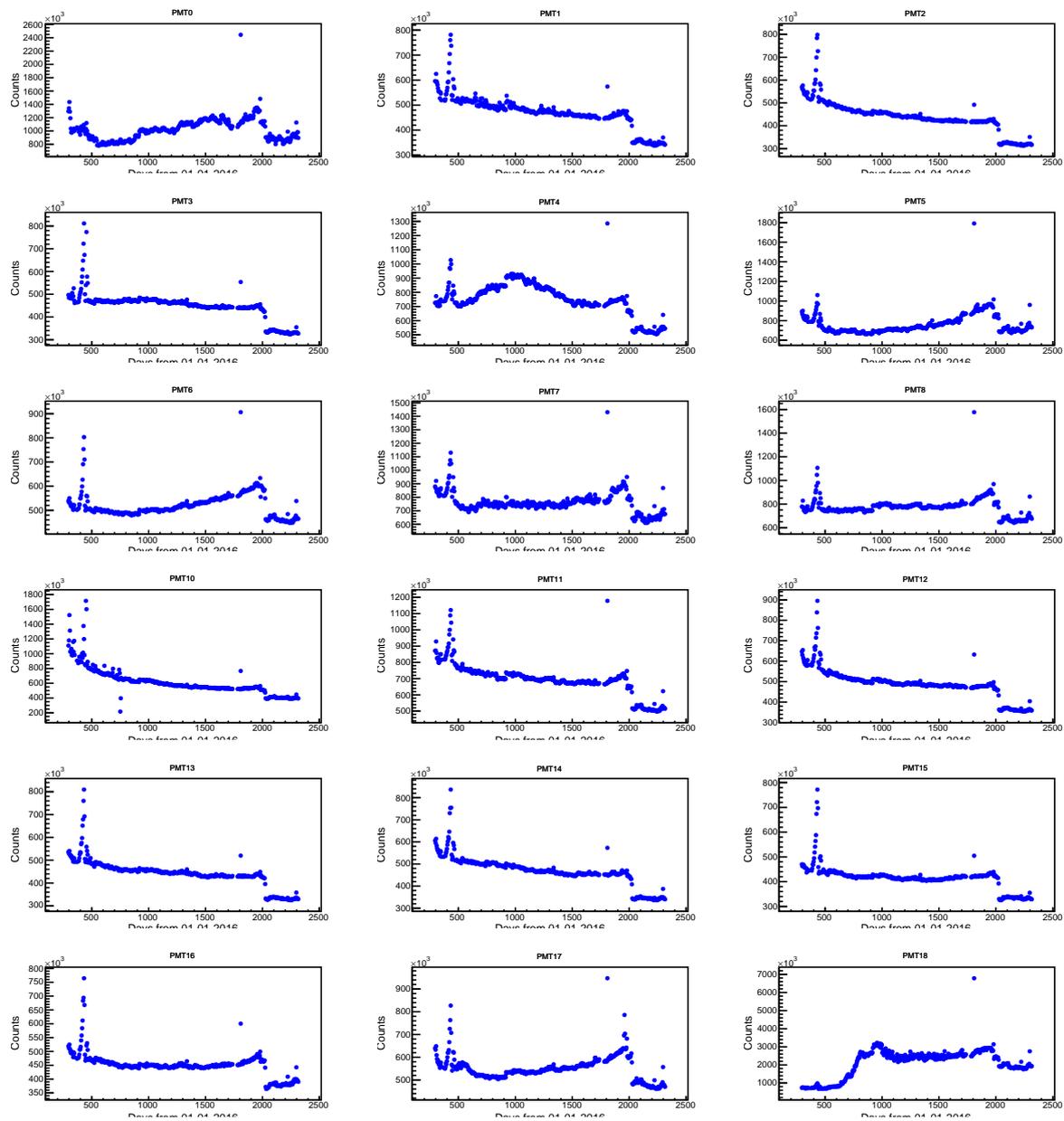


Figure 5.11: The number of events above 100 ADC counts (to remove obvious PMT noise) for each of the LS PMTs in 50-subrun bins from October 2016 to June 2022 (excluding calibration runs). The x-axis is in days from 01/01/2016.

the event rate of the PMTs slowly decreases after this initial spike, as would be expected as the radiogenic backgrounds in the detector with relatively short half-lives deplete over time. The cause of the single-point spike at around 1700 days is unknown. Generally the PMT rate decreases by around 0.5×10^5 to 1×10^5 counts per bin over the period of decreasing rate before the change in event selection. An uptick in the event rate before the threshold changes in the PMT of C5 and C8 is also seen, presumably as a result of the increasing noise-rate in the C8 and C5 PMTs triggering additional LS events. However, not all the PMTs follow this general trend. In the cases of PMTs 5, 6, 7, 8 and 17 there is no discernible decrease in event rate before the change in crystal event selection and instead a flat or gradually rising event rate is observed. Furthermore the uptick before the change in crystal event selection is not equally large in all LS PMTs, though this is potentially explained by the detector geometry, with PMTs closer to C8 (PMTs 7, 8, 17, and 18) showing a larger effect.

PMTs 0, 4, 10, and 18 have a less consistent event rate profile relative to the other PMTs. The event rate of PMT0 briefly drops after the initial spike, but then begins to rise. It also exhibits a less smooth change in event rate compared to the other PMTs with a rising event rate as well as a higher event rate than all the other PMTs aside from PMT18. Apart from during the initial event rate spike, the rates of the other PMTs largely do not exceed 10^6 counts per bin, whereas in PMT0 the event rate is predominantly over 10^6 counts per bin.

The rate of PMT4 rises and falls in a way not seen in any of the other PMTs with the event rate steadily rising and peaking around September 2018 before dropping again in a manner resembling some of the other PMTs. The amount by which the event rate of PMT4 rises and falls (around 2×10^5 counts per bin) is somewhat larger with that of most of the other PMTs. The shape of PMT10's rate distribution over time is broadly consistent with other PMTs with a decreasing rate profile, however the degree to which the rate drops, particularly at early times is unusual.

While the divergent behaviours of the PMTs described above are not well understood, it is thought that their effects are not significantly problematic for the use of the LS as a tool for event selection and they can be largely rectified with the overall LS gain correction described in Section 5.1. This is partly from studying the noise rate of the PMTs by plotting the event rate of the low energy events below 100 ADC counts that are excluded in Fig. 5.11 and observing they are consistent across all the PMTs aside from PMT18, shown in Fig. 5.12, and partly from observing that while the signal event rates of individual PMTs are fairly changeable, these divergent behaviours are largely smoothed out in the overall rate of LS events (excluding PMT18), as seen in Fig. 5.13.

Fig. 5.13 also reveals a feature not immediately obvious in the individual PMT rate plots: a discrete increase in the overall rate of PMT events at around 930 days. This directly corresponds to the end of a ^{60}Co calibration run (note that data from calibration runs is not shown in the PMT rate plots). It is not clear what caused this increase in rate; however, as noted in Section 5.1.3, calibration runs can be disruptive and destabilising for detector components. Finally, as expected, a clear decrease in the LS event-rate is observed towards the end of the data-taking period, corresponding to C5 and C8 no longer being used in the selection of LS events.

CHAPTER 5. LONG TERM STABILITY OF THE COSINE-100 LIQUID SCINTILLATOR VETO SYSTEM

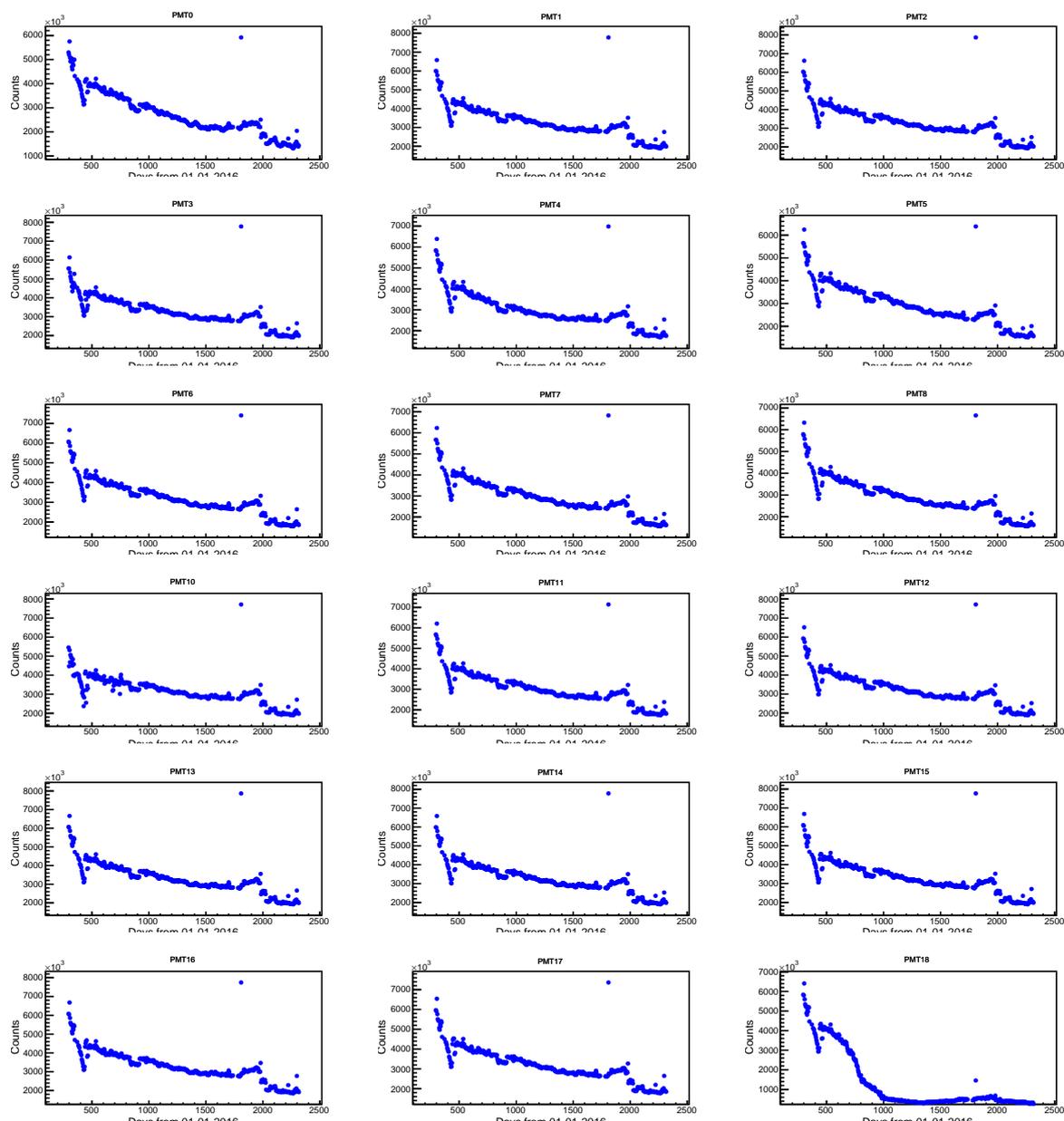


Figure 5.12: The rate of PMT noise in each of the LS PMTs as determined by the number of events below 100 ADC units in 50-subrun bins from October 2016 to June 2022 (excluding calibration runs) in units of days from 01/01/2016.

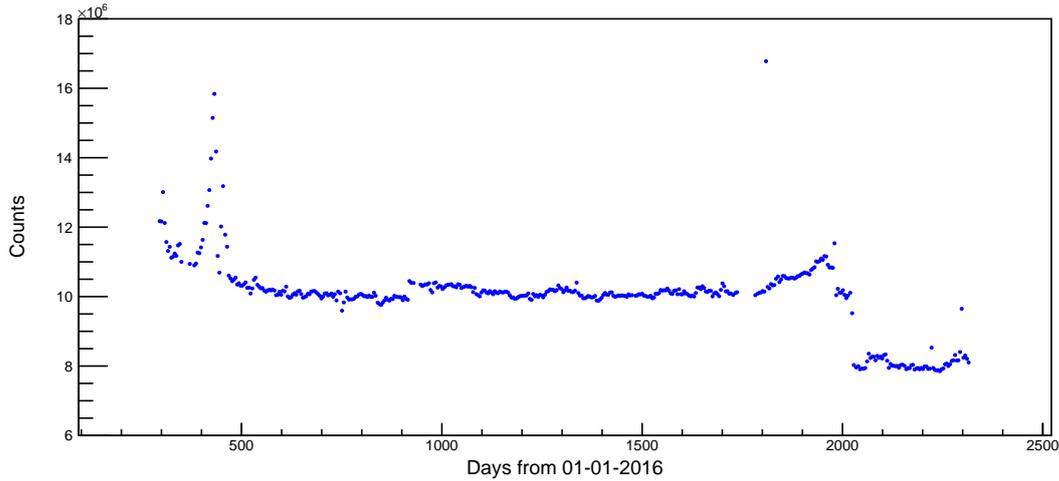


Figure 5.13: The total rate of LS events as calculated by summing the number of events above 100 ADC units for a given PMT for all PMTs except PMT18 in 50-subrun bins between October 2016 and June 2022 (excluding calibration runs).

5.2.2 Studies of unusual behaviour in PMT18

As has been alluded to in the previous section and is shown in Fig. 5.11 and Fig. 5.12, the behaviour of PMT18 of the LS is particularly problematic. Specifically, in Fig. 5.11 it can be seen that the number of events above 100 ADC units per bin for PMT18 is initially typical, i.e. slightly below 10^6 counts per bin, but then the “event” rate dramatically rises to nearly triple that of the other PMTs. At the same time the rate of events below 100 ADC units, tagged as noise events, drops to almost zero. In effect PMT noise events in PMT 18 that would previously have been measured as very low energy and consequently discarded are leaking into the signal region, potentially affecting the single-hit searches in the crystals. This behaviour is far more dangerous to physics searches than the behaviour of the other PMTs, as in those cases it appears that no additional noise is being introduced into the LS signal region. As such the behaviour of PMT18 must be studied in more detail to attempt to understand the cause of the issue, and to see if it is resolvable through some change in analysis techniques, or if the data from the PMT must be discarded.

Plotting the energy spectra of PMT18 for 100 subruns of data at different points throughout the analysis period, as shown in Fig. 5.14, shows the changing shape of the PMT18 energy spectrum over time. In this plot, several interesting features are shown. The first is, as expected from Fig. 5.11 and Fig. 5.12, a shifting of the low energy/noise peak into higher energies. Meanwhile at higher energies until September 2020 the spectra of PMT18 behave roughly as would be expected from observing the behaviour of the other PMTs in Fig. 5.11 in that there is a drop in event rate. However, at some point between September 2020 and August 2021, the number of events in PMT18 significantly increases across all energies above the low energy noise peak. In addition to simply increasing in event rate, the shape of the spectra also changes significantly relative to both the previous spectra, and the reference spectrum from the start of data taking. Finally another unexpected feature seen in the plot is the shifting of the saturation events, located in the spike in event rate at the high end of the spectrum, to successively higher points throughout the data taking period.

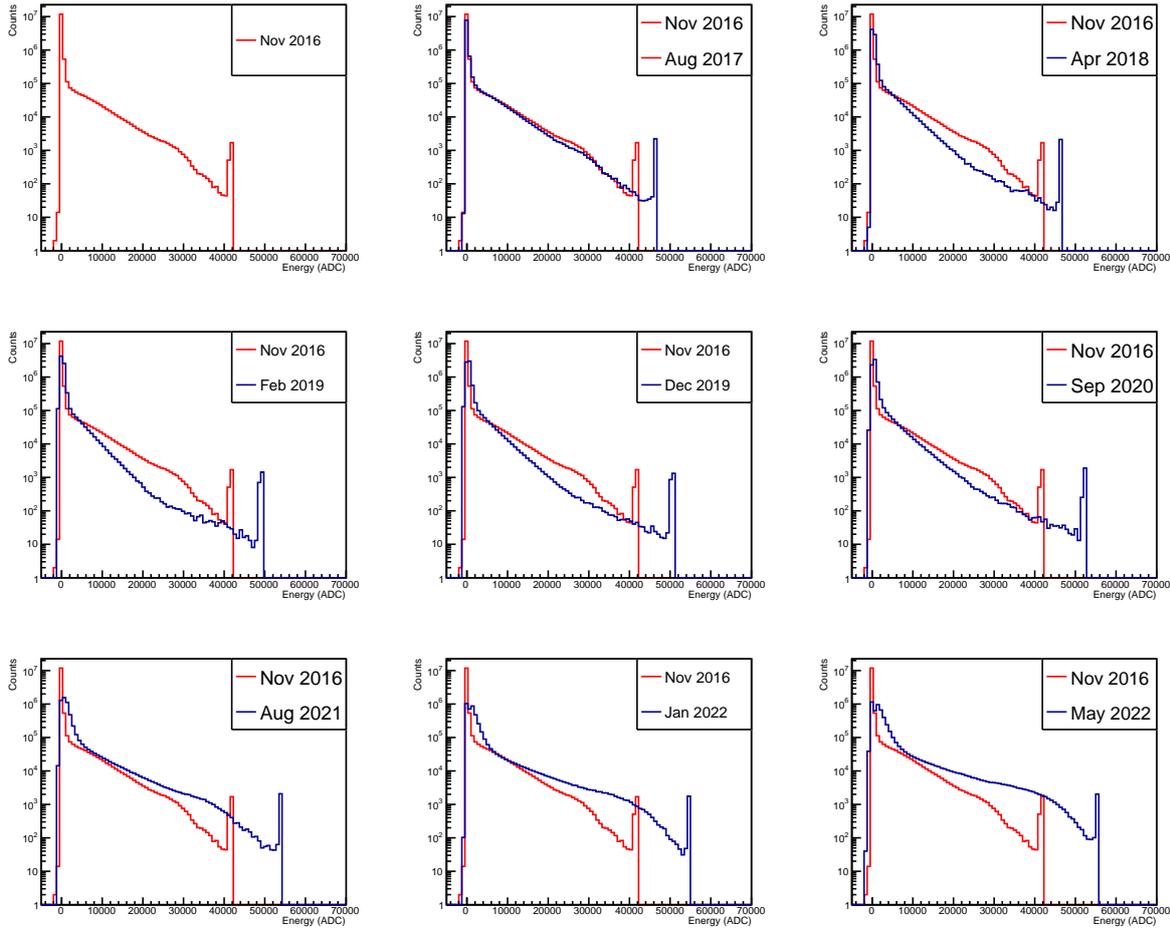


Figure 5.14: The energy spectra for 100 subruns of LS PMT18 data throughout the data taking period (blue spectra) compared to a reference spectrum taken at the start of the data taking period (red spectra).

Since this period includes the ^{22}Na calibration run (from the 6th of October 2020 to the 17th of November), it is natural to first check whether the calibration run has caused some immediate change in behaviour in the PMT. However, as is shown in Fig. 5.15, there is no large shift in the location of the spectra before and after the calibration period. Instead, the rate seems to increase rapidly but steadily after the calibration run. This hypothesis is confirmed by counting the number of events in 50-subrun bins in PMT18 above 10000 ADC units from July 2020 to August 2021, shown in Fig. 5.16, where it becomes clear the calibration run (the gap in the data) has not had a significant impact on the rate of high energy events in PMT18. Instead the event rate is already increasing before the calibration run and continues to increase at a similar rate after the calibration run before slowly starting to rise more steeply around 50 days later.

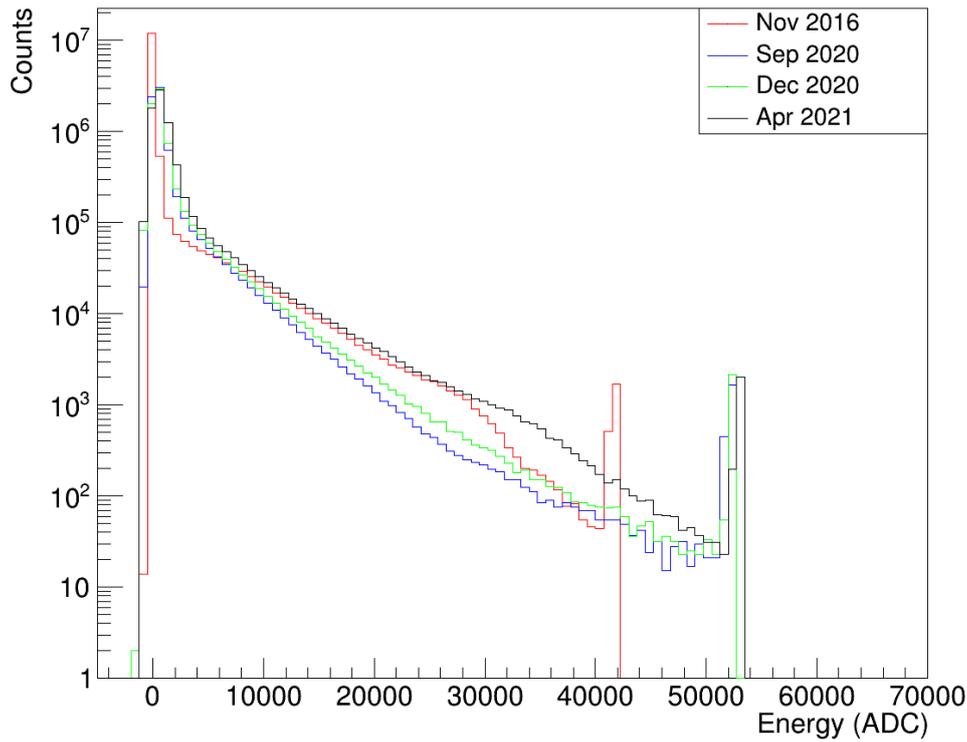


Figure 5.15: The energy spectra for 100 subruns of LS PMT18 data at different times in data taking, highlight the rapidly changing shape of the spectra after the end of December 2020, but not directly after the calibration run in October and November 2020.

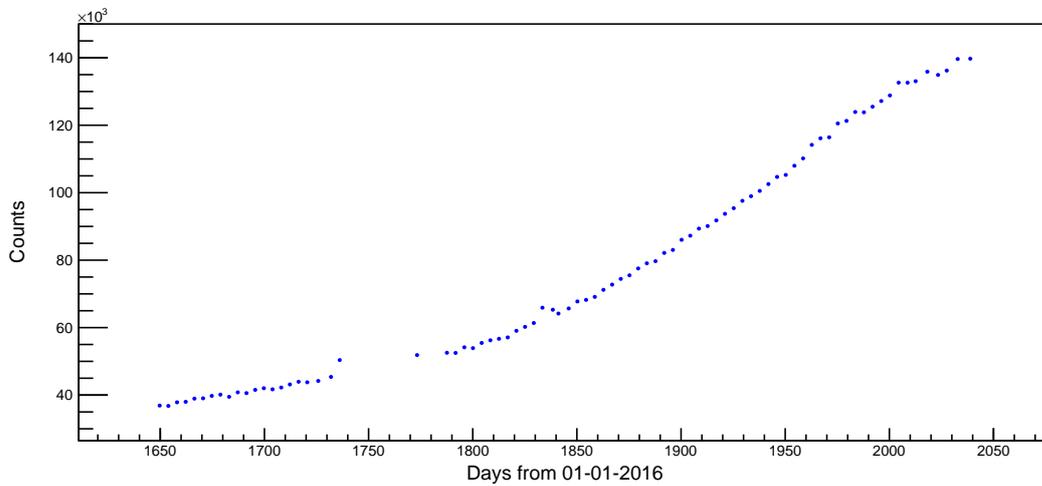


Figure 5.16: The number of events above 10000 ADC units in PMT18 in 50-subrun bins from July 2020 to August 2021. The gap in the data is the ^{22}Na calibration run and since the rate of change of the event rate appears unchanged before and after the calibration run, it can be inferred that the calibration run did not directly cause the increase in the change of rate at later times.

It is unclear what physical effect is driving these changes in behaviour in PMT18. Although it may be possible to salvage data from PMT18 up until the ^{22}Na calibration run, most likely through some process of raising the energy threshold of the PMT to some

conservative value to ensure no single-hit events in the crystal are incorrectly tagged as multi-hit events by PMT noise in PMT 18, it is less clear how the change in spectral shape and increase in rate after September 2020 might be resolved. This combined with the fact that the entire PMT18 energy spectrum is affected by this latter change in behaviour meant a decision was taken to remove signals from PMT18 entirely from the calculation of LS variables in order to ensure the purity of LS events. As such, PMT18's signals no longer affects the overall LS charge variable, the variable on which single/multi-hit decisions are made in the context of crystal/LS coincident events. It would perhaps have been possible to only remove PMT18 from later data, however this could potentially cause difficulties in interpreting single/multi-hit signals before and after the point at which the PMT's signals were removed. This combined with the fact that the functionality of the LS is not significantly affected by removing a single PMT's signals means removing the signals from PMT18 entirely is the safest choice.

5.2.3 LS energy calibration and threshold without PMT18

The removal of PMT18's signals from the overall energy spectrum of the LS necessitates a recalculation of the gain correction, energy calibration, and low energy threshold of the LS since the original work carried out to obtain these values were based on LS spectra that included PMT18. As discussed in Section 5.1.3, the gain correction process is independent of the energy calibration, since the process is done using ADC units. It is also independent of the low energy threshold, since the ^{40}K Compton edge that is fitted is comfortable above the low energy threshold. As such, the new gain correction without PMT18 is carried out before the energy calibration and threshold recalculation, as described previously in Section 5.1.3.

The initial process of energy calibration in the LS involved simulating and fitting LS data both for physics data at the start of data taking in October 2016 and a ^{60}Co calibration run the following December using Monte Carlo (MC) simulations. Then by comparing these simulations (which are in units of keV) with the real data (still in units of ADC), an ADC to keV calibration value can be obtained. Rather than a lengthy process of repeating these MC simulations and fitting processes after updating the simulation to remove PMT18, the assumption was made that each PMT contributed an equal number of counts to the overall LS spectrum (i.e. one eighteenth of the total counts) and as such the original energy calibration factor of 143.8 counts/keV could simply be reduced by a factor of 17/18, or approximately 6%. This was verified by plotting energy spectra, both with and without PMT18 for physics data at the beginning of taking as well as ^{60}Co calibration data from the calibration run in December 2016, then comparing the location of spectral features. In the case of the physics data the Compton edge from the 1461 keV γ from the decay of ^{40}K is fitted as described previously. In the case of the ^{60}Co calibration data, two Compton edges are fitted (as shown in Fig. 5.17), arising from the 1173 keV and 1332 keV γ s associated with the decay of ^{60}Co [73]. In all cases the Compton edges are fit using the phenomenological function described in Eq. 5.2. The results of these fits are summarised in Table. 5.1. The errors on the location of the fit value given by the fitter are of the order of 0.1% of the fitted value and as such are not significant in this estimation of energy calibration.

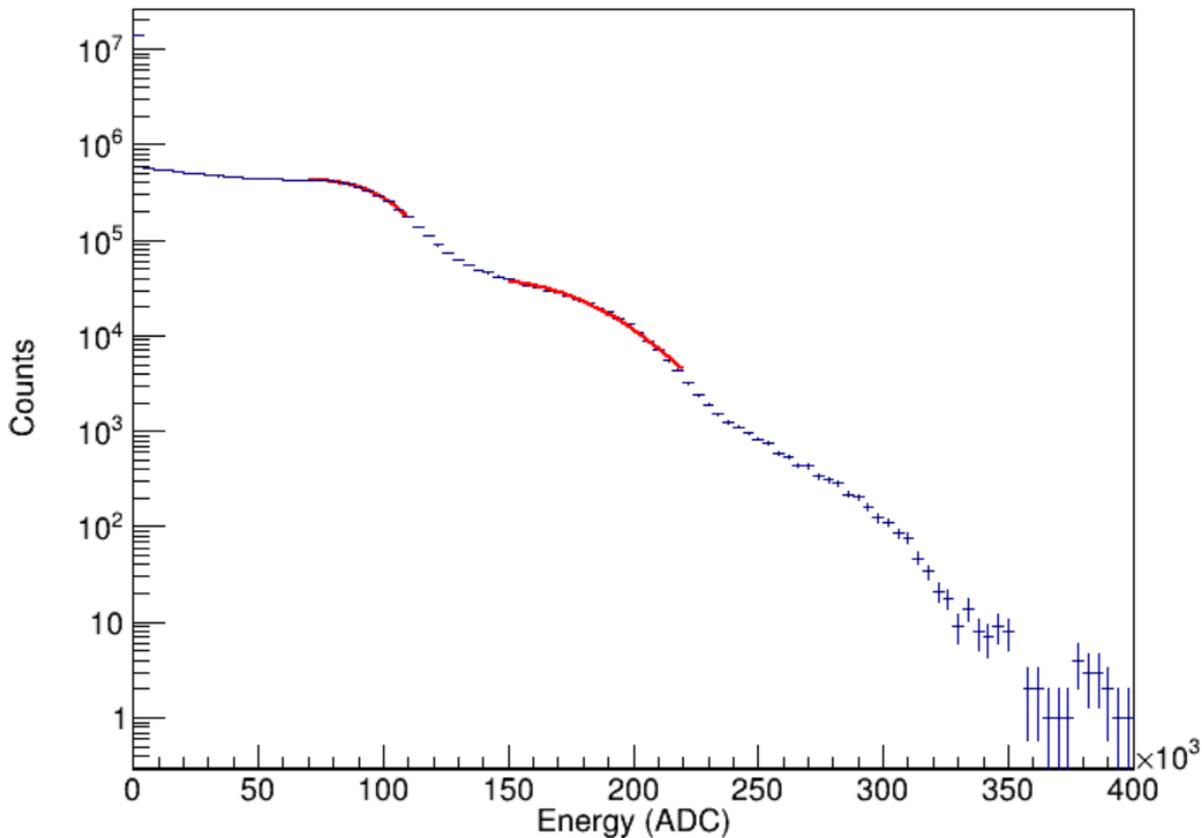


Figure 5.17: Energy spectrum for 100 hours of ^{60}Co calibration data in December 2016 with fits (using Eq. 5.2) to the Compton edges from the 1173 keV and 1332 keV γ s associated with ^{60}Co decay.

Spectral Feature	With PMT18 (ADC)	Without PMT18 (ADC)	Percentage Drop
^{40}K	2.09×10^5	1.97×10^5	5.8%
^{60}Co first edge	1.00×10^5	9.40×10^4	6.0%
^{60}Co second edge	1.75×10^5	1.71×10^5	2.3%

Table 5.1: The locations as determined by fitting of the 1461 keV Compton edge in physics data from ^{40}K decay, and the 1173 keV and 1332 keV Compton edges in ^{60}Co calibration data from ^{60}Co decay for LS spectra with and without PMT18.

The locations of the ^{40}K Compton edge in physics data and the 1173 keV ^{60}Co Compton edge in calibration data both fall by around 6% when PMT18 is removed from the LS energy spectrum as expected from the initial prediction. However, the second ^{60}Co edge from the 1332 keV γ only falls by around 2%. This is unlikely to be due to any non-linearity effects given this feature falls between the ^{40}K edge and the other ^{60}Co edge in energy.

A further complicating factor comes from the gain correction performed without PMT18 described in Section 5.1.3. Both the physics and calibration data used for the above calibration test do not have any gain correction applied to them. This is the case for the physics data as it is the first data after the start of data taking and therefore taken

as the reference value against which the rest of the data is calibrated for, and for the calibration data since it is also very early in the data taking period and because data from the LS is generally not used during calibration runs. However, the gain correction for the rest of the data is calculated based on fits to LS energy spectra in units of ADC, without any consideration of the potential change in calibration, and this spectrum restored to the reference value.

This means that in effect, the gain correction is automatically providing some or all of the re-calibration required after removing PMT18. Were it the case that the reference value used for the gain correction after removing PMT18 was directly equivalent (in terms of fit location and bin size) to that of the initial gain correction which included PMT18, then the ratio of reference values before and after the removal of PMT18 could be used to inform the change in energy calibration i.e. if the new reference value was around 6% lower than the old one then the re-calibration is effectively done automatically. However, the fit to the location of the Compton edge is non-trivially affected by choices of fit location and binning², and unfortunately, these choices were not saved for the initial reference value so a direct comparison in reference values is not possible.

Interpreting these factors correctly is non-trivial, and for an exact ADC to keV calibration value to be found, ideally the process of MC fitting to physics and calibration data should be repeated. However, it is relatively safe to assume that the energy calibration factor for the LS will be somewhere within the bounds of the original value (143.8 ADC/keV) plus or minus a few percent. This would be a potentially problematic conclusion for any studies that specifically required precise LS energies, fortunately however, the primary function of the LS is to tag events from radiogenic decays and as such veto any crystal events in coincidence with these LS events as multi-hit events. This process does not require precise ADC to keV calibration in the LS, and so long as the LS low energy threshold is set sufficiently high to avoid any chance of accidentally recording PMT noise as a physics signal, a small uncertainty in LS energy calibration should not have a significant effect on analysis involving the crystals, the main analysis performed in COSINE-100. Separately during the three years annual modulation search [1], as it was initially thought the gain correction described in Section 5.1.2 would not be applied in time, the LS low energy threshold was raised from 20 keV to 80 keV. Ultimately this analysis was delayed and the gain correction was applied, but the LS threshold was not lowered back to 20 keV in the analysis. This 80 keV low energy threshold in the LS is sufficiently high to avoid concerns about noise events leaking into the LS signal region and incorrectly vetoing crystal events and therefore is kept for the analysis described in this work.

5.3 Concluding remarks

The changing behaviour of the LS and subsequent attempts to restore its stability thereafter have been described in this chapter. Gain changes in the LS generally occur in a linear fashion, and by tracking the location of the Compton edge in the LS spectrum from the 1461 keV γ from ⁴⁰K decay, this gain change has been quantitatively measured and accounted for. The nature of the behaviour of the individual LS PMTs has also been studied and it has been shown that broadly they behave in an acceptable, if not entirely consistent fashion. An exception to this is PMT18 which shows unresolvable changes in

²Note that although the fit value is sensitive to these parameters, so long as the parameters are kept constant, the fit is stable.

its energy spectrum, introducing a large amount of noise into the LS spectrum and consequently its signals must be removed from the calculation of LS variables and subsequent event selection. This in turn required a re-examination of the ADC to keV energy calibration, and of the low energy threshold in the LS. Overall though, the LS has shown to be stable over the entirety of the first period of COSINE-100 data-taking from October 2016 to March 2022, after the corrections described in this Chapter have been applied. The LS with a low energy threshold of 80 keV is therefore utilised in single/multi-hit event selection throughout the remainder of this thesis.

Chapter 6

Investigation and Mitigation of Intermittent PMT Noise

At low energies, photomultiplier tube (PMT) noise events become the dominant source of background events at COSINE-100. There are two types of PMT noise which appear consistently throughout the data-taking period. These are known as “Type-I noise” and “Type-II noise”. Above 2 keV, Type-I noise is the dominant source of noise events in the detector and is relatively easy to separate from signal events at these energies due its characteristically sharp decay time. Below 2 keV, Type-II noise takes over as the dominant source of noise events. As a result of Type-II noise events being closer in pulse-shape to signal events than Type-I noise in combination with fewer photo-electrons (PEs) being present in events at low energies, it is a more complex process to separate Type-II noise from signal events. Multivariate analysis methods in the form of boosted decision tree (BDT) variables have been developed to separate signal events from these types of noise, first focusing on Type-I noise at a 2 keV low energy threshold [2], then later on Type-II noise allowing for a 1 keV threshold to be reached [105]. These are also described in more detail in Section 9.2.

A third type of PMT noise is also found at COSINE-100, but unlike Type-I noise and Type-II noise, which have roughly constant rates over time, this type of noise only appears intermittently in most of the NaI crystal PMTs. This intermittent noise is known as “Gaussian noise” due to the similarity of its pulse-shape to a Gaussian distribution. Gaussian noise is more similar to signal events than the other types of noise when characterised by the variables used in the training of the main BDT, making it harder to tag and reject. For a four week period in April and May 2020, a significant increase in the event-rate in C2 (the eight NaI crystals are labelled “C1”, “C2” etc) was measured. The cause of this was studied and found to be an increase of Gaussian noise events.

The initial identification of Gaussian noise in early COSINE-100 data is described in Section 6.1 with reference to an initial attempt at separating the events from signal. Then a study of the new Gaussian noise period is presented in Section 6.2.

6.1 Initial identification and rejection of Gaussian noise in early data

The presence of this Gaussian noise was first observed in C1 at the start of data-taking, after an unexpectedly high event-rate was observed in the crystal. The majority of noise events measured in C1 had a different pulse-shape to noise events in other crystals and

therefore could not be removed by the event selection used for other crystals. Examples of typical waveforms for signal, Type-I noise, and Gaussian noise (referred to as “discharge” noise in the figure) are shown in Fig. 6.1. Gaussian noise appears in C1 throughout the data-taking period, and appears intermittently in other crystals.

Both signal and Type-I noise are characterised by a fast rise-time in measured counts in the PMT, but unlike in signal events which typically the count-rate then drops relatively slowly over a few hundred ns, in Type-I noise, the count rate typically drops down to the pedestal value within 200 ns. Type-II noise (described in more detail in Section 9.2.1) has a pulse-shape similar in shape to signal and Type-I noise events with a similarly sharp rise-time, but with an intermediate decay thereafter. Gaussian noise on the other hand has a slower rise-time than the other types of events which then decays away roughly symmetrically. The shape of the decrease in counts of Gaussian noise is similar to that of signal events.

There are multiple potential origins for noise events generated in a PMT, for example from signals from thermoelectrons, current leakage or any other process that produces an electron in the PMT circuitry (“dark current”), and from after-pulses from larger events [73, 103]. It is therefore reasonable, given the different characteristic pulse-shapes of the different types of noise, shown for Type-I noise and Gaussian noise in comparison with the pulse-shape from signal, to assume the three types of PMT noise have different sources and should be dealt with independently.

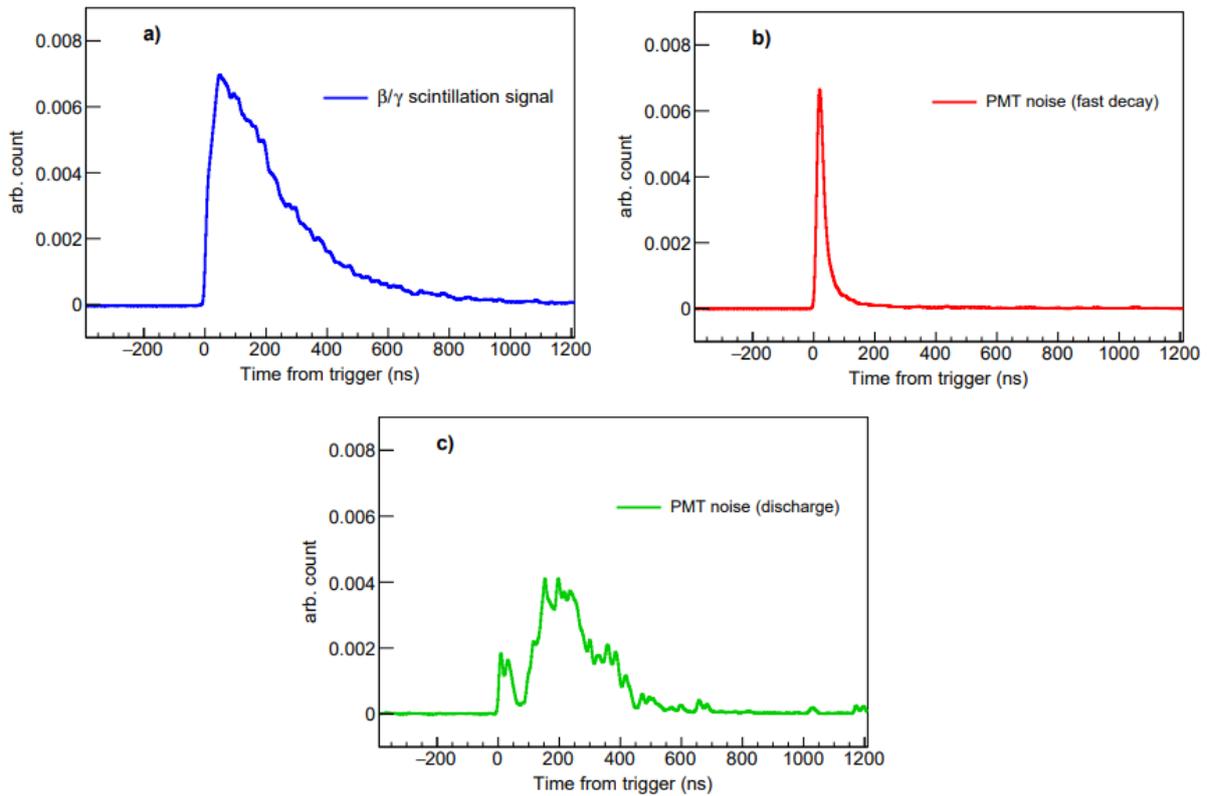


Figure 6.1: Typical waveforms of signal events from a calibration run (a), Type-I noise PMT noise events (b), and Gaussian noise events (c). Adapted from [75].

The initial event selection used to discriminate between signal and Type-I noise events was effective at separating these two types of event, but not at separating Gaussian noise

from signal events. Type-II noise events were not yet considered due to predominantly occurring at energies below the low energy threshold of 2 keV, used at the time. The BDT variable used to veto Type-I noise events used input parameters largely designed to characterise the rate at which the charge of an event decayed over time, since this is the main difference in pulse-shape between signal and Type-I noise. For example, the “fast” and “slow charge (“X2” and “X1”), the fraction of the total charge measured in the first 50 ns from the initial pulse of the event ($t = 0$ in Fig. 6.1), and from 100 ns to the end of the event respectively, where each event is 600 ns long starting from $t = 0$. This parameter is also one of the main parameters used for noise rejection at the DAMA experiments [69]. All of the parameters used in this BDT are described in full in Section 9.2. Consequently, the BDT was ineffective at rejecting Gaussian noise since the decay time of Gaussian noise is similar to signal events.

To address this, a new BDT variable, known as the “BDTA” variable, was developed using input parameters designed to more effectively separate signal and Gaussian noise events. Specifically, the two new input parameters with the greatest separation power for signal and Gaussian noise events were parameters based on the time at which 95% of the total event charge has been accumulated (the “CAT95” parameter), and the variance of the charge-weighted mean time in the first 1000 ns of the event from the first pulse (the “*nvt* parameter”) [106, 107]. The *nvt* parameter is defined as

$$nvt = \frac{\sum_{i=0}^T q_i t_i^2}{\sum_{i=0}^T q_i} - \left(\frac{\sum_{i=0}^{T_s} q_i t_i}{\sum_{i=0}^T q_i} \right)^2, \quad (6.1)$$

where $T = 1000\text{ns}$, the index i is the binning of the event in time starting at the time of the first pulse, q is the charge, and t is the time.

The BDTA variable was trained on a crystal-by-crystal basis using a sample of the first 200 hours of data from the beginning of data-taking, selected using the original BDT variable as the signal sample. The noise sample was also selected using the original BDT variable, but events were selected from a sample of 268 hours of data with high Gaussian noise [106]. After the training, a cut based on the BDTA variable was generated.

While this BDTA variable was able to remove some Gaussian noise events from samples of signal data, it was clear that in regions of high Gaussian noise, signal samples were still contaminated. This was evident as the rate of events passing the event selection with the BDTA cut applied was still elevated in these regions and the event-rate in C1 remained volatile at all times. Furthermore, the separation power of the BDTA variable decreases as a function of energy as the waveforms of events contain fewer PEs and have a less clear shape. Consequently, regions of known high Gaussian noise, including all data from C1, was excluded from physics analyses. The event selection in early COSINE-100 analyses [2, 75] utilised a cut based on the BDTA variable, but one chosen with a relatively high selection efficiency to remove obvious Gaussian noise events that might have been missed in event selection from the original BDT.

6.2 Recurrence of Gaussian noise

The initial work on the identification and rejection of Gaussian noise took place in the early stages of data-taking at COSINE-100. At this point, two prolonged periods of elevated Gaussian noise had been observed outside of C1 and removed from data-sets used in dark matter searches. These were the first 59.5 days of data in C2 (starting on the 20th of October 2016), and from the 12th of January to the 28th of March 2017 in C7 [75]. Then

followed a prolonged period of stability within which no periods of elevated noise were identified. This lasted until an approximately three-week-long period between April and May 2020 where the event-rate below 50 keV in C2 increased significantly compared to previous periods. The event-rate in the other crystals was unaffected.

6.2.1 Investigating a new period of high event rate

An example of the evidence of high event rate in the period is shown in Fig. 6.2. An energy spectrum of single-hit events from four weeks of data centred on this period is shown in comparison with a period of equal length prior to the period of increased rate. Both spectra are also shown in comparison with a reference spectrum from earlier in the data-taking period. A clear excess of events above the reference in the period being investigated can be seen. These spectra are examples of plots checked by collaboration members on weekly monitoring shifts and therefore were some of the first indications of the period of increased noise. No noise rejection event selection is applied in these spectra.

Also as shown in the standard monitoring plots is the low-energy (below 20 keV) energy spectra in each crystal where noise rejection event selection has been applied, shown in Fig. 6.3. This is effectively the energy spectrum for the events used in dark matter searches. As in Fig. 6.2 there is an excess of events in the period being investigated, hence the events causing the increased event rate are not all being removed by the event selection in the signal region of interest (ROI). Instead a clear excess of events above the reference spectra can be seen for events below approximately 6 keV, directly coinciding with the main 1-6 keV ROI used in annual modulation searches.

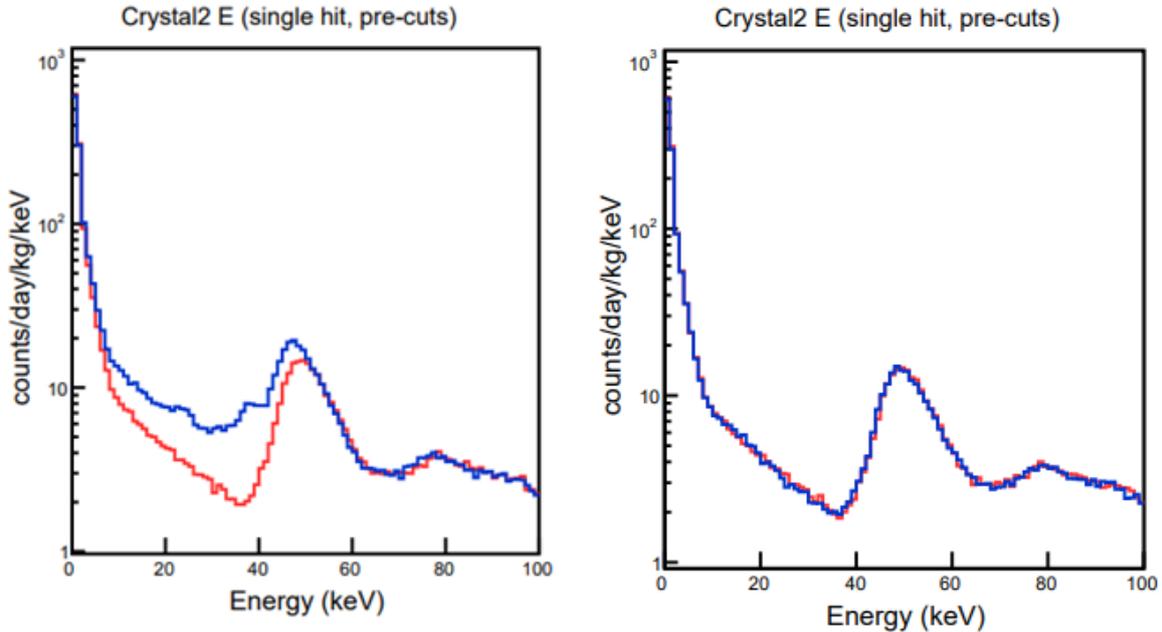


Figure 6.2: Energy spectra of single-hit events in C2, prior to any other event selection (blue spectra) for a four week period during a period of high Gaussian noise (left) and a stable period (right). The red spectra are reference plots from earlier in the data-taking period. These plots are taken from the plots generated for the continuous monitoring of COSINE-100.

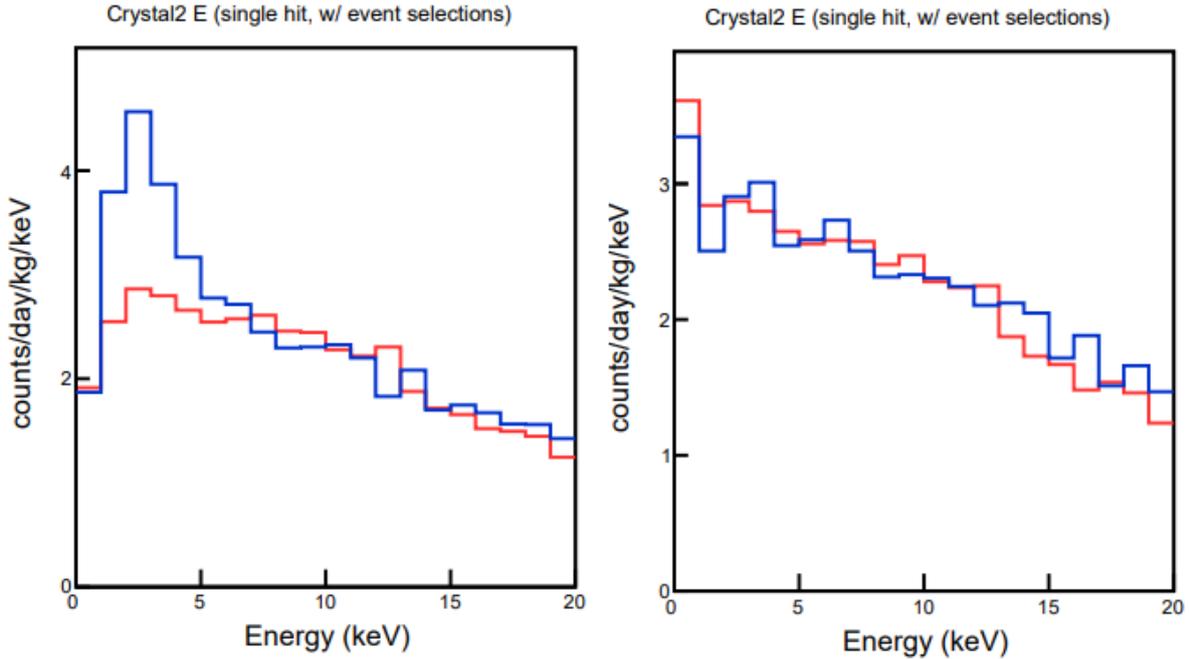


Figure 6.3: Energy spectra from the same data-sets as in Fig. 6.2, shown with noise rejection event selection applied and at a lower energies. A clear excess of events at low energies in the high-rate region (left plot) can be seen. These plots are also taken from the standard monitoring plots.

6.2.2 Properties of the new high-rate events

To determine if this period of increased event-rate was caused by a recurrence of Gaussian noise events, a 2D histogram of the distribution of BDT and BDTA values was generated for 200 hours of data in the region and then compared to equal length periods in stable regions, and in previous regions of high Gaussian noise in C2. These distributions are shown in Fig. 6.4. The BDT variable used in this case is the updated BDT variable which allows for a 1 keV low energy threshold (described in more detail in Section 9.2 and [108]). This is the same event selection used to select single-hit events in dark matter searches at COSINE-100 beyond 2019. In this distribution, events between 1-10 keV were selected with all the standard noise rejection cuts (described in more detail in Section 9.2) applied, with the exception of the BDT cut. The degree to which the excess events are leaking into the signal region is therefore shown.

In Fig. 6.4, the islands of events in the top right of the distributions are the signal-like events, separated from the noise-like events in the bottom half of the distribution by the BDT variable. In the upper left regions of distributions for the period of elevated event-rate and known Gaussian noise, an excess of events in the signal region of BDT parameter space can be seen. It is therefore clear that the observed period of increased event-rate does correspond to a period of Gaussian noise which is not separated from signal events using the updated BDT variable.

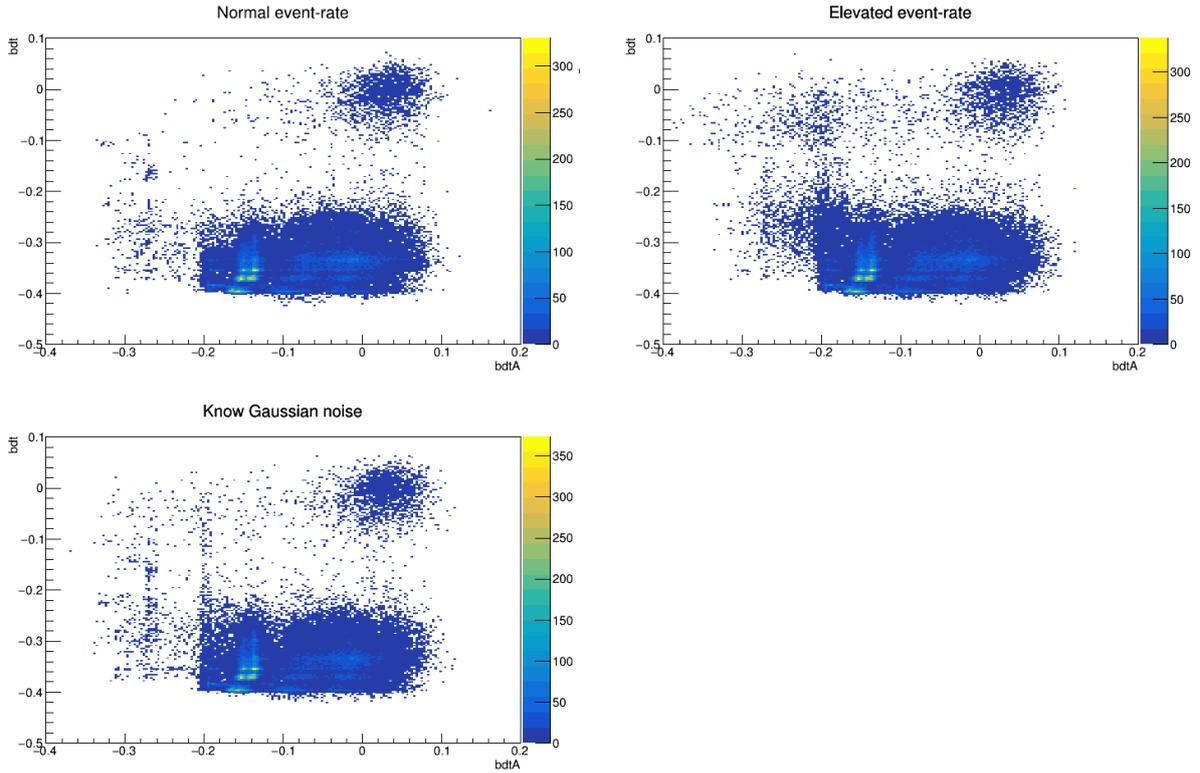


Figure 6.4: 2D distributions of BDT and BDTA values for single-hit events passing signal-like event selection criteria, aside from a BDT cut, in a region of normal event-rate, the region of high event-rate described in the text, and a region of known Gaussian noise. Signal-like events are concentrated in the upper right of the distribution whereas Gaussian noise appears in the upper left.

Applying the standard noise rejection cuts from the BDT variable (roughly speaking, this selects events in the top half of the distributions in Fig. 6.4) and plotting the distribution of event energy and BDTA score, as in Fig. 6.5, shows the energies and BDTA scores of events passing the standard noise rejection event selection. In this case, unlike in previous regions of high Gaussian noise where almost all of the Gaussian noise events were concentrated in energies below 2 keV in C2, the events causing the elevated event-rate in the period studied here are evenly distributed between 1 keV and 6 keV.

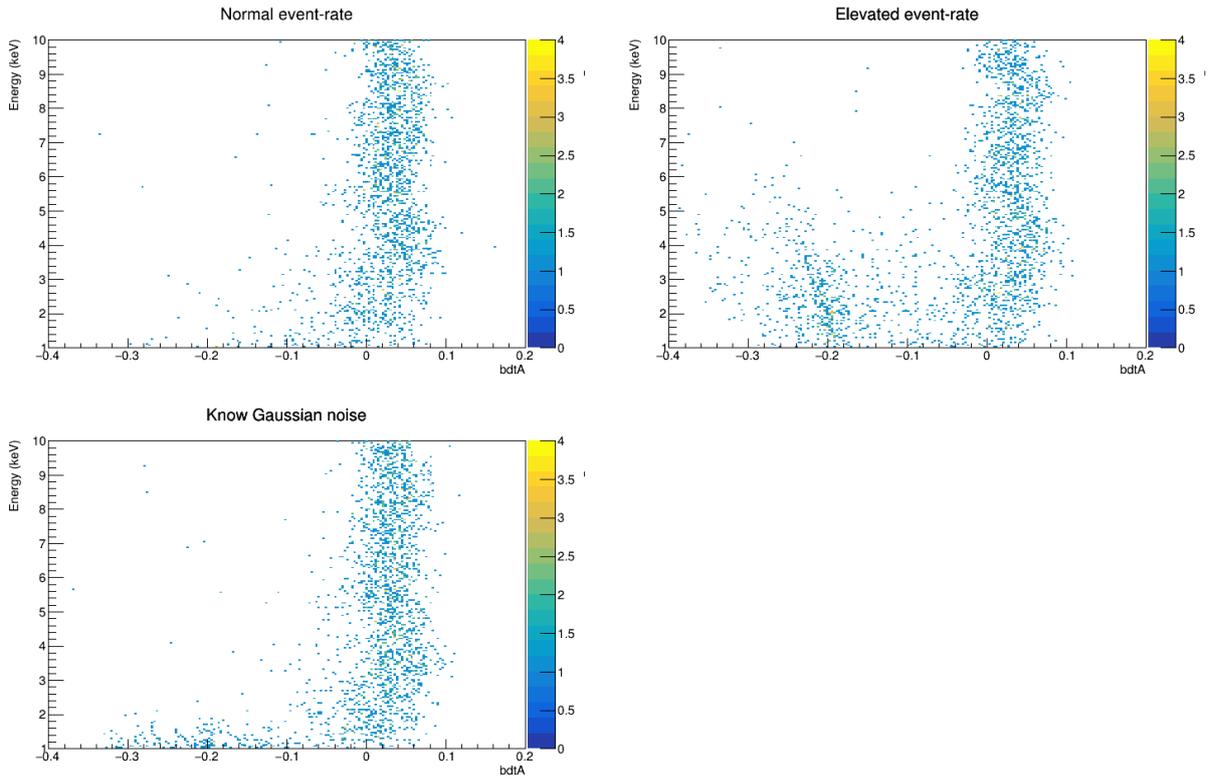


Figure 6.5: 2D distributions of energy and BDTA values for the same data-sets as in Fig. 6.4 with the standard noise rejection BDT cuts applied.

Using the distribution of events observed in Fig. 6.4, event selection in BDT/BDTA space can be chosen to quantify the amount of Gaussian noise present in a given data-set. Specifically, events with BDT scores above -0.14 , and BDTA scores below -0.04 were selected from events with the above single-hit event selection in 10 subrun (20 hour) long bins, and the number of events in this region were counted and converted to units of counts/day/kg/keV (dru).

The rate of candidate Gaussian noise-like events selected using the above cuts are shown for two full runs of COSINE-100 data in Fig. 6.6. The first, run 1873, contains the period of high event-rate being investigated, the other, run 1719, is a period of stable event rate which covers from 18/04/2018 to 10/06/2018. A rise and fall in the rate of Gaussian noise-like events in run 1873 from approximately subrun 100 to subrun 360 can be seen where the rate of such events is several times higher than the normal event-rate, seen either at later times in the run, or in the stable run shown below. The error bars are from the propagated uncertainties calculated from the number of events in the bin.

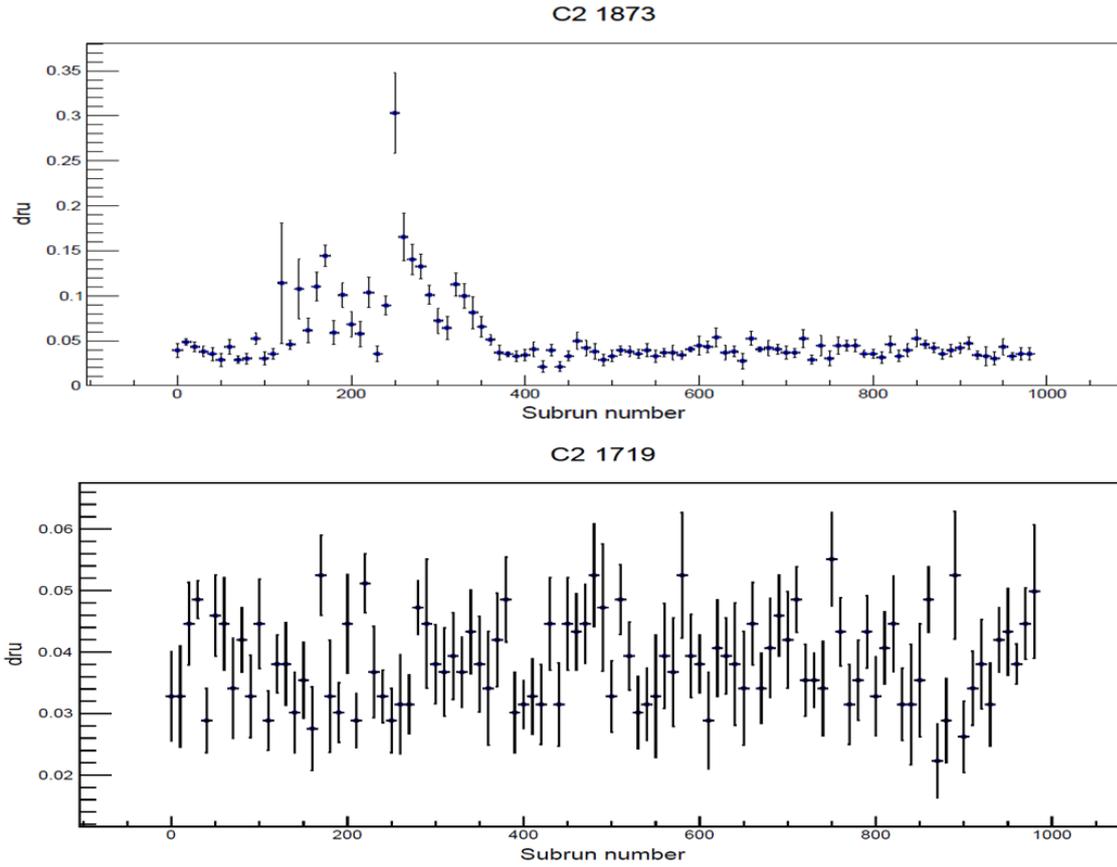


Figure 6.6: The rate of Gaussian noise candidate events in C2 in a period of high event-rate (top, run 1873), and stable detector operation (bottom, run 1719). Events are selected from single-hit events passing obvious noise rejection cuts, and then using the BDT/BDTA box cuts described in the text.

The hypothesis that these excess events were Gaussian noise was verified by visually inspecting the waveforms of a sample of the selected events. Fig. 6.7 shows a typical waveform in the Gaussian noise-like region, compared to a signal-like event in a stable period of data-taking. A clear difference in the waveform shapes can be observed, where the signal-like event exhibits the typical fast rise-time and slow decay, whereas the Gaussian noise-like event has a roughly symmetrical shape.

In Fig. 6.7 waveforms of events that are visually clearly either from signal or Gaussian noise are shown. However, as can be seen in Fig. 6.4, some events have an intermediate value BDTA where the signal-like events merge into the Gaussian noise-like events. Even in regions of normal event-rate, some events with BDT and BDTA values typical of Gaussian noise are present. This is because even in regions of normal event-rate, a low, constant rate of Gaussian noise-like events is observed, for example as shown in Fig. 6.6. The rate of these events is consistently around 0.04 dru in C2 during regions of stable detector operation. Examining the waveforms of samples of these events to try and determine if they are Gaussian noise or signal does not provide a clear solution, since many of the waveforms are visually similar to both types of event in some way, for example as in Fig. 6.8 where in one of the PMT the event is more signal-like and in the other it is more noise-like. The true provenance of these events is therefore unclear.

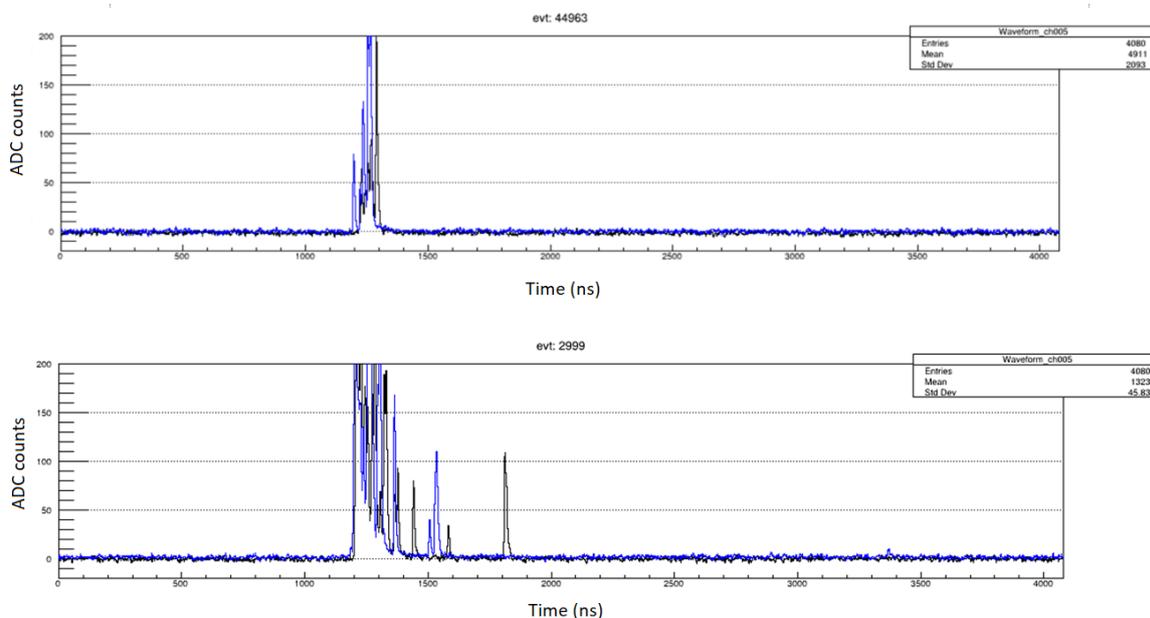


Figure 6.7: A typical waveform of an event meeting the Gaussian noise-like criteria in run 1873 (top), compared to signal-like event (bottom). The blue and black curves represent the signals from each of the crystal PMTs.

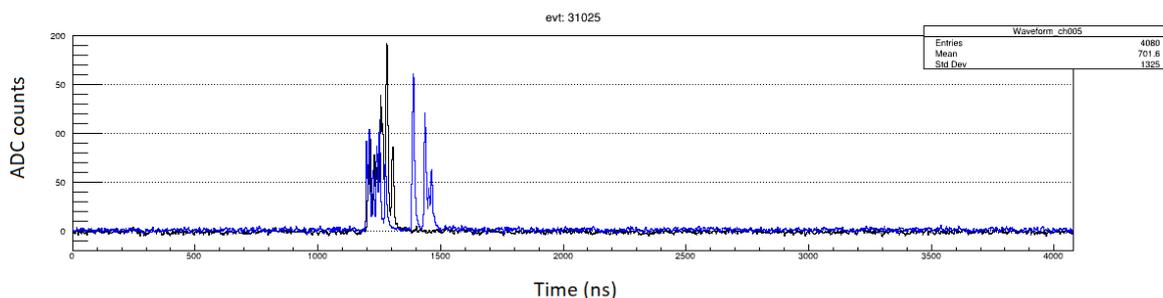


Figure 6.8: An example of an event with an unclear provenance.

To further investigate whether the new Gaussian noise-like events measured in C2 differed in behaviour from previous examples, particularly with a potential goal of improving the BDTA variable in mind, events passing signal event selection criteria as in Fig. 6.5 were also plotted in 2D distributions with energy and the CAT95 and nvt parameters described in Section 6.1. This is shown in Fig. 6.9 and Fig. 6.10. The CAT95 parameter is the time taken from the first pulse in the event for 95% for the total charge of the event to be accumulated and the nvt parameter is a measure of the variance of the charge-weighted mean-time of the waveform, as defined in Eq. 6.1. Note that in the BDTA training the CAT95 and nvt parameters are calculated for each of the two crystal PMTs, whereas in Fig. 6.9 and Fig. 6.10 an average value for the two PMTs is shown.

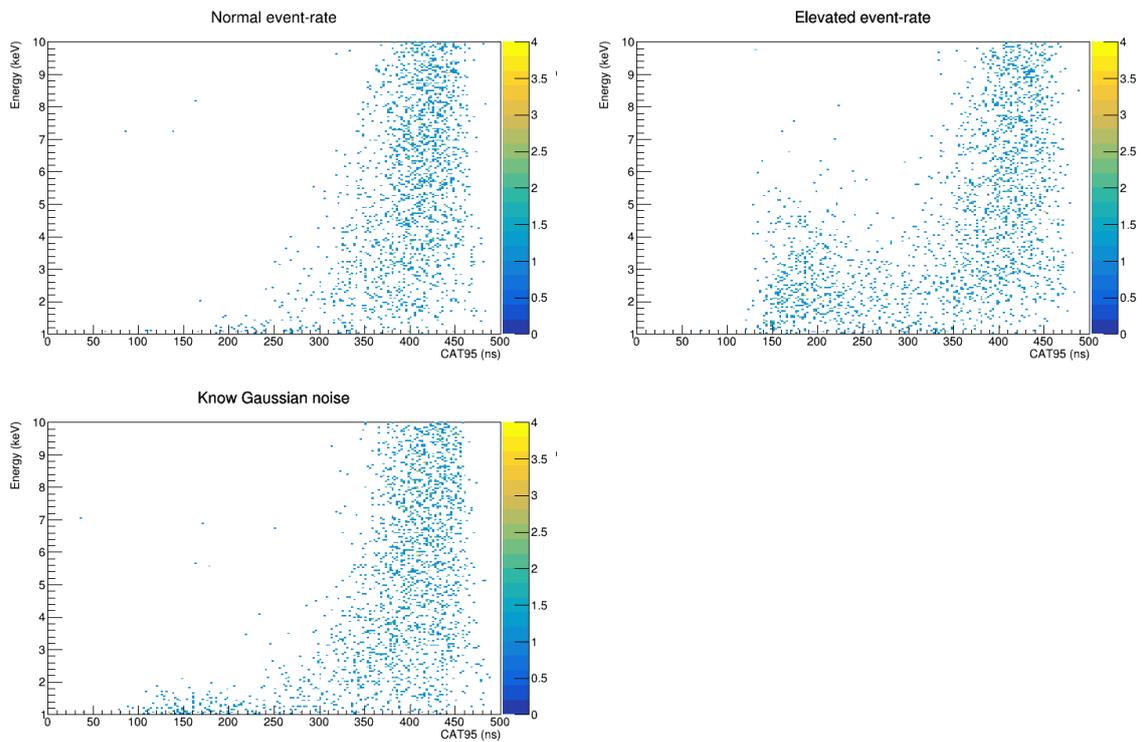


Figure 6.9: 2D distribution in energy and the CAT95 variable for the same samples of signal data shown in Fig. 6.5. Signal events are located in the band of events in the right of the distributions.

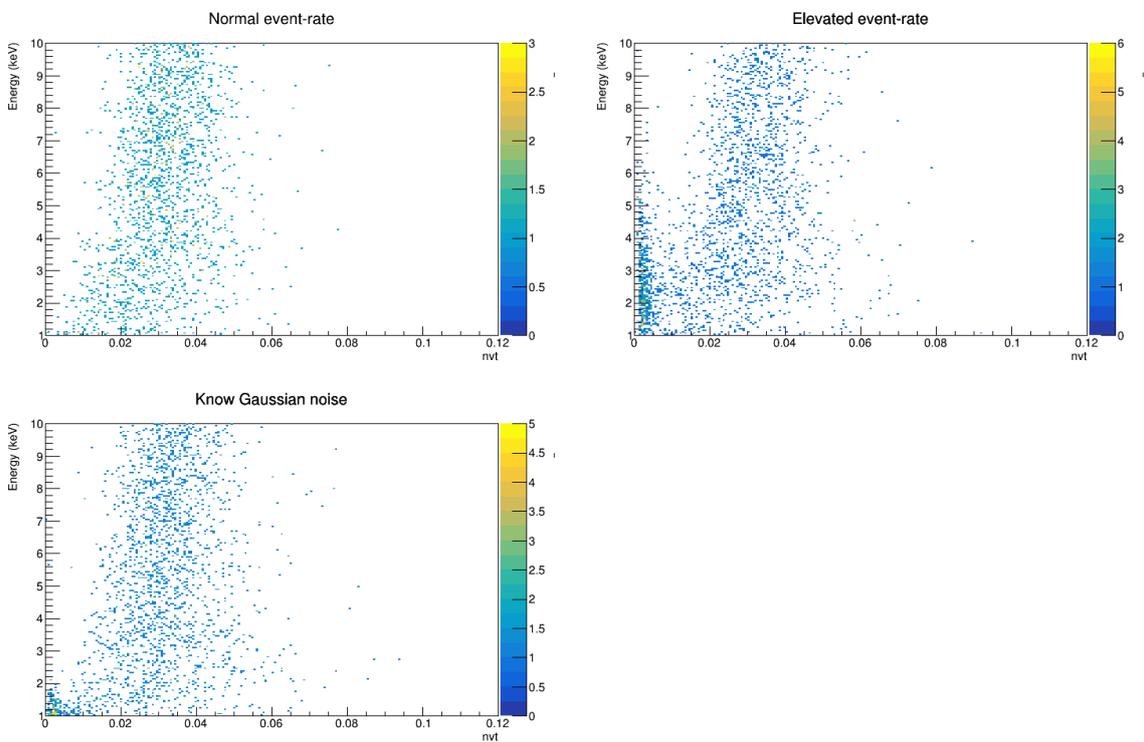


Figure 6.10: As above, but for the nvt parameter. Here signal events are the wide band of events between 0.02 and 0.04 nvt values.

In the distributions showing the CAT95 variable, the true signal events are mostly

found in the band of events above approximately 350 ns as shown by the distribution of events in the normal event-rate period, whereas the Gaussian noise events are generally spread across lower values. This is an expected result, given the shapes of the waveforms shown in Fig. 6.1 and Fig. 6.7 where the signal events tend to last longer in time than the Gaussian noise events.

In the nvt variable distribution, the Gaussian noise events are generally found to have a low nvt score and signal events a higher nvt score. Again, this is expected from the definition of the parameter and the known shapes of the different types of events. As expected from Fig. 6.5, in the more recent period of Gaussian noise events there is a much wider spread of Gaussian noise event energy than in the previous period. In both cases as also seen in Fig. 6.4 and Fig. 6.5, there is a significant overlap between regions of signal and noise events, highlighting the difficulty in separating Gaussian noise from signal.

Work to improve the separation power of the BDTA was started using a sample of Gaussian noise events from run 1873, selected using the described above. The aim being that the new sample of Gaussian noise events, combined with the greater selection power of the new regular BDT variable to help select signal and noise events would allow for the BDTA variable would be able to improved such that regions of high Gaussian noise would no longer have to be removed from dark matter searches. However, this study did not produce such a variable and regions of high Gaussian noise remain unusable. In the selection of data-sets for dark matter searches, a data-quality check is performed to identify potentially problematic regions. This process, described in Section 9.2.2, does not use the BDTA variable or the procedures described in this chapter. However, the event selection described in the Chapter is used after the data quality checking process is complete as a final test of the SET4 data-set.

6.3 Concluding remarks

A period of elevated event-rate in C2 in April and May 2020 was investigated and found to be caused by a re-occurrence of intermittent ‘‘Gaussian’’ noise that had previously been observed in periods of C2 and C7 data. These Gaussian noise events were studied to gather more information about the properties of these events. Using event selection mainly based on the already developed ‘‘BDT’’ variable (mainly for rejecting other types of PMT noise), and ‘‘BDTA’’ variable (an early attempt at a parameter to reject Gaussian noise), a sample of the Gaussian noise events causing the elevated event-rate were selected and studied. It was initially hoped that in selecting a sample of Gaussian noise events in this way, an updated BDTA variable could be developed which, unlike the original BDTA variable, had sufficient separation power to fully reject Gaussian noise. This would allow for regions of high Gaussian noise to still be used in dark matter searches, rather than having to be discarded, as in previous searches. However, limited progress was made in this study and the region of high Gaussian noise ultimately had to be rejected from physics searches in this analysis. The event selection described in Section 6.2.2, used to select Gaussian noise events, was also applied in the process of selecting the SET4 data-set, as described in Section 9.2, to validate the data quality checks carried out.

Chapter 7

Development of the Third Generation COSINE-100 Annual Modulation Fitting Procedure

The primary goal of the COSINE-100 experiment is to provide a model independent test of the positive annual modulation signal from the DAMA experiments which they attribute to signals from dark matter [91,92] (see Section 3.2 for more details of the DAMA result). In order for the analysis of COSINE-100 data to be a truly model independent test of the DAMA results, the main principles of the data collection and analysis, i.e. searching for an annual modulation in single-hit events in NaI(Tl) crystals, must be functionally identical to those used in the DAMA experiments. However, once the data has been collected, the process of data analysis need not exactly follow that used by DAMA, provided no assumptions are made about the nature of dark matter itself. Indeed at COSINE-100 a different approach is taken to estimate the degree of any annual modulation compared to that taken by DAMA. A separate study where DAMA's procedure were followed as closely as possible when analysing COSINE-100 data was also carried out to investigate whether DAMA's analysis techniques had any impact on the results found. This is described in more detail in Section 3.3.2 and in [5].

The primary method of searching for annual modulation at COSINE-100 has evolved from one analysis to the next, and between the second-generation, 2.8-year analysis ("SET3") and third-generation, 4.7-year analysis ("SET4", this work) the process was improved in light of identifying two issues with the SET3 procedure. Fortunately, the SET3 annual modulation results remain valid in light of these issues; nevertheless, the requisite changes to the annual modulation fitting procedure represent a significant upgrade.

In this chapter, this analysis procedure for the third generation annual modulation analysis at COSINE-100 is described in detail. In Section 7.1, the SET3 analysis method is described, since it provides the framework for the SET4 analysis. Thereafter, the two issues with the SET3 analysis and their subsequent solutions are described. Specifically, in Section 7.2.1 the alterations to the process of scaling the data/model to account for detector live-time being below 100 % are described, and then on the updated treatment of ^{210}Pb in the model in light of issues uncovered at the end of the SET3 analysis is described in Section 7.2.2.

7.1 Overview of the second generation fitting procedure

At the DAMA experiments, single-hit data in a given energy range is collected in 15 day bins based on the various event selection procedures of a given experiment. Residual event-rates are then calculated by subtracting the calculated background rate in roughly yearly cycles, starting in September, to give the final rate in a given crystal over time [91, 92]. These values are then summed for all crystals and this is the resultant data-set within which dark matter signals are searched for by way of fitting an annual modulation function to the data. The potential pitfalls of this method, first reported in [88] and shown quantitatively in COSINE-100 data in [5], are described in Section 3.3.2 and so given these potential issues, a different analysis method is chosen at COSINE-100.

7.1.1 SET3 data preparation

While much of the DAMA method of searching for annual modulation is also used at COSINE-100, specifically using the same energy binning (1-6 and 2-6 keV) and time binning (15 days), unlike the DAMA analyses, the main results of annual modulation searches at COSINE do not come from the calculation and fitting of a residual event-rate. Instead, a full energy-dependent background model for each crystal at COSINE-100 was developed, based on study of the background event-rate in the first 1.7 years of data-taking [74]. The full single-hit data-set is fitted using this background model, with background components treated as nuisance parameters. The fitting at COSINE-100 takes place at a crystal-by-crystal level, before the data is recombined across the detector as in the DAMA method. In this process the modulation components are kept fixed across each crystal while the background components are allowed to vary, since each crystal has its own unique makeup in terms of radiogenic backgrounds but the modulation component from dark matter is expected to be constant across each crystal, once normalised for crystal mass.

The initial data-sets for the SET3 modulation analysis at COSINE-100 are all events passing the single-hit event selection (described in full in Section 9.2) between 1 and 20 keV in the five crystals with acceptable light yields and noise levels (C2, C3, C4, C6, and C7). These data-sets are then split into 0.25 keV wide windows to match the binning of the energy efficiency correction (see Fig. 9.10 in Section 9.2.1 for more information). Each window is then divided into 15 day long bins where the uncertainty on the number of events per bin is approximated as the square root of the number of entries. The number of events and its associated uncertainty are then corrected by scaling the number of events by the inverse of that energy's event selection efficiency. The uncertainties in the values for the event selection efficiency had only a negligible affect on the event-rate so were not considered. These 0.25 wide keV windows were then recombined for each crystal in either 1-6 keV and 2-6 keV regions, or 20 1 keV-wide windows depending on the analysis being carried out.

The rates and uncertainties of individual time bins were then scaled a second time, this time to account for the fact a given 15 day bin may not have 100% live-time. This was achieved by calculating the total live-time of each time bin and scaling the energy efficiency-corrected number of events by the inverse of the fraction of the bin's true live-time over the ideal 15 days live-time. Dead-time may be caused by detector wide shut-downs, either for very short periods of time (e.g. due to a muon event) or longer periods of time (e.g. a DAQ issue or calibration run), or on a crystal-by-crystal basis, for example

as a result of data from a specific crystal failing the data quality check (described in full in Section 9.2.2) for a period of time. Finally the event-rate in each bin is converted from counts to daily rate units (dru), i.e. counts/(day·keV·kg), assuming a full 15 days of live-time in addition to the predefined energy window and the known crystal mass.

The result of this process, shown for SET3 data in Fig. 7.1, is a collection of data-sets representing the time dependent rate R of signal-like events in a given energy range in 15 day bins for the i^{th} crystal. Data points with larger uncertainties correspond to time bins with fewer raw counts. The gaps in C2 and C7 correspond to regions of high noise in the respective crystals and these bins are omitted from the modulation analysis. The general trend of decreasing event-rates over time is a result of the majority of the selected events coming from radioactive backgrounds with relatively short half-lives. It is these data-sets shown in Fig. 7.1 which are modelled by the following function with an annual modulation component

$$R^i(t | S_m, \alpha^i, \beta^i) = \alpha^i + \sum_{k=1}^{N_{bkgd}} \beta_{0,k}^i e^{-\lambda_k t} + S_m \cos(\omega(t - t_0)). \quad (7.1)$$

The arguments of this equation can be separated into two parts. One pertaining to the dark matter-induced annual modulation component with the amplitude S_m , phase t_0 , and period $T = 365.25$ days, where $\omega = 2\pi/T$, the other describes the background components of a given crystal. Specifically α^i is the constant background rate in the i^{th} crystal and $\beta_{0,k}^i$ is the rate at time $t = 0$ (defined here as midnight on the first of January 2016) in the i^{th} crystal of the k^{th} background component (of which there are eight) with the decay constant¹ λ_k . The annually modulating components are fixed per unit mass across all crystals since any dark matter component measured at COSINE-100 is expected to be constant (per unit mass) across all crystals. On the other hand the background event-rate of a given crystal is independent of other crystals and as such, background event-rates are allowed to vary across the crystals. The origin of the background components in this model and the method for dealing with the many nuisance parameters that arise from including these backgrounds in the model to recover the desired best-fit parameters associated with the annual modulation component will now be described.

¹Note that decay constants rather than half-lives are used in the annual modulation fit model at all times. However, for the sake of clarity and since decay-constant and half-life are inversely proportional, the decays constant will often be described in terms of half-life in this work.

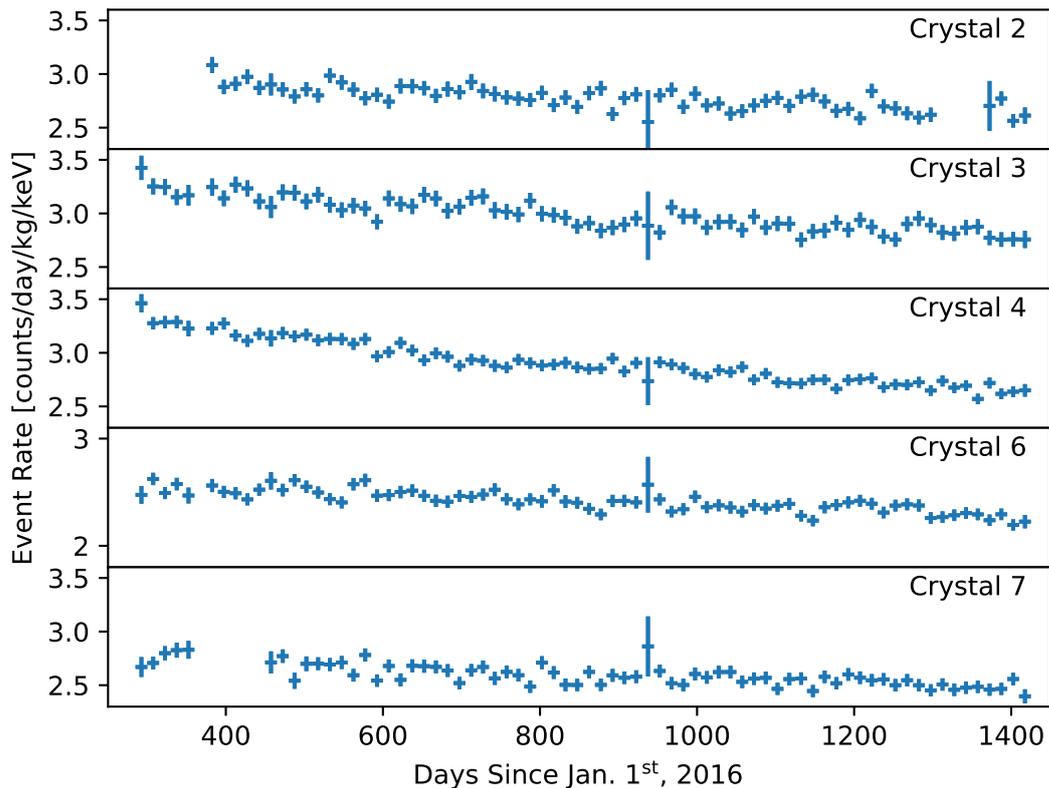


Figure 7.1: The rate vs time data in 15 day bins for the 1-6 keV single-hit data presented in the annual modulation search with three years of COSINE-100 data [1]. The data are scaled to give the event-rate assuming 100% live-time in each 15 day bin. The gaps in C2 and C7 correspond to regions of high noise in the crystal, resulting in those regions not being used in the analysis.

7.1.2 The background model used in the three year annual modulation search

In order for COSINE-100 to fully verify or refute the DAMA dark matter signal, the time-dependent background spectrum must be fully understood. A full understanding of the background event-rate also removes the need to calculate residuals and hence any potential issues arising from such a calculation are avoided. To that end, Monte Carlo (MC) simulations of the time-dependent energy spectrum using the GEANT4 toolkit were carried out and fit to the first 1.7 years of COSINE-100 data with a 1 keV low energy threshold to build a full time-dependent model of the background energy spectrum [74].

This background model represents a significant improvement on that used in dark matter searches prior to the search with SET3 data [1]. Specifically, in the first annual modulation search [75], all the background components were modelled as a single exponential for each crystal. Given the background energy spectrum of a given crystal at COSINE-100 will have significant contributions from many different radioisotopes with a wide range of half-lives, this single exponential model is a somewhat limited description of the background rate over time.

To address this in the new background model, all radioisotopes with half-lives less

than one million years were selected for modelling. Isotopes with half-lives greater than one million years, or in decay chains in secular equilibrium with parent isotopes with half-lives of over one million years were collected for each crystal and treated as a single flat/constant component which appears as the α^i term in Eq. 7.1. The components consequently considered are shown along with their half lives in Table 7.1.

Component	Half life
^{210}Pb	22.20 years
^{60}Co	5.27 years
^{125}I	59.4 days
^{121}Te	19.3 days
$^{121\text{m}}\text{Te}$	154 days
$^{123\text{m}}\text{Te}$	119.2 days
$^{125\text{m}}\text{Te}$	57.40 days
$^{127\text{m}}\text{Te}$	9.35 hours
^3H	12.32 years
^{109}Cd	461.3 days
^{22}Na	2.60 years
^{113}Sn	115.08 days

Table 7.1: The significant radioisotopes with half lives below one million years in the NaI(Tl) crystals at COSINE-100, shown with their half-lives.

Each component is treated separately in each crystal and is assumed to follow the normal decay equation (where here the decay constant λ is used), i.e. the rate β_k^i of the k^{th} component in the i^{th} crystal is given by

$$\beta_k^i(t) = \beta_{0,k}^i e^{-\lambda_k t}. \quad (7.2)$$

Where $\beta_{0,k}^i$ is chosen as the event-rate of the component at midnight on the first of January 2016 and t is the time since that date.

The background modelling described above and in [74] does not return activities at a specific time, but rather the average event-rate of each component in each crystal over the 1.7 year time period studied. This value can be generally represented as

$$\beta_{0,k}^{\bar{i}} = \frac{1}{\Delta t} \int_{t_0}^{t_0+\Delta t} \beta_{0,k}^i e^{-\lambda_k t} dt = \frac{\beta_{0,k}^i e^{-\lambda_k t_0}}{\lambda_k \Delta t} (1 - e^{-\lambda_k \Delta t}). \quad (7.3)$$

Solving for $\beta_{0,k}^i$ gives

$$\beta_{0,k}^i = \frac{\beta_{0,k}^{\bar{i}} \lambda_k \Delta t}{1 - e^{-\lambda_k \Delta t}} e^{-\lambda_k \Delta t}. \quad (7.4)$$

By choosing $\Delta t = 1.7$ years and t_0 as the time between the start of 2016 and the 1.7 year data-set on the 21st of October 2016, $\beta_{0,k}^{\bar{i}}$ becomes the value returned by the background model and the other values are known constants, thus generating the time-dependent activities for each time-varying background component in each crystal, as required in Eq. 7.1. These values are shown in Fig. 7.2 where the uncertainties are found by propagating the uncertainties from the background model in [74].

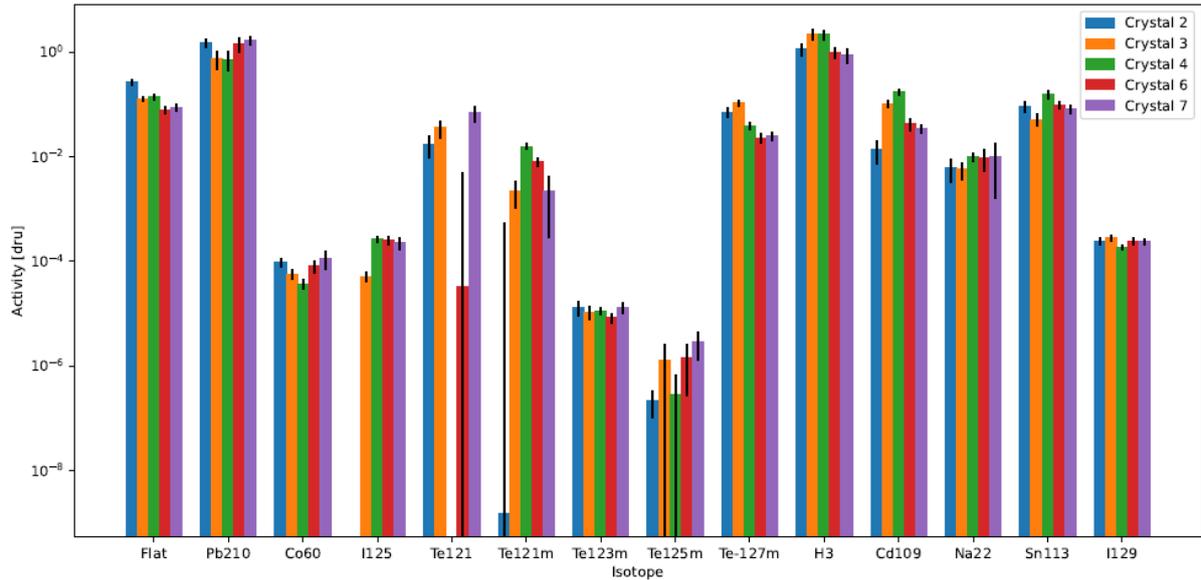


Figure 7.2: The activities and associated uncertainties in each crystal on the 21st of October 2016 (the start of physics data taking at COSINE-100) for all of the radiogenic components considered for inclusion as backgrounds in the annual modulation model. From [107].

As Fig. 7.2 shows, the significance of the background components in terms of their contribution to the overall background event-rate varies considerably between backgrounds. Therefore to reduce the number of nuisance parameters in the model, the components with $\beta_{0,k}^i$ values below 10^{-4} dru were not considered in the background model, thus removing the ^{60}Co , ^{125}I , $^{123\text{m}}\text{Te}$, and $^{125\text{m}}\text{Te}$ components. This 10^{-4} dru cutoff is a conservative choice considering that 10^{-4} is roughly two orders of magnitude below the annual modulation amplitude reported by DAMA, as well as significantly lower than the expected uncertainty in the annual modulation value at COSINE-100. Summing these remaining eight components returns the $\sum_{k=1}^{N_{bkgd}} \beta_{0,k}^i e^{-\lambda_k t}$ term in Eq. 7.1 which is fit to the data. The activities of each of the background components were allowed to float in the fit and is independent for each crystal. The decay constants of each component were fixed based on the standard values for the given component. The energy spectra of full background model across the SET3 data-set is shown with the contributions from the eight backgrounds included in the time-dependent annual modulation fit in Fig. 7.3 with very good agreement between the data and the predictions from the background model.

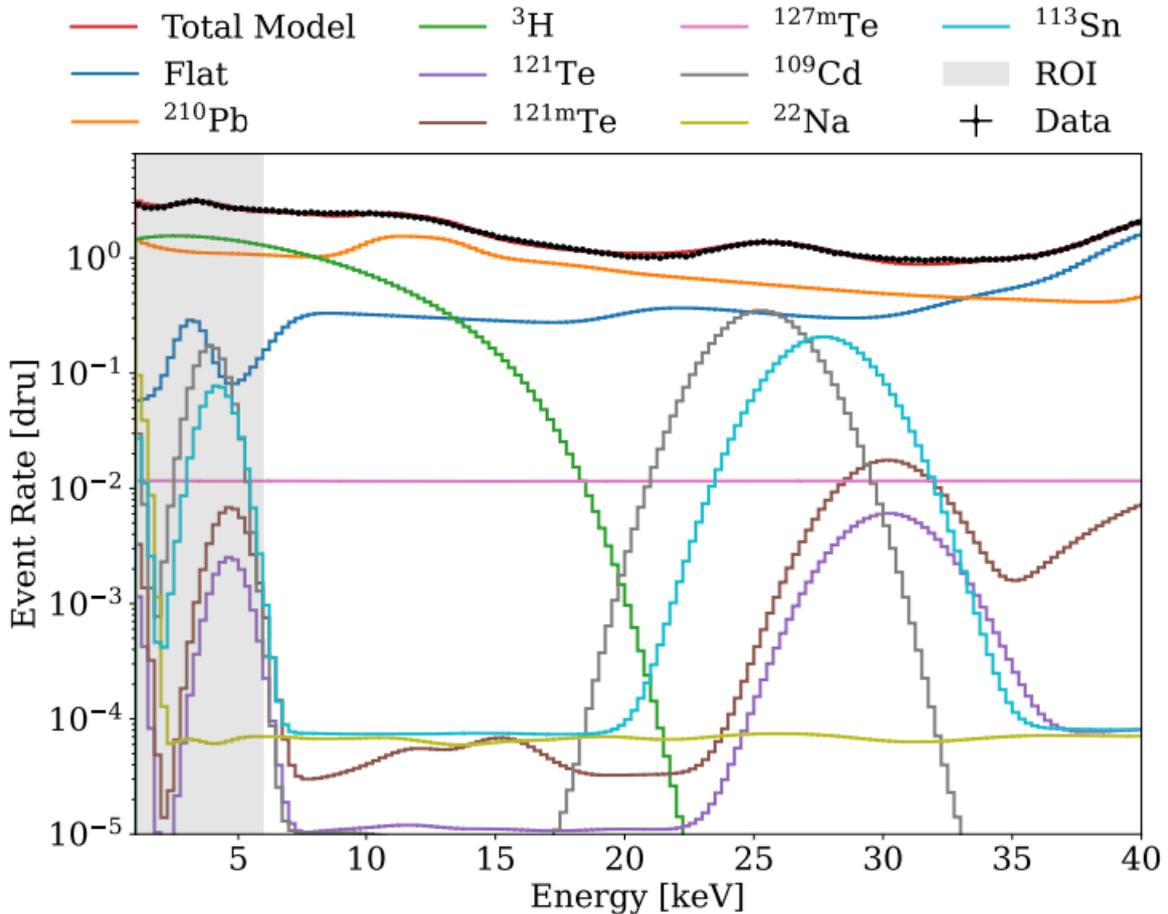


Figure 7.3: The energy spectrum over the SET3 time period for single-hit signal like data (black data points) compared to the background model described in [74]. The eight background components included in the time-dependent annual modulation fit are shown separately, and the grey shaded region highlights the 1-6 keV ROI. From [107].

7.1.3 Recovering best-fit modulation parameters

As a direct result of building a background model that properly considers all relevant background components in the annual modulation search at COSINE-100, a large number of nuisance parameters are introduced into the model that is fit to the signal-like data. Specifically a new nuisance parameter is introduced for every individual background component in each crystal. To deal with these nuisance parameters and recover the annual modulation component of the fit model, a Bayesian analysis framework using Markov Chain Monte Carlo techniques is used to calculate marginalised posterior distributions for the annual modulation parameters as required. Specifically, the Metropolis-Hastings algorithm [109, 110] with either 10^7 (for the fixed-phase analysis) or 10^8 (for the floated-phase analysis) iterations is used to carry out the multi-variable integration of the posterior probability density function for a given parameter of interest. More iterations are required in the floated-phase analysis to generate smooth contours when calculating highest-density credible regions of the two-dimensional posterior distributions. The step size between each iteration is chosen to give an acceptance ratio of between 20% and 25% and varies slightly depending on the specific fit being carried out (i.e. different energy ROIs and fixed vs floated-phase fits).

For example, the posterior distribution of the annual modulation amplitude S_m given

the observed data \mathbf{x} is a probability density function described by

$$P(S_m | \mathbf{x}) = N \int d\boldsymbol{\alpha} \int d\boldsymbol{\beta} \mathcal{L}(\mathbf{x} | S_m, \boldsymbol{\alpha}, \boldsymbol{\beta}) \pi(S_m, \boldsymbol{\alpha}, \boldsymbol{\beta}), \quad (7.5)$$

where N is a normalisation constant, $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are vectors representing the constant and time-varying background components described in Section 7.1.2 (i.e. the nuisance parameters). The priors for these distributions and for the annual modulation amplitude distribution are contained in $\pi(S_m, \boldsymbol{\alpha}, \boldsymbol{\beta})$. Finally the term $\mathcal{L}(\mathbf{x} | S_m, \boldsymbol{\alpha}, \boldsymbol{\beta})$ represents the following binned likelihood of the observed data given the annual modulation and background components, built as a product of Gaussian probabilities:

$$\mathcal{L}(\mathbf{x} | S_m, \boldsymbol{\alpha}, \boldsymbol{\beta}) = \prod_i^{N_{crys}} \prod_j^{N_{bin}^i} \exp \left[-\frac{1}{2} \left(\frac{x_{ij} - \mu_{i,j}}{\sigma_{ij}} \right)^2 \right]. \quad (7.6)$$

Here the indices i and j act over the $N_{crys} = 5$ crystals used in the annual modulation search, and the N_{bin} time bins respectively, $\mu_{i,j}$ is the expected rate in the j^{th} time bin in the i^{th} crystal with the associated uncertainty σ_{ij} . The prior distributions for the activities of background components are Gaussian with means and standard deviations based on the results of the time-dependent background model described in Section 7.1.2, whereas the annual modulation components are given flat prior distributions.

In the primary analyses for the SET3 and SET4 data, the phase of the annual modulation component in Eq. 7.1 is kept fixed at 152.5 days (in line with predictions from the standard halo model [65]) and the only parameter of interest is the modulation amplitude. In this case a one-dimensional posterior distribution is produced for the modulation amplitude, and to find the best-fit value of the parameter the mean of the distribution is computed. The 1σ uncertainty of this value is given by the standard deviation of the distribution. Since the posterior distributions have a high number of entries (10^7 entries in the case of searches with physics data) and are very well described by a Gaussian, the $\pm 1\sigma$, $\pm 2\sigma$, $\pm 3\sigma$ e.t.c. represent a very close approximation of the 68.3%, 95.5%, 99.7% etc highest-density intervals (HDIs) and are used instead of calculating the HDIs numerically.

An annual modulation search with the phase as a floated parameter in the model is also carried out. In this analysis it is still possible to produce a one-dimensional posterior distribution for both the modulation amplitude and phase as in the fixed-phase case and this is again how the best-fit values are calculated. Likewise the uncertainties for each parameter can be calculated from the standard deviation of the Gaussian posterior distribution. It is also possible to combine the posterior distributions of both values to draw a two-dimensional distribution and calculate highest-density credible regions (HDRs) therein to represent the uncertainties of the distribution in two-dimensional amplitude-phase space.

7.2 Development of the SET4 fitting procedure

While the modulation analyses carried out thus far at COSINE-100 are thought to be robust and provide a good description of the measured data, the methods used must regularly be examined and, if necessary, updated to add improvements or correct for any issues found. Between the SET3 and SET4 analyses, a potential issue was uncovered in the treatment of background components in the model used to fit the data when searching for annual modulation and, separately, an improvement in the treatment of the data with

respect to live-time scaling was developed. Neither of these changes results in the validity of the SET3 result being compromised, but it was necessary for them to be addressed to ensure the robustness and validity of the process for future COSINE-100 analysis. This section outlines these improvements to the fitting procedure.

7.2.1 Live-time scaling

The model to which the SET3 data was fit (shown in Eq. 7.1) does not contain any information about the live-time of any given time bin, or in other words, assumes 100% live-time in any given time bin. This, would be an unrealistic model if the fact that many bins have less than 100% live-time was not somehow incorporated into analysis procedure. As such, the process of scaling the event-rate to full 15 day bins as described in Section 7.1.1 was necessary for the model described to fit the data and this scaling was carried out for annual modulation searches at COSINE-100 until the end of the SET3 search.

While this process of scaling the data should be functionally equivalent to incorporating the real live-time information into the model and is expected to produce the same results with regards to the best-fit values and uncertainties of the modulation parameters, there are other factors which must be considered in the construction of the model and the treatment of the data. Specifically, the primary goal of COSINE-100 is to test the claim of a dark matter-induced signal from the DAMA experiments. This is a controversial result, in part due to the reluctance of the DAMA collaboration to publish their full rate vs time data-set and instead only reporting the residual event-rate (as discussed in Section 3.3.2). Consequently, COSINE-100 as a test of the DAMA signal is best served by presenting results of analyses as completely and as transparently as possible. Reporting/modelling the rate against time data after this process of live-time scaling runs contrary to that principle, even if this method is theoretically sound.

It is also important to test and validate any assumptions made when building models to describe the behaviour of data. There are well motivated reasons to assume that the periods of less than 100% live-time should not have any significant effect on the annual modulation fit result, since these periods are randomly distributed throughout the data-taking period. However, at no point prior to this work was the validity of this assumption tested, since when building simulated annual modulation data, aka “pseudo-data” (described for the SET4 analysis in Chapter 8), the main tool used for validation of the fit model, 100% live-time was assigned for each bin. Consequently no information about the real live-times of the bins in the physics data were included in this validation process. To have full confidence in the COSINE-100 annual modulation fitting results, and therefore in COSINE-100 as a test of the DAMA result, this assumption about the negligible impact of the live-time effects on the modulation result must be tested in the case of already published results, and ideally removed entirely from future analyses.

To resolve this potential issue, the model to which the data is fit in the SET4 analysis is updated to include the real live-time information of each bin in the data, rather than scaling the data to assume the rate in the case of 100% live-time. This has the result of adding a factor to the model described in Eq. 7.1 as follows:

$$R^{i,k}(t | S_m, \alpha^i, \beta^i, T^{i,k}) = \frac{T^{i,k}}{T_{full}} \left[\alpha^i + \sum_{k=1}^{N_{bkgd}} \beta_{0,k}^i e^{-\lambda_k t} + S_m \cos(\omega(t - t_0)) \right]. \quad (7.7)$$

The new variable $T^{i,k}$ represents the live-time of the k^{th} time bin in the i^{th} crystal, and

$T_{full} = 15$ days is the maximum possible live-time in any given time bin. The term $T^{i,k}/T_{full}$ therefore represents the fractional live-time of each time bin and the model is scaled accordingly.

The efficacy of this change to the fitting method was verified by re-analysing the SET3 data with this updated method, allowing for the new method to be tested with a very comparable data-set to that which will be used in the SET4 search without un-blinding any SET4 data. This has the additional benefit of verifying the assumption made during the SET3 analysis that scaling the data as described above does indeed have the same effect as scaling the model. The non-live-time-scaled SET3 rate vs time data as fit by the scaled annual modulation model described in Eq. 7.7 is shown in Fig. 7.4.

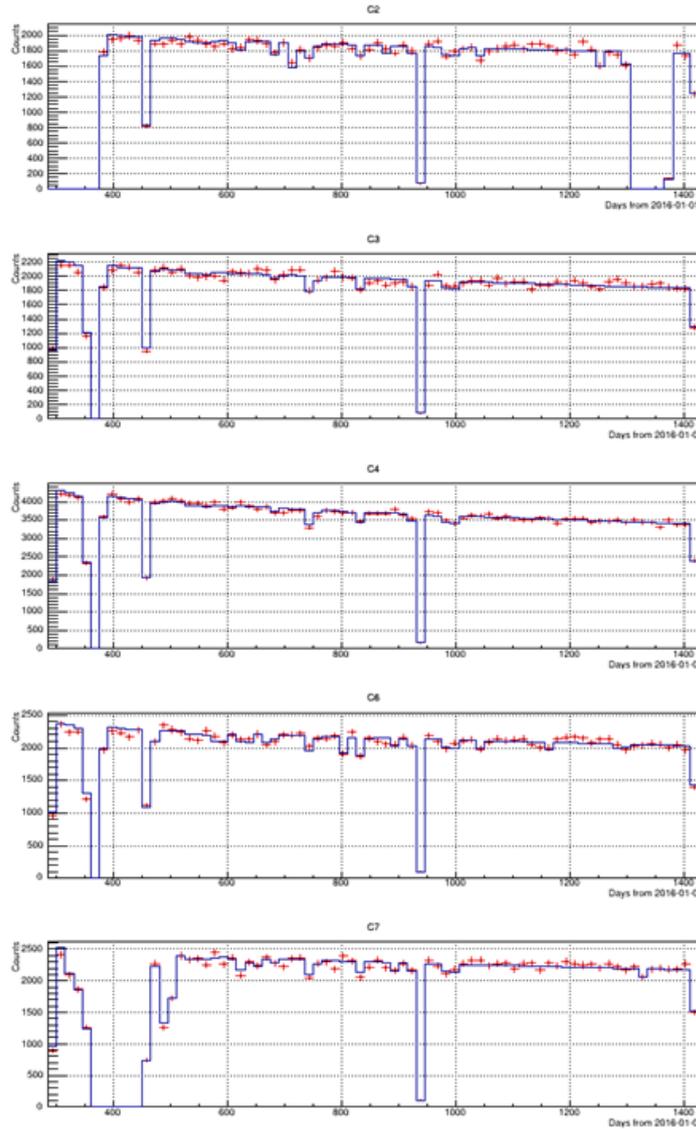


Figure 7.4: SET3 rate vs time data for 1-6 keV single-hit events (red data points) and associated fixed phase annual modulation fit from modelling the data with Eq. 7.7 (blue line). This analysis differs from that published in [1] in that the model fit to the data is scaled, rather than the data itself being scaled.

The best-fit parameter for the modulation amplitude was extracted from the mean of the Gaussian posterior distribution and found to be negligibly different (less than 2×10^{-4}

dru) from the value published in [1], i.e. less than an order of magnitude different from the uncertainty on the parameter. Likewise the uncertainty associated with the best-fit modulation amplitude (the standard deviation of the posterior distribution approximating a 68.3% HDI) is found and also shown to be similarly unchanged. This result has the significant benefit of validating the assumption made in the SET3 analysis about the scaling of the data to account for bins having less than 100% live-time.

7.2.2 Accounting for surface ^{210}Pb effects

While the fit results of the annual modulation data are usually marginalised over the posterior distributions of the annual modulation-specific parameters, it is equally valid to marginalise over the posterior distributions of the activity of any given background component to obtain its best-fit parameter and associated uncertainty. Indeed, during the process of re-evaluating the SET3 annual modulation techniques in preparation for the SET4 analysis, the efficacy and reliability of the annual modulation model in the context of correctly describing the activities of background components in the data was tested. This was achieved by obtaining these best-fit background activities from the SET3 annual modulation fit results (obtained using the updated live-time scaling procedure described above) and comparing them to similar best-fit values found through different methods in different analyses. This check and subsequent update to the fit model in light of the results are described in this subsection.

Identifying an issue with the best-fit values of background activities

The best-fit values of the initial activities of the background components in the 1-6 keV SET3 annual modulation search were compared to the equivalent activities from the following sources: the best-fit activities from the model-dependent WIMP search with SET2 data at COSINE-100 [3], the initial activities from the time-dependent background model [74] which are used as prior distributions for the background activities in both the SET3 annual modulation analysis and model-dependent WIMP search, and the best-fit activities from fitting an ensemble of annual modulation pseudo-data with background activities chosen to match the input values from the time-dependent background model, fit with the same model as the SET3 analysis. All these analyses had the same region of interest (1-6 keV). This is shown for both logarithmic and linear activity y axes in Fig. 7.5. In this figure (and in all subsequent discussions), the individual activities for each crystal have been summed, and their uncertainties summed in quadrature to give a single value and uncertainty for each component, however, the values from individual crystals were also checked.

The process of generating annual modulation pseudo-data is described in full in Chapter 8. In simple terms, an ensemble of 1000 data-sets each with 15 day-long bins with the same start and end points, and the same live-time scaling as the physics data are created. The rate in each bin is determined by the activities expected from the time-dependent background model, but where the exact activity of each component is selected randomly from the one sigma uncertainty bands of each component for each data-set. Poissonion fluctuations are then applied to each bin. In the full pseudo-data analysis an annual modulation component is added to the rate and five ensembles of pseudo-data are generated with modulation amplitudes ranging from zero to the amplitude reported in the DAMA experiments. In this case where the goal is to compare the best-fit values of background activities, no annual modulation component is added for the sake of simplicity. The best-

fit values of any given parameter for the 1000 pseudo-experiments will follow a Gaussian distribution, and the means and standard deviations of these distributions give the values quoted for the activity and associated uncertainty of a given background component in this analysis.

The generation and comparison of pseudo-data against physics data serves two purposes here. The first is to give an indication of the spread of best-fit values for the background activities given by this fit model, i.e. the standard deviation of the Gaussian distribution described above. This will give an indication of whether it is likely that any deviations in activities seen are a result of random fluctuations. The second purpose is to check that the model used to search for annual modulation is correctly able to find the values for the activities of the background components (i.e. the values chosen when generating the pseudo-data).

In the model-dependent WIMP search, the activities of the ^{121}Te components are not reported individually and instead combined with the flat component and called “other”. To make a direct comparison with the annual modulation analysis, the best-fit values of the ^{121}Te activities are summed and added to the flat component with their uncertainties being summed in quadrature. It should be noted that the activities of the ^{121}Te components are very small and have a negligible effect on the outcome of both this comparison and the overall analysis.

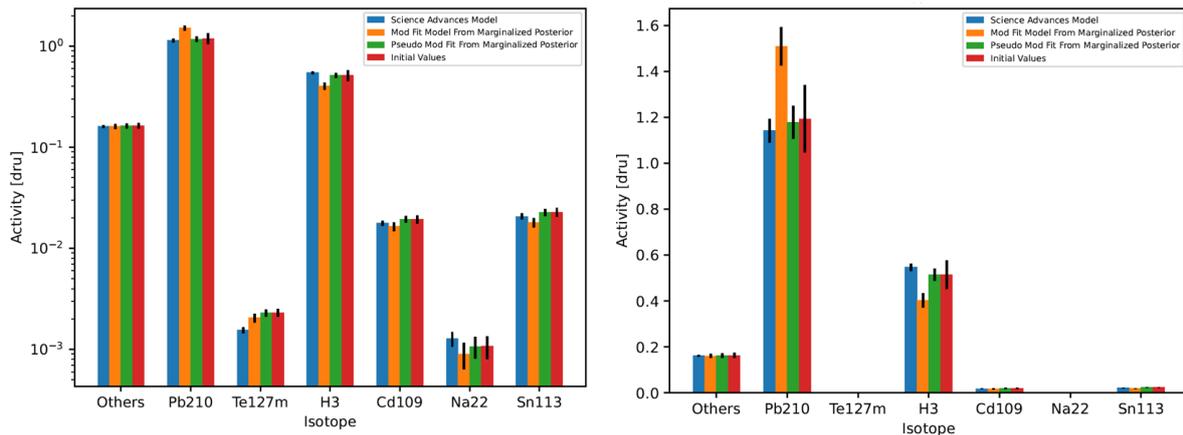


Figure 7.5: The best values for the background components from different analyses of 1-6 keV single-hit COSINE-100 data. These analyses are the model dependent WIMP search with SET2 data [3] (blue bars), the modulation analysis with SET3 data [1] (orange bars), simulated data of the SET3 annual modulation analysis (green bars), and the values found in the initial study of the COSINE-100 backgrounds (red bars). The left and right plots show the same data with logarithmic and linear y axes respectively.

As can be seen in particular on the linear-y-axis version of Fig. 7.5, there is a large disparity, well above what would be expected from the measured uncertainties, in the best-fit values for the activities of the ^{210}Pb and ^3H components in the SET3 annual modulation results compared to the other results. Specifically, the activities of ^{210}Pb and ^3H were found to be 2.3σ and 1.8σ away from the value expected from the pseudo-experiments. This disparity is too large to ignore in the context of how well the other analyses agree with the inputs for these components.

Because of the difference in these activities, there is a 1.2σ disparity in the total activities of the SET3 results compared to the pseudo-data where the total activity of the SET3 data is found by summing the activities of the individual components and

this value is then compared to the (Gaussian) distribution of the summed activities of the 1000 pseudo-experiments. Since the uncertainties on the activities of the pseudo-data components are higher than that of the values from the other two analyses, the SET3 values are in further tension with the results from these analyses. This level of tension is unacceptably high level of tension. There are smaller but non-zero deviations between the SET3 activities and other activities in the other components besides ^{210}Pb and ^3H but these are within the expected uncertainties and so are not considered problematic. The activity of $^{127\text{m}}\text{Te}$ in the model-dependent WIMP search is also non-negligibly low compared to the values from other analyses, but the overall contribution from $^{127\text{m}}\text{Te}$ is very small and the details of the results of the model-dependent WIMP search are beyond the scope of this work so this effect is not considered for further analysis.

In addition to the plot shown in Fig. 7.5 where the best-fit values for each crystal are combined, the best-fit values were checked at a crystal-by-crystal level to determine if the difference in fitted activities in the combined plots was a result of a mismatch in individual crystals or an issue affecting the whole model. However, no obvious dependency on any specific crystal was found, indicating an issue that affected the fitting method itself, rather than an issue with a specific crystal. Note that the best-fit values for the activity of the pseudo-data very closely match the input values, therefore it can be inferred that the fit model does on average return the correct values for background activities when the correct inputs are given to the model (which is the case by definition with pseudo-data).

Updating the model to account for extra surface ^{210}Pb

The cause of the over-estimation of the ^{210}Pb activity and subsequent under-estimation of the ^3H activity in the SET3 analysis was found to be from a poor modelling of the exact behaviour of the ^{210}Pb . The overall contribution made by ^{210}Pb in NaI crystals comes from two parts, the bulk ^{210}Pb i.e. from ^{210}Pb that could not be removed from the powder used to grow the crystals and is now contained inside the crystal, and secondly the surface ^{210}Pb which was not present during the crystal growing process, but instead is later deposited onto the crystal surface as a result of the decay of airborne ^{222}Rn . Rn is found in the Earth's atmosphere despite its relatively short half-life as a result of constant replenishment from the decay of other, longer lived radiogenic materials, specifically ^{238}U . The majority of atmospheric radon will be ^{222}Rn with a half-life of 3.8 days. The daughter isotopes of ^{222}Rn decay have half-lives of 27 minutes or less until ^{210}Pb with a half-life of 22.2 years is reached.

The crystals will inevitably be exposed to air before their final installation at COSINE-100 and as a result ^{222}Rn and therefore ^{210}Pb which will be deposited on the crystal surface. Additionally, despite efforts to filter out as much atmospheric Rn from the COSINE-100 detector room as possible, there remains around approximately 30 Bq/m^3 of Rn measured in the detector room at any given time, which may result in additional small amounts of ^{210}Pb being deposited onto the crystals after installation if the Rn can permeate into the detector to reach the crystals. The ratio of bulk to surface ^{210}Pb in the five analysis crystals was found to be between less than two in C4 and over 20 in C7 in the time dependent background study [74]. The exact depth profile of the surface ^{210}Pb was studied in detail in the COSINE-100 crystals in [111].

In the model used for the annual modulation fit of the SET3 data, the surface and bulk ^{210}Pb components are treated as a single exponential component as described by one of the $\beta_{0,k}^i e^{-\lambda_k t}$ terms in Eqs. 7.1 and 7.7. Here the initial ^{210}Pb activity is modelled as the sum of the surface and bulk components and the decay constant as the standard

value for ^{210}Pb (i.e. calculated from a half-life of 22.2 years). This carries the built in assumption that the expected rate of ^{210}Pb events decreases in the standard exponential way. However, this assumption will not hold in the case where some of the surface ^{210}Pb is continually replenished by ^{222}Rn entering the detector shielding. In this case, the half-life of the surface ^{210}Pb will effectively be extended beyond the expected 22.2 years used in the model used to fit the SET3 annual modulation data. Since the decay constants of the background components are fixed parameters in the fit, the extended half-life from the surface ^{210}Pb is incorrectly modelled as increased ^{210}Pb activity. This in turn lowers the modelled activity of the ^3H component, the second largest component behind ^{210}Pb to redress the overestimation of the activity of ^{210}Pb .

To calculate an estimate for the revised half-life and hence decay times of the surface ^{210}Pb (the surface alpha events) from around 2100 days of COSINE-100 data (from the start of data taking in October 2016 to June 2022) were selected and split into 30 day bins for each crystal. The activities in each crystal were then summed and the resulting activity over time was fit to an exponential decay function. Initially it was intended to model each crystal individually, but there were insufficient statistics to get meaningful results for all the crystals. Using this combined analysis the effective half-life of the ^{210}Pb is found at 34 ± 8 years; notably higher than the standard value.

Using this new-found value for the half-life of the surface ^{210}Pb the background model used in the annual modulation searches described above is updated to split up the original ^{210}Pb component into separate surface and bulk ^{210}Pb components. Unlike the decay constants of the other components, which are fixed when fitting the data to the models described above, the decay constant of the surface ^{210}Pb is allowed to float independently for each crystal and given a Gaussian prior distribution with a mean and standard deviation equivalent to a half-life of 34 ± 8 years for all crystals. Ideally, each crystal would have its own prior distribution for the decay constant to represent the fact that individual crystals may have different susceptibilities to extra deposition of surface ^{210}Pb . However, these differences are small and not expected to affect the overall result a significant amount.

Furthermore, the calculation of the prior distributions for the activity of the surface ^{210}Pb subtly differs from that of the activity of the other background components in that the time dependent background model [74] provides the activity of each component in 1 keV wide bins from 1-2 keV to 19-20 keV, and these bins are summed to give the prior distributions for the initial activity of a given search region of interest. For example, in the case of the 1-6 keV search region of interest, the activities of each component in the five 1-2 keV, 2-3 keV etc bins are summed (with their uncertainties summed in quadrature), to inform the mean (and standard deviation) of the Gaussian prior distribution. However, in the case of the separate surface and bulk ^{210}Pb only the activity and uncertainty in the 1-6 keV bin is recorded (see Table 2 in [74]).

Therefore in the case where the annual modulation search region of interest is not also 1-6 keV, an inference about the activities of the surface and bulk ^{210}Pb must be made at other energies. The fractional activities of the surface and bulk ^{210}Pb were assumed to be equal in all energy regions for each crystal and so equal to that in the 1-6 keV region. The combined surface + bulk ^{210}Pb activities in the other energy regions were therefore converted to single surface and bulk activities accordingly. It should be noted that the primary energy ROI of this work is the 1-6 keV region where no inference is required.

Testing efficacy of changes on SET3 data

As with the improved method of scaling the model used to fit the annual modulation data by the detector live-time, the SET3 data provides an excellent test-bed for trialling the new split ^{210}Pb background components in the fit model without un-blinding the SET4 data, while also highlighting the degree to which this issue may have affected the published SET3 annual modulation search result. As the method of scaling the model to match the real data live-time was shown to be equivalent to the original method of scaling the data to assume 100% live-time in Section 7.2.1, the fit model was scaled with the live-time of the data in this check.

As described above, the fit model is updated to split the previously single ^{210}Pb component into two separate surface and bulk ^{210}Pb components. Here the decay constant for the bulk ^{210}Pb component is the standard value for ^{210}Pb and is fixed in the model, whereas the surface ^{210}Pb decay constant is floated independently across the five crystals. All crystals have the same added Gaussian priors with a mean and standard deviation based on a half-life of 34 ± 8 years. The activities of the surface and bulk ^{210}Pb are fit for each crystal with the same method as the other background components and have the same Gaussian prior distributions described above. The best-fit values for the background activities are found and compared to the input values (i.e. the values from the time-dependent background model) as shown in Fig. 7.6. In this comparison, it can be seen that fitting the ^{210}Pb components separately has largely resolved the tension between the input and output activities of the fit.

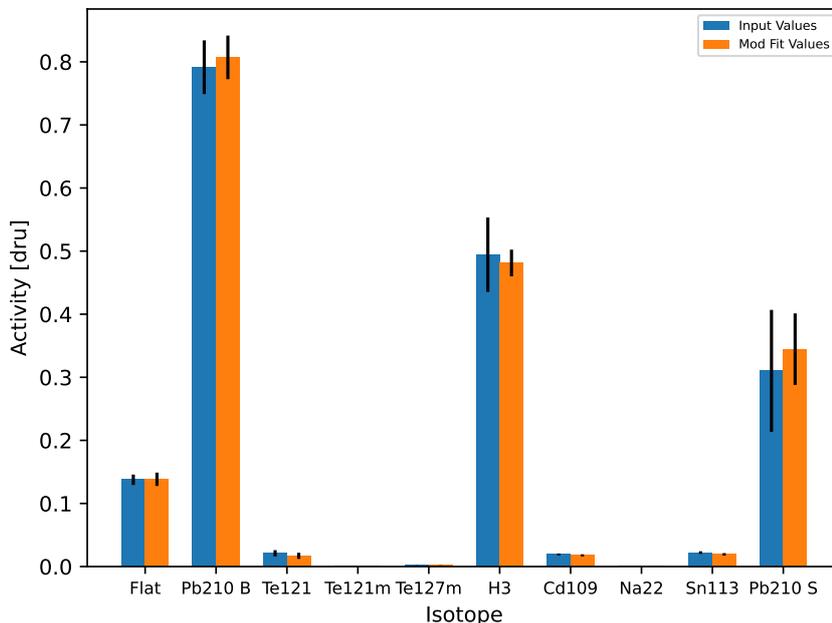


Figure 7.6: The best-fit values for the activities of the background components in the SET4 annual modulation model with split surface and bulk ^{210}Pb denoted by “Pb210 S” and “Pb210 B” respectively (orange bars) compared to the input values used in the fit from the time-dependent background model [74] (blue bars).

To compare the values from the updated fit model with split ^{210}Pb components with the values from the model-dependent WIMP search with a single ^{210}Pb component, shown

in Fig. 7.7, the surface and bulk ^{210}Pb components of the new model must be recombined into a single parameter. This is achieved by summing their values and summing their uncertainties in quadrature. Similarly, all of the best-fit values for the activities of all the components and their associated uncertainties are combined in the same way to give the total value of the initial activities for each analysis. This is also shown in Fig. 7.7. As expected, the effect of recombining the ^{210}Pb components in the new annual modulation model slightly eases the tension in the comparison between the updated annual modulation model and the other analyses as the relatively large surface ^{210}Pb uncertainty is folded into the uncertainty of the bulk ^{210}Pb . Furthermore, the alterations to the ^{210}Pb parameter in the fit have also had the effect of increasing the best-fit value of the ^3H component as it is no longer suppressed by the over-counting of surface ^{210}Pb events.

Overall when summing all of the components in this way, the total initial activities of the models are within the calculated uncertainties of one another and show good agreement across all the components. The above analysis was repeated for the case where rather than the effective half-life of the surface ^{210}Pb being uncorrelated across the crystals, the value is instead correlated for all the crystals. In this case there is no significant change in the best-fit values of the activities.

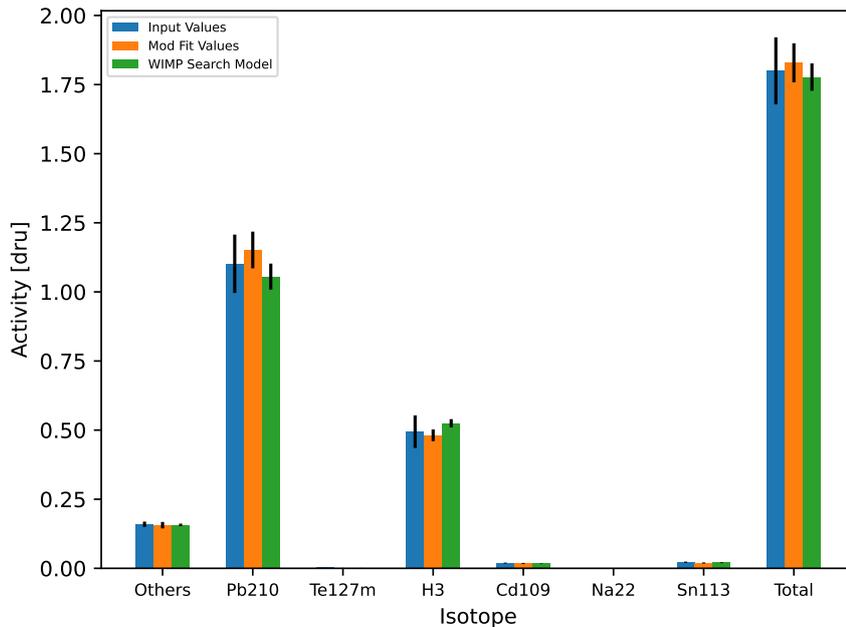


Figure 7.7: The best-fit values for the activities of the background components in the SET4 annual modulation model with split surface and bulk ^{210}Pb with the individual ^{210}Pb components recombined (orange bars) for comparison with the values from time-dependent background model [74] (green bars), and the model-dependent WIMP search [3] (blue bars). The summed best-fit activities of each method is shown by the bar labelled “total”.

In addition to checking the best-fit values of the activities in the SET4 annual modulation fit model, the best-fit values for the effective decay constant of the surface ^{210}Pb in each crystal was also extracted and converted to half-life to compare the values to the inputted half-life of 34 ± 8 years. The results for each crystal are shown in Table. 7.2 and represent good agreement with the expected value. Additionally, there is also good

agreement with the results of the values obtained for the effective half-lives of individual crystals in the case where there was sufficient data for a given crystal to get a meaningful result for the effective half-life from the surface alpha data. However, as discussed in more detail in the following chapters, an issue with the calculation of the surface ^{210}Pb best-fit values was later found and as such, the values in Table. 7.2 are not expected to be correct. Nevertheless, these issues were not uncovered until later and therefore the values found at the time of analysis are included for posterity.

Crystal	Half life (years)
C2	30.1 ± 7.8
C3	32.7 ± 8.1
C4	32.2 ± 7.9
C6	36.9 ± 8.0
C7	31.9 ± 7.9

Table 7.2: The best-fit values for the effective half-life of surface ^{210}Pb when the updated annual modulation model for split bulk and surface ^{210}Pb is fit to SET3 data.

While the above discussion has been centred around ensuring the best-fit values of the background activities (and effective half-lives) are accurate and well described by the model used in annual modulation searches, it must be remembered that the main goal of the COSINE-100 experiment is to search for an annual modulation signal. As such, the effect of the above inconsistencies and subsequent changes in the fitting of background values should also be analysed through the lens of the effect on the annual modulation result. Fortunately, it was found that the best-fit modulation amplitude was largely unaffected by the changes to the background model described above. Various iterations of the model used to search for annual modulation in the SET3 data were trialled, including the models described above, before a final version was reached. In these various models, when a like-for-like comparison was made with the already published [1] value for the annual modulation amplitude of 0.0067 ± 0.0042 dru for the 1-6 keV ROI, the new value and associated uncertainty was never more than $\pm 5\%$ from the published SET3 value. Given that the uncertainty on the best-fit modulation amplitude is over an order of magnitude larger than these differences, it can be said that these changes to the treatment of ^{210}Pb do not have a significant impact on the results of the annual modulation analysis, and the results published for the SET3 analysis are robust. Nevertheless, these changes represent a significant improvement in the robustness of the annual modulation fit model at COSINE while also addressing a mis-modelling issue that could not be left unresolved.

7.3 Concluding remarks

The changes to the COSINE-100 annual modulation fitting procedure from SET3 to SET4 have been described, focusing on treatment of the data.fit model in the light of not every time bin having 100% live-time, and also the updates to the treatment of ^{210}Pb in the fit model. Specifically it was found that the surface ^{210}Pb , which previously had been combined with the bulk ^{210}Pb and treated as a single parameter in the model, was being somewhat replenished through the deposition of extra ^{210}Pb onto the crystal surface as a result of ^{222}Rn permeating the detector, extending the effective half-life of the surface ^{210}Pb . While these changes ultimately did not represent a significant change

to the annual modulation result when they were tested on SET3 data against the already published result, they do make the results of the SET3 and future results more robust, and therefore a better test of the DAMA results. The changes described above also represent a significant re-working in the machinery used for the annual modulation analysis of COSINE-100 data.

Overall, these improvements to the fit model which will be used to fit the COSINE-100 single-hit data as described in the upcoming chapters make for a more robust and accurate fit model than the models used previously. Furthermore in testing these improvements the already published COSINE-100 annual modulation results are verified.

Chapter 8

Pseudo-Data Validation of Analysis Techniques

The generation and subsequent testing of analysis techniques on simulated data (aka “pseudo-data”) at COSINE-100 is a crucial step in the modulation analysis procedure described in this thesis. By testing the annual modulation fitting procedure on pseudo-data, not only can the procedure be verified in terms of its ability to correctly identify relevant fit parameters, but any issues with the procedure should be identified before applying the procedure on the physics data. Consequently the physics data can be kept blinded until the point at which it is analysed, making the final results more robust and safer from unconscious bias.

In this chapter, in Section 8.1, the need for pseudo-data production and analysis is described, followed by a description of the method of production of the pseudo-data in Section 8.2 and the results of testing the annual modulation fitting procedure on the pseudo-data in Section 8.3. In testing the fitting procedure, some issues in the procedure were uncovered. These are described and discussed in Section 8.4. Finally, a novel method for interpreting the results of the floated-phase analysis is described in Section 8.5.

8.1 The motivation for pseudo-data analysis

The most important results of the annual modulation searches at COSINE-100 are the best-fit values associated with the annual modulation parameters in the fit model (for example Eq. 7.7) applied to the single-hit event rate in the physics data. It is these best-fit parameters that characterise the observed annual modulation in the COSINE-100 data and allow for a comparison with the DAMA annual modulation results. However, when building and applying a model to observed data, there will always be some uncertainty in the efficacy of the fit model as without totally (unrealistically) accurate event reconstruction, the origin of events is not known on an event-by-event basis. Rather a guess is made based on the nature of the events. As such, even if the fit model used appears to be a good description of the observed data, it could be the case that certain events are mis-identified and this could cause the best-fit parameter of interest to be incorrectly measured. Consequently it is necessary to understand in detail the performance of a fit model applied to measured data.

One such method of testing a fit model is to test the model on simulated data (aka “pseudo-data”). Unlike in a real physics data-set where the provenience of any given event, the contributions of background components, the efficiency of event selection etc. can only

be estimated (even if it is to a high degree of precision and accuracy), these features are known with absolute certainty in the case of pseudo-data. This results in pseudo-data being an extremely useful tool in testing the performance of the annual modulation fit model at COSINE-100 as the analysis procedure can be tested in what is effectively a totally controlled environment. This means the best-fit parameters outputted by the fit model can be compared to the known input parameters to test for any bias in the fit procedure where the model inaccurately describes the input parameters.

A further benefit of this type of pseudo-data analysis is the results produced should give reliable estimates of where the best-fit values, associated uncertainties, chi-squared values, etc. are expected to fall when the physics data is analysed. In the case where the results of the analysis applied to the physics data do not match those from pseudo-data, it would be an indication that an issue has occurred in the analysis, assuming the pseudo-data correctly describes the physics data.

8.2 Production of the pseudo-data

In order for the tests of the effectiveness annual modulation fitting procedure on pseudo-data to be useful, the pseudo-data generated for the tests must match the physics data as closely as possible. Furthermore, enough variability must be introduced to test different analysis procedures across all the parameter space where the final physics result might fall. To achieve this, five different ensembles of pseudo-experiments were initially created, with each ensemble containing 1,000 pseudo experiments and each ensemble having a different injected modulation amplitude in order to test the response of the annual modulation fit procedure in different regions of parameter space. These injected modulation amplitudes were chosen as 0 dru, 0.0025 dru, 0.005 dru, 0.0075 dru, and 0.0105 dru (“daily rate units”), measured as counts/day/keV/kg. The injected amplitudes correspond to the no-modulation case, the annual modulation reported by the DAMA collaboration for a 1-6 keV region of interest (0.0105 dru) [49], then three intermediate values of 0.0025, 0.005, and 0.0075 dru to test the response of the fit model at a range of possible modulation amplitudes. To match the DAMA region of interest (ROI) and subsequently the main energy ROI in this work, the pseudo-data described here is chosen to recreate COSINE-100 physics data in the 1-6 keV ROI, i.e. the background activities used to determine the rate of each time bin (in addition to the modulation component) in each pseudo-experiment are based on the 1-6 keV activities from the time-dependent background model of COSINE-100 data [74]. These are the same values as the priors used in the 1-6 keV annual modulation fit.

The time binning of the pseudo-experiments is chosen to exactly match that of the physics data, that is 15-day bins from the 12th of October 2016 to the 29th of June 2021. This start date does not exactly correspond to the date that data taking started at COSINE-100, which is the 21st of October 2016. Instead, this date was chosen such that the start of the first bin fell an integer number of 15 day bins from the start of 2016. This has also been the time binning convention of the first two annual modulation analyses at COSINE-100 [1, 75] and is effectively a choice of convenience for the analyser, since in the analysis framework there is a convention of the activities of background components in the model being entered at the start of 2016. This choice has the effect of immediately under-filling the first bin of the data-set as the first nine days of the bin occur before the start of data taking. In order to verify that the choice of the start date of the time binning did not have a significant effect on the results, a second set of five ensembles of

pseudo-data were created with time binning starting at the same time as data-taking for comparison. The two sets of ensembles are identical aside from the difference in binning.

Since the pseudo-data is designed to accurately simulate the physics data, only the rates in the five crystals used in the single-hit physics analysis are simulated. Once the binning of the pseudo-experiments has been established, each bin is populated separately for each crystal using the following formula:

$$R^{i,k}(t \mid S_{inj}, \alpha_{fluc}^i, \beta_{fluc}^i, T^{i,k}) = \left\{ \frac{T^{i,k}}{T_{full}} \left[\alpha_{fluc}^i + \sum_{k=1}^{N_{bkgd}} (\beta_{fluc,0,k}^i) e^{-\lambda_k t} + S_{inj} \cos(\omega(t - t_0)) \right] \right\}_{fluc}. \quad (8.1)$$

Note that the model used to generate the pseudo-data is the same as the fit model described in Eq. 7.7, but updated to introduce random fluctuations and an injected modulation amplitude. The final term inside the square brackets is the annual modulation component with phase t_0 , a period given by $T = 365.25$ days, and angular frequency $\omega = 2\pi/T$. An injected annual modulation amplitude S_{inj} is chosen for each ensemble as described above. The other terms in the square brackets describe the background components where, the α^i component is the constant background rate in the i^{th} crystal and $\beta_{0,k}^i$ is the rate at time $t = 0$ (defined here as midnight on the first of January 2016) in the i^{th} crystal of the k^{th} background component (of which there are nine) with the decay constant λ_k . The nine varying background components are the same as those described in Section 7.2.2 and include a split surface and bulk ^{210}Pb component where the surface ^{210}Pb component has an extended half-life (decreased decay constant).

The *fluc* subscript on the α^i and β^i components denotes that the initial activities of the background components in a given crystal are allowed to vary randomly based on a Gaussian distribution with a mean and standard deviation given by the activity and uncertainty of the component in that crystal in the 1-6 keV ROI in the time-dependent background model [74]. Note that these are the same distributions as the prior distributions used in the annual modulation search model. These random fluctuations are different for every component and every pseudo-experiment. The new effective background activities of each component are saved after the fluctuations are applied for comparison with the best fit values after the fit has been carried out to test whether they are recovered correctly by the fit procedure. The values of the activities of the background components are varied in this way to test the fit model's performance over the full possible parameter space where it is likely the fit results to the physics data will fall. Although the half-life of the surface ^{210}Pb component is also a floated parameter in the fit model, it does not have the same random fluctuations applied as the initial activities of the background components.

An example of the output rates in a single pseudo-experiment for each of the five crystals, in this case with zero injected modulation amplitude, is shown in Fig. 8.1. Unlike in previous pseudo-data analyses, the rate in each bin is corrected based on the live/dead-time measured in the corresponding bin in the physics data-set, hence the large fluctuations in rate observed.

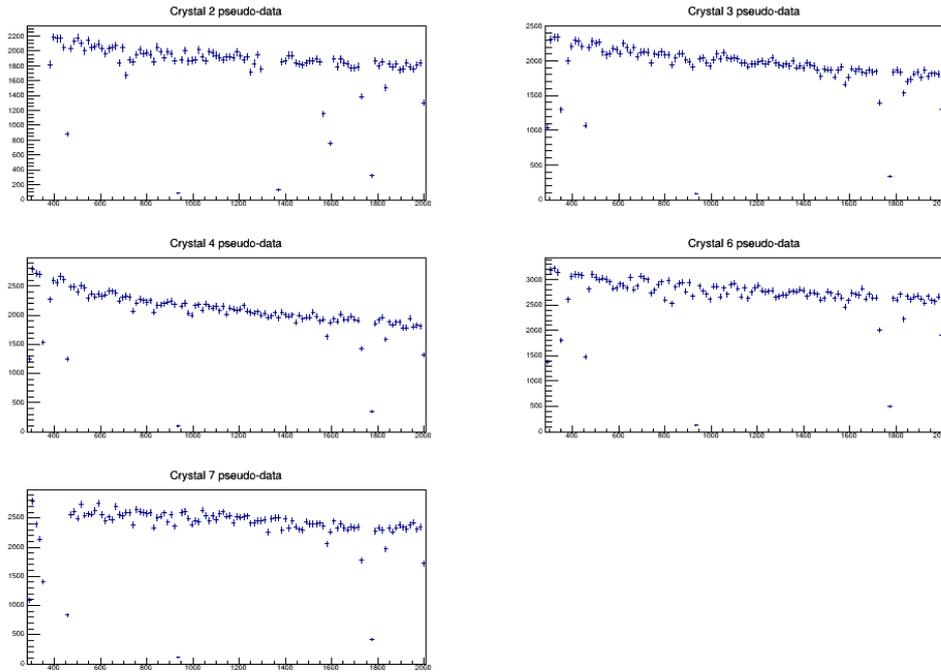


Figure 8.1: An example of the rates against time of a single pseudo-experiment, populated by Eq. 8.1, simulating single-hit physics data in the five analysis at COSINE-100 with an injected modulation amplitude of zero. As with the physics data, each data point represents a 15 day bin. Each bin is scaled by the live-time of the corresponding bin of the physics data so the pseudo-experiments more accurately represent the physics data.

In previous annual modulation analyses at COSINE-100, the individual background activities were varied in an uncorrelated manner, i.e. it was possible for most or all of the activities to be significantly larger (or smaller) than the input values from the background model. This was useful in testing a wider range of parameter space, but these pseudo-experiments with maximally high or low varied activities are unlikely to realistically represent the COSINE-100 data, since the overall event-rate is well known. Therefore inaccuracies in the time-dependent background model are likely to be correlated such that if one background activity is overestimated, another is likely to be underestimated to account for this and vice versa. An example of this is the overestimation of the ^3H activity as a result of the underestimation of the ^{210}Pb activity in the SET3 annual modulation analysis, shown in Fig. 7.5. During the SET4 pseudo-data analysis described in Section 8.3, it was found that the pseudo-experiments with the highest and lowest, and therefore least realistic, activities were disproportionately likely to have non-convergent fits. It was therefore decided that the fit method would be better tested if these highly uncorrelated pseudo-experiments could not be produced in the pseudo-data generation stage. Consequently, the ensembles of pseudo-data were re-generated such that the summed activities in each crystal of a given pseudo-experiment could not exceed $\pm 25\%$ of the summed input activities.

The *fluc* subscript on the right hand side of Eq. 8.1 denotes that the rate of a given crystal in a given bin has Poissonian fluctuations applied, where again the fluctuation is drawn from a Gaussian distribution, now centred around the initially calculated rate and with a standard deviation equal to the square root of the original rate. These fluctuations are applied to ensure that each pseudo-experiment has a realistic distribution of events in time, rather than the unphysical case where the event rate exactly followed the input

function.

As with the third-generation fit model described in Eq. 7.7, each bin is scaled by the real live-time information of each bin. These live-time scaling values are obtained from the calculation of the histograms for the rates of real physics data, however, it is important to note that this process of extracting the live-time information from the physics data prior to the completion of the pseudo-data analysis does not equate to un-blinding the physics data as no information about the rate of the physics events is extracted. This is the first analysis of annual modulation pseudo-data at COSINE-100 to use the real live-time information of the physics data to scale the pseudo-data. Therefore this pseudo-data should provide a more accurate description of the real COSINE-100 data where any effects the distribution of dead-time in the physics data are uncovered before the physics data is un-blinded.

8.3 Testing the fixed-phase modulation analysis procedure on the pseudo-data

Once the ensembles of pseudo-data have been generated, they are analysed with the annual modulation fitting method described in Section 7.2, the same method applied to the physics data-set. Each pseudo-experiment is modelled by Eq. 7.7 which takes the same form as Eq. 8.1 but with the additional fluctuations removed. Using a Bayesian analysis framework utilising Markov Chain Monte Carlo techniques, a marginalised posterior distribution for a given parameter of interest (in this case the modulation amplitude, S_m) is created and the best-fit value of the parameter is extracted from the mean of this distribution. The 1σ uncertainty on the best-fit value is calculated as the standard deviation of the distribution and very closely approximates the 68.3% highest-density interval (HDI) of the distribution. The modulation amplitude is fixed per unit mass across the crystals. The results of this fit to the pseudo-experiment shown in Fig. 8.1 is shown in Fig. 8.2.

In this case, the annual modulation fitting procedure is being tested with the phase of the annual modulation component fixed at 152.5 days, in line with the standard halo model [65] and the headline results from the DAMA and previous COSINE-100 analyses. The only practical difference in the analysis procedure of a pseudo-experiment compared to the physics data is that in the physics data analysis, the Metropolis-Hastings algorithm used to marginalise the posterior distributions is applied with 10^7 iterations, which is reduced to 10^6 iterations for a pseudo-experiment in order to reduce the computational time required. As a result, the number of entries in the posterior distribution of each parameter in the pseudo-data will be 10^6 , rather than 10^7 as is the case in the physics data.

This slightly increases the chances that the fit will not converge and the results from that pseudo-experiment have to be discarded, as well as producing posterior distributions with a factor of ten fewer entries than the physics data posterior distributions resulting in distributions which less closely follow a perfect Gaussian distribution and are therefore more likely to have slightly skewed means. Fortunately, only around 1% of pseudo-experiments do not converge and need to be discarded. Consequently the degree by which the best-fit parameters may be affected by the smaller number of entries in the posterior distribution is not only very small, but also likely to be averaged out over the many experiments in each ensemble.

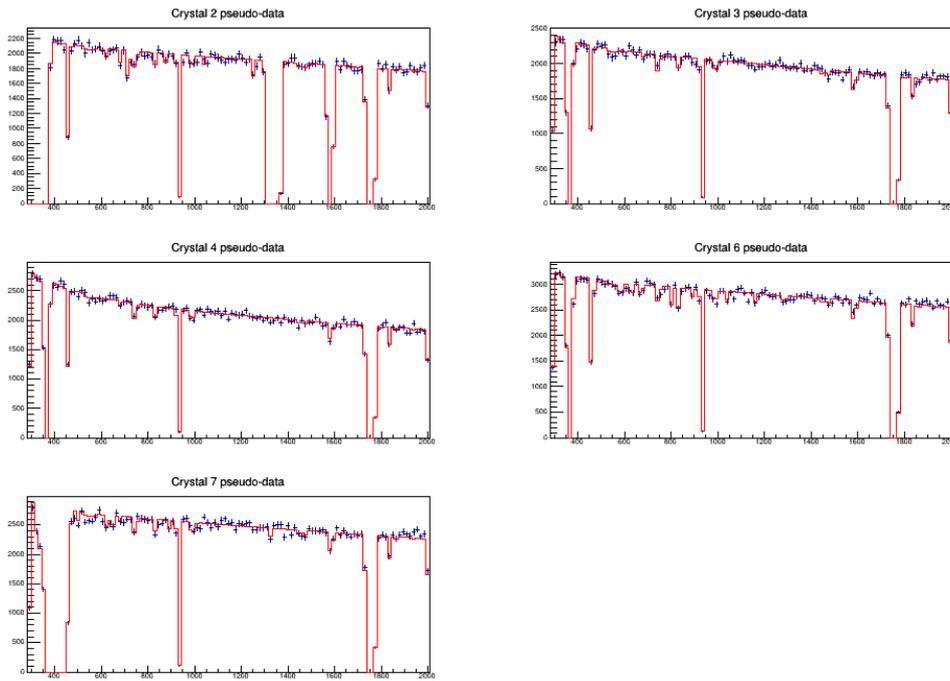


Figure 8.2: The example pseudo-experiment with an injected modulation amplitude of zero, shown in Fig. 8.1 (blue data points) with the results of the fitting procedure (red line).

8.3.1 Modulation amplitude bias check

For any given ensemble of pseudo-data, best-fit modulation amplitudes are collected for all 1,000 pseudo-experiments with convergent fits and used to populate a histogram (for example Fig. 8.3). Fitting a Gaussian function to this histogram (using the standard fitting package in the ROOT Data Analysis Framework [112]) and finding its mean, then subtracting the injected modulation amplitude of the ensemble returns the “bias” of the fitting procedure for data with that annual modulation amplitude. The mean of this distribution therefore gives an estimate of the value that can be expected were the physics data to have a modulation amplitude equal to the injected amplitude of the pseudo-data. The bias then represents the degree of imperfection in the fit, which can potentially be corrected for. The uncertainty of this bias is given by the uncertainty on the mean of the Gaussian returned by the fitter. This process is carried out to find the bias for all ensembles of pseudo-experiments and the bias of each ensemble (with the earlier choice of start date) is presented visually in Fig. 8.4.

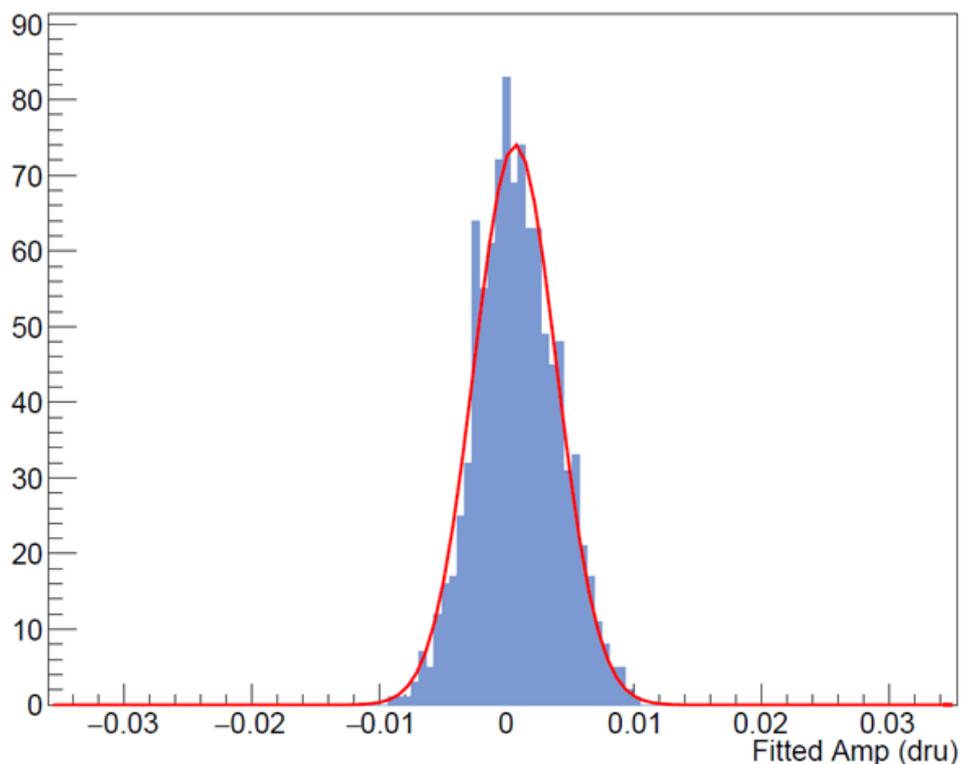


Figure 8.3: A histogram of the best-fit modulation amplitudes of an ensemble of 1,000 pseudo-experiments (blue bars) with an injected modulation amplitude of zero, fitted with a simple Gaussian (red line) to obtain the best-fit value expected by the COSINE-100 annual modulation fit method for data with the given injected modulation amplitude. The results shown are for an ensemble of pseudo-data with the earlier of the two start times described in Sec 8.2.

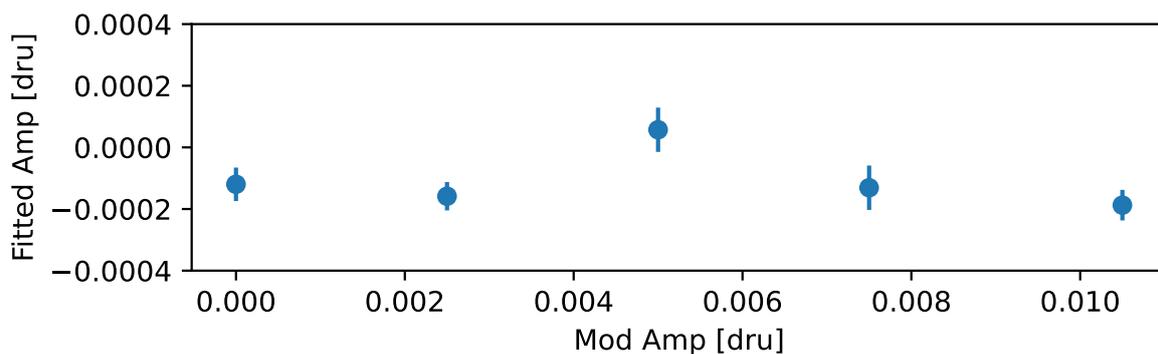


Figure 8.4: The bias on the modulation amplitude for five different ensembles of 1,000 pseudo-experiments with modulation amplitudes of 0, 0.0025, 0.005, 0.0075, and 0.0105 dru, as determined from the difference between the mean of a Gaussian function fitted to a histogram containing the best-fit modulation amplitudes for that ensemble (e.g. Fig. 8.3) and the injected modulation amplitude. The results shown are for an ensemble of pseudo-data with the earlier of the two start times described in Sec 8.2.

In the case where the annual modulation search procedure exactly returns the injected modulation amplitude, the bias calculated will equal zero. However, as can be seen in Fig.

8.4, a negative bias of between -2×10^{-4} dru and 0 dru is obtained in four of the five ensembles. Given that this bias is over an order of magnitude lower than the expected uncertainty on the best-fit value of the modulation amplitude, it is safe to ignore this bias. Also note that this bias is very similar to that reported in the pseudo-data analysis of the SET3 annual modulation search [1]. A similar result is found when using the alternative binning described in 8.2, implying the choice of binning has no impact on the best-fit values as expected. This bias check confirms that the SET4 fit procedure described in Section 7.2 is able to detect annual modulation in COSINE-100 data with a high degree of accuracy. As such, the main goal of the pseudo-data validation of the analysis technique is achieved.

8.3.2 Modulation amplitude uncertainty check

In addition to the best-fit modulation amplitudes themselves, the uncertainty on each best-fit modulation amplitude is also saved for each pseudo-experiment and entered into a histogram which is fitted in the same way as the distribution of best-fit modulation amplitudes to get an estimate of the uncertainty that can be expected on the best-fit modulation values. The mean of this distribution is 3.1×10^{-3} dru with a very small uncertainty for all ensembles tested, suggesting with relatively high certainty the uncertainty expected for the best-fit modulation amplitude of the SET4 physics data will be close to this value. This value also very closely matches the fitted width of the distribution of the modulation amplitudes shown in Fig. 8.3, further suggesting that 3.1×10^{-3} dru is roughly the uncertainty that can be expected from the analysis of the 1-6 keV physics data. This is an improvement on the 4.1×10^{-3} dru uncertainty reported in the SET3 analysis for the same ROI [1].

It should also be noted that were the best-fit modulation amplitude in the SET4 physics data to be measured as 0.0 dru with the above uncertainty, the DAMA signal would be excluded at the level of 3.4σ . The SET4 analysis therefore represents the first test of DAMA signal by COSINE-100 that would be capable of excluding the DAMA signal by over 3σ , were zero modulation to be measured.

The pull of each pseudo-experiment is also recorded where the generic and COSINE-100-annual-modulation definitions of pulls are

$$g = \frac{x - \mu}{\sigma} = \frac{S_{fit} - S_{inj}}{\sigma_{fit}}, \quad (8.2)$$

where x is the measured value, in this case the best-fit value for the modulation amplitude S_{fit} , μ is the expected value for this parameter or the injected modulation amplitude S_{inj} , and σ_{fit} is the uncertainty of the best-fit modulation amplitude, given by the width of the posterior distribution of S_{fit} . As with the biases, a histogram is drawn with the pull values and this distribution is fit to a Gaussian. These distributions have a very similar form to that of the best-fit modulation amplitudes shown in Fig. 8.3.

The mean of this pull distribution will clearly be proportional to the mean of the bias distribution since the pull of a given pseudo-experiment can be thought of as its bias divided by its uncertainty. The width of the pull distribution (referred to as the ‘‘pull RMS’’) however, reveals more information. In the desired case where the fit results are well matched to the given inputs, the pull RMS will have a mean of zero and unit width. However, if the value of the pull RMS is significantly above or below one then it could imply that the uncertainties on the best-fit modulation amplitudes are under and over-estimated respectively. The pull RMS values for the same five pseudo-data ensembles

as previously shown are shown in Fig. 8.5, where again the error bars represent the uncertainty on the fitted value of the width of the Gaussian fitted to the distribution. As with the bias check, it was found that the choice of binning with regards to start date does not have a significant effect on the measured values.

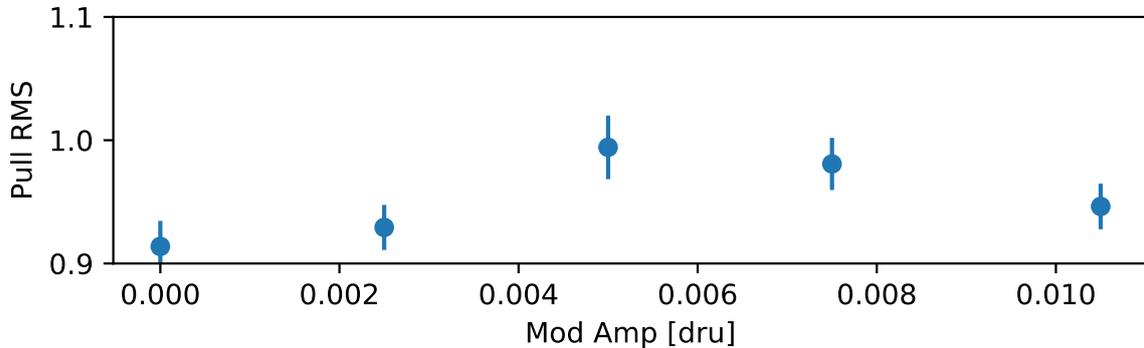


Figure 8.5: The pull RMS (defined in the text) values for the same five ensembles of pseudo-data as shown in Fig. 8.3.

Here the pull RMS values are slightly lower than the expected value of one, potentially suggesting an overestimation of the uncertainty of the best-fit modulation amplitude. Since the values of the pull RMS are not a significant distance away from the expected value of one, the uncertainty on the modulation amplitude is therefore not changed as a result of this finding.

8.3.3 Goodness of fit check with chi-square values

In addition to generating distributions of modulation amplitudes, pulls, and best-fit uncertainties for each ensemble of pseudo-data, the chi-square (χ^2) value for each fit to a pseudo-experiment is also calculated to measure the goodness of each fit. The χ^2 value is calculated as

$$\chi^2 = \sum_k \sum_i \frac{(O_i - E_i)^2}{\sigma_i^2}, \quad (8.3)$$

where O_i and E_i are the rate in the i^{th} time bin returned by the fitter and the true rate for that the pseudo-experiment respectively. These values are summed across all time bins for a single crystal with index k , then the total values from each crystal are also summed. More specifically, the O_i values correspond to a histogram with the same time binning as the input pseudo-data, populated by inputting the relevant time and best-fit parameters into the fit function. This histogram is generated in the fitting procedure and an example can be seen in the red curve in Fig. 8.2.

As with the distributions of the modulation amplitudes, pulls, and best-fit uncertainties, the distribution of the calculated χ^2 values are expected to be well-approximated by a Gaussian-like distribution, given the large number of degrees of freedom, and therefore give an estimate for the χ^2 value obtained from the analysis of the physics data. This follows the process carried out in past annual modulation analyses at COSINE-100. However, in this case with the new SET4 fitting procedure that includes the live-time scaling and a split surface and bulk ^{210}Pb component, a symmetrical Gaussian distribution is not

obtained when plotting a histogram of the χ^2 values of an ensemble of pseudo-data as expected. Instead, as shown in Fig. 8.6, the distribution is right-skewed with proportionately more high χ^2 values than low. This shape of distribution is found to be typical for the χ^2 s of all the ensembles of pseudo-data analysed using the method described at the beginning of this section, that is including the live-time and split surface and bulk ^{210}Pb component modifications. Also note that the upper limit of the horizontal axis in Fig. 8.6 has been limited at 1000 for the sake of clarity, however, a handful of higher χ^2 values were also measured.

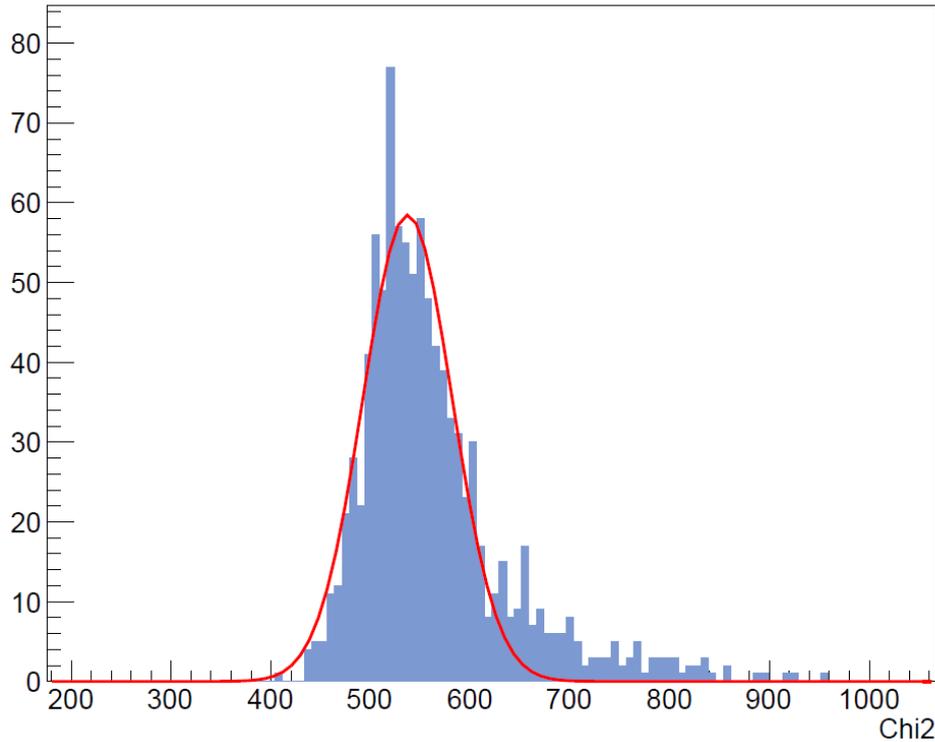


Figure 8.6: The typical distribution of chi-square values (blue bars) for the annual modulation fits to an ensemble of pseudo-data when a split surface and bulk ^{210}Pb component is used. This distribution is fit to a Gaussian function (red line).

The high proportion of pseudo-experiments with fits with high χ^2 values when using the new split- ^{210}Pb analysis method is indicative of an issue with this new fit method, and the source of this issue should be identified.

8.4 Investigating the issue with the reproduction of background activities in the SET4 fitting method

Once it was established that the introduction of the split- ^{210}Pb component in the fit model was the cause of the unexpected shape of the distribution of χ^2 values, attention turned to finding which specific element of the fit to the pseudo-data had performed poorly and was therefore causing high χ^2 values in the fits to some of the pseudo-experiments. This was done first by looking at the best-fit function and values of the initial activities in specific pseudo-experiments with high- χ^2 fits, then at the best-fit values for the half-lives of the surface ^{210}Pb components across a full ensemble of pseudo-data. Then the efficacy

of the fitter was testing in terms of producing the correct surface ^{210}Pb half-life. These steps are described in turn, then finally the interpretations of the findings are discussed.

8.4.1 Locating the source of the issue in the fit method

Since this right-skewed distribution of χ^2 values was not observed prior to the SET4 analysis, it follows that one of the additions to the analysis method, either the live-time scaling or the splitting of the ^{210}Pb component has caused the issue. To test this, the same ensembles of pseudo-data as described in Section 8.2 were re-analysed with fitting methods that in one case used the old method of accounting for dead-time in the data but still used a split ^{210}Pb component, and another where only a single ^{210}Pb component was considered but the new live-time scaling method was used. This check showed that when only a single ^{210}Pb component was used, the distribution of χ^2 values for a given ensemble of pseudo-data followed the expected Gaussian shape as shown in Fig. 8.7, whereas no change was observed in changing the live-time scaling method, therefore indicating an issue in the updated treatment of the ^{210}Pb component where a significant proportion of pseudo-experiments are being poorly fitted.

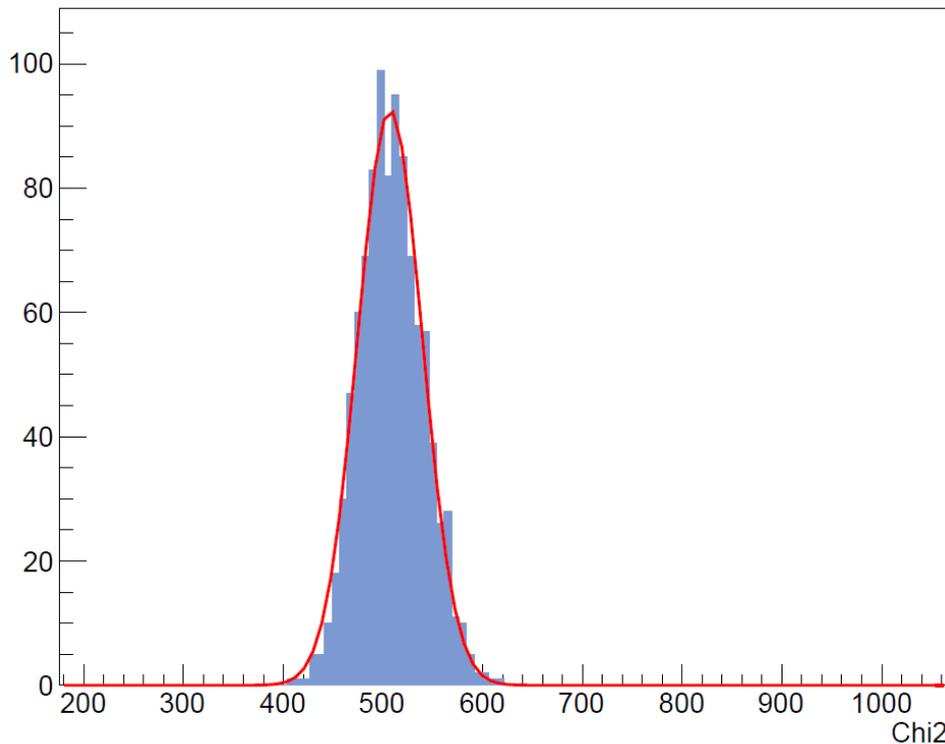


Figure 8.7: The distribution of chi-square values for the same ensemble of pseudo-data as shown in Fig. 8.6, but fit using a model that only considers a single ^{210}Pb component.

As it seems that high χ^2 values causing the skewed distribution shown in Fig. 8.6 result from the new splitting of the ^{210}Pb component in the fit model, it is useful to briefly restate what this alteration to the fit model (described in Section 7.2) entails. In studies of the best-fit background activities measured is the SET3 annual modulation analysis, it was discovered that the ^{210}Pb activity was consistently over-fitted and consequently the ^3H activity being under-fitted to compensate. This was investigated and the likely cause was found to be ^{210}Pb on the crystal surface getting replenished over time by the decay

products of ^{222}Rn decay, effectively extending the half-life of the surface ^{210}Pb . To account for this, the ^{210}Pb component was split into separate bulk and surface components. The bulk component is treated in the same way as the previous combined ^{210}Pb component but with a lower initial activity, whereas since the effective half-life of the surface ^{210}Pb is unknown, this value is allowed to float in the fit.

As shown in Sections. 8.3.1 and 8.3.2, the best-fit values related to the modulation component are returned with the expected values, it is therefore anticipated that the issue causing the high χ^2 fits lies with the fitting of the background components, rather than the modulation components. To ensure the choice of fitting method did not have a large influence on the main parameters of interest in the modulation search, the process of calculating the biases and pulls from the fits to the pseudo-data ensembles was repeated for the fit results using the single- ^{210}Pb framework. A comparison of the observed bias for the two methods is shown in Fig. 8.8. Here it can be seen that the single- ^{210}Pb model fit consistently returns a slightly higher value for the best-fit modulation value than the split- ^{210}Pb model by around 10^{-4} dru. This indicates the split- ^{210}Pb model has a very slight negative bias relative to the single- ^{210}Pb model, but this is sufficiently small as to be ignored.

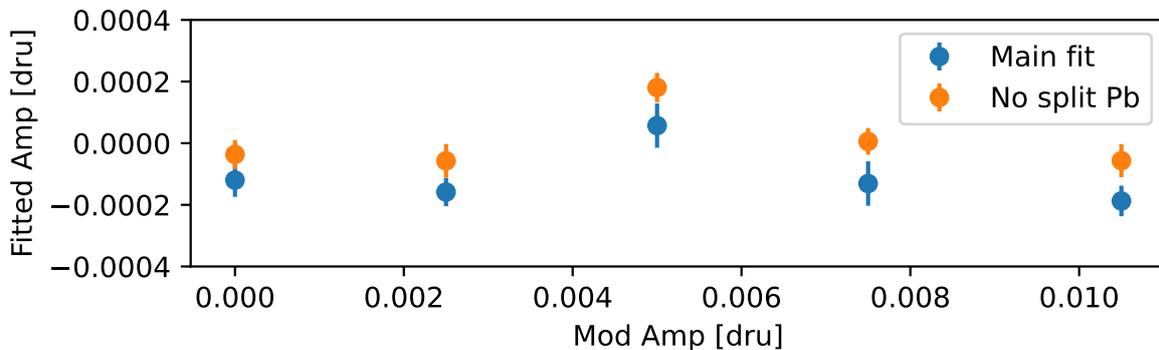


Figure 8.8: The biases of fit results to the same five ensembles of pseudo-data described in Section 8.2 using a fit model with and without (blue and orange points) a split ^{210}Pb component.

It is also useful to test the performance of the two fitting procedures on an experiment-by-experiment level, and therefore get an indication of the degree by which the best-fit modulation parameters would be expected to vary between methods were they applied to the physics data. This was done by finding the difference between the best-fit modulation amplitudes from each method in every pseudo-experiment in an ensemble. The resulting distribution of differences is shown for an ensemble of pseudo-data with zero injected amplitude in Fig. 8.9 and where a positive value represents the case where the value from the split- ^{210}Pb method is higher than from the single- ^{210}Pb method. A Gaussian has been fitted to the distribution as in previous examples and is a good description of the distribution. The means and standard deviations across the five ensembles of these Gaussians are consistently around -1×10^{-4} dru and 5×10^{-4} dru respectively.

This small, negative bias on the best-fit modulation amplitude value in the split- ^{210}Pb model relative to the single- ^{210}Pb model is expected from the previous bias test shown in Fig. 8.8. These distributions of the differences between the best-fit modulation amplitudes give an indication of the maximum difference that can be expected if the physics data is analysed using both models. Were the values produced in such an analysis to vary outside

the distribution shown in Fig. 8.9, it would be indicative of an issue in the procedure.

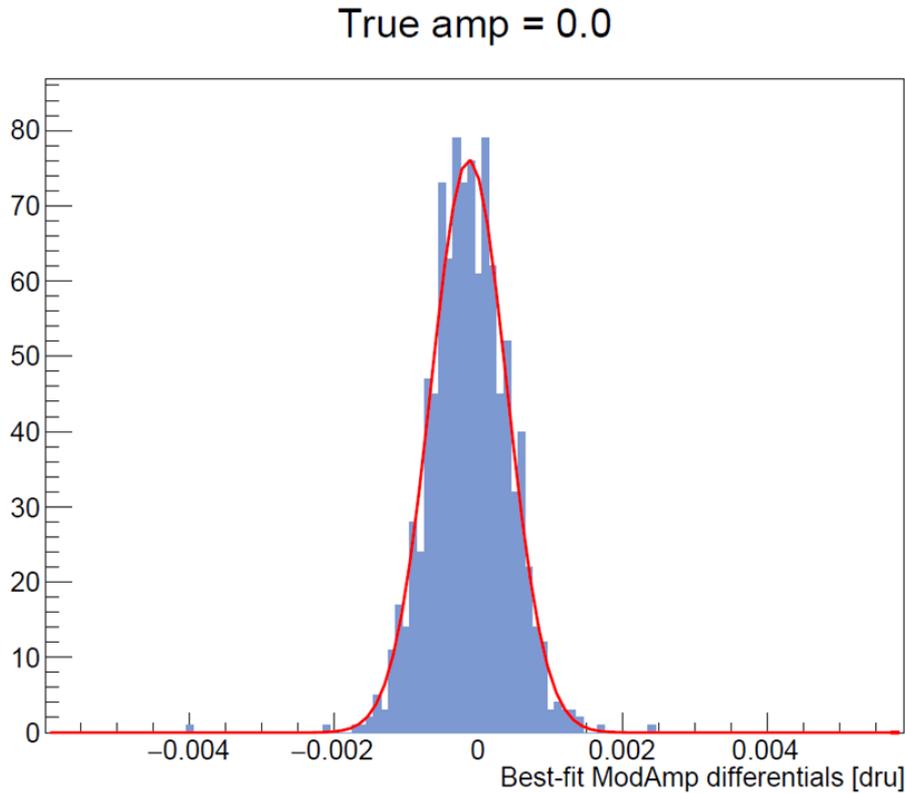


Figure 8.9: The distribution of the differences between the best-fit modulation amplitudes outputted by the single and split ^{210}Pb fit model for each data-set in an ensemble of 1,000 pseudo-experiments (blue bars) with zero injected modulation, shown as a representative example. Negative values correspond to a lower best-fit value in the split- ^{210}Pb analysis. A Gaussian has been fit to this distribution (red line).

Since the expected behaviour in the χ^2 values was returned using the same pseudo-data and analysis framework (including calculating the χ^2 value, producing the above plots), aside from the difference in the treatment of the ^{210}Pb component in the fitting procedure, it seems unlikely that the issue has come from the alterations in the production of the pseudo-data. It is perhaps possible that some issue does exist with the production of the pseudo-data which in combination with the fit model being somehow more sensitive to such issues as a result of splitting the ^{210}Pb component that has caused the high χ^2 values, but there is no specific reason to think this should be the case.

Furthermore, the values used for the activities and extended half-life of the surface and bulk ^{210}Pb components in generating the pseudo-data are the same as the means of the prior distributions used in the fit model for these parameters. Since the fit model to this pseudo-data generally performs well even when only a single ^{210}Pb component is assumed, it is less likely that it is this choice of prior distributions that is causing the issue in the split ^{210}Pb fit model. This is because were the prior distributions for the split ^{210}Pb components a poor description of the data, the pseudo-data generated using these same values for the split ^{210}Pb components would likely be poorly fit using the fit model that only uses a single ^{210}Pb component, and therefore this fit model would likely perform worse on this sample of pseudo-data than observed.

8.4.2 Tests of individual high- χ^2 fits

As an initial test of this issue, a selection of pseudo-experiments with high χ^2 values was selected and the event rates and fit function plotted to give a visual indication of the potential origin of the high χ^2 value. A representative example pseudo-experiment with a χ^2 value of 1060 associated with its fit is shown in Fig. 8.10. This example is typical of the fit functions with high- χ^2 values, inasmuch as typically in these cases three or four of the crystals have fit functions that visually agree well with the data, then the other one or two fit functions show a consistently high or low rate relative to the pseudo-data. In the case where there are poor fits in multiple crystals there appears to be no relationship between the relative sizes of the poor fits. This is expected given the activities of each crystal are uncorrelated in the fit model.

Since a consistent under/over-fitting of the data in a pseudo-experiment is typically indicative of an under/over-estimation of the initial activities of the background components in the fitting process, the best fit-values of the initial activities for the background components were plotted relative to the post-fluctuation input activity used to generate that pseudo-experiment. These best-fit initial activities for the fits to four of the five crystals in the pseudo-experiment in Fig. 8.10 are shown in Fig. 8.11.

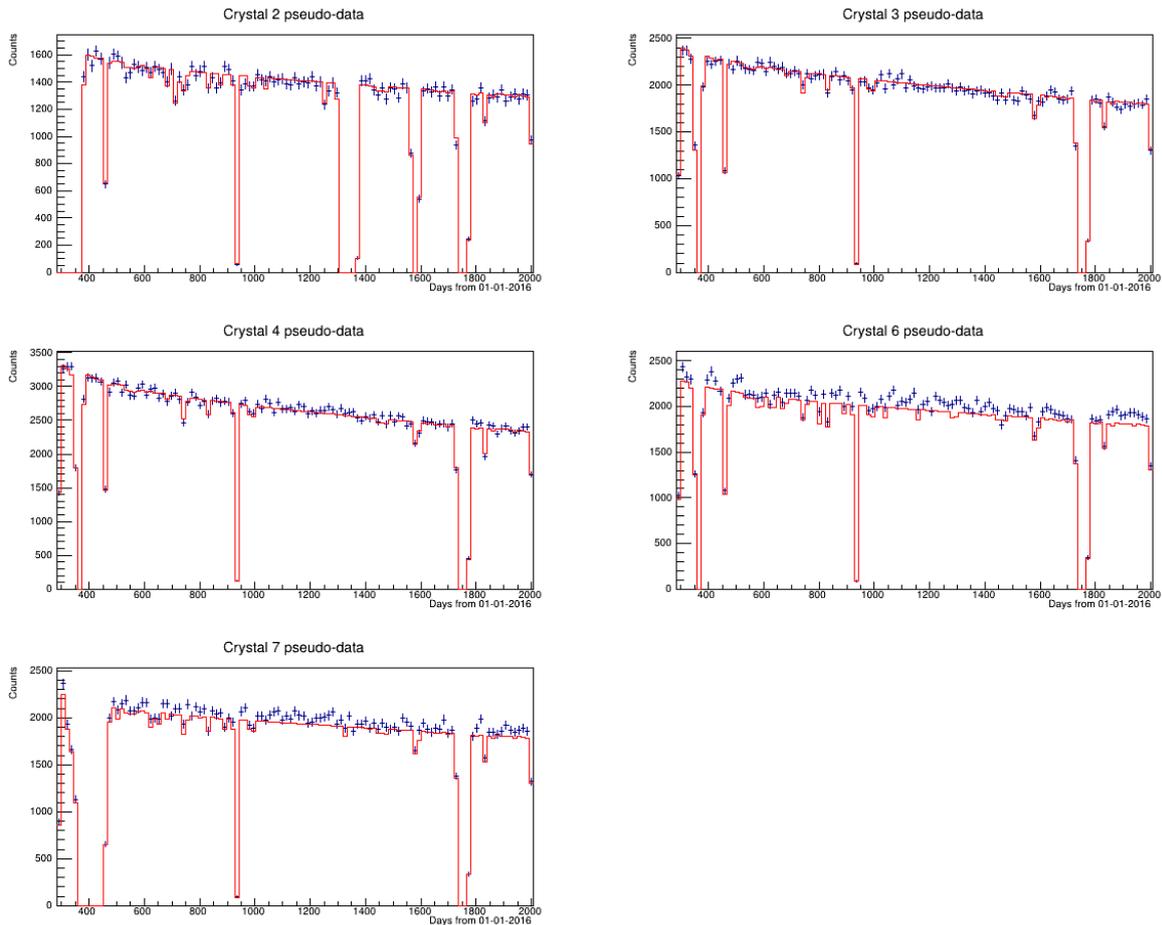


Figure 8.10: An example of a pseudo-experiment with a high chi-square value associated with the annual modulation fit (red line) to the pseudo-data (blue data points). Here there is an under-fitting relative to the data in C6 and C7, however, both under and over-fitting is observed in all crystals in different pseudo-experiments.

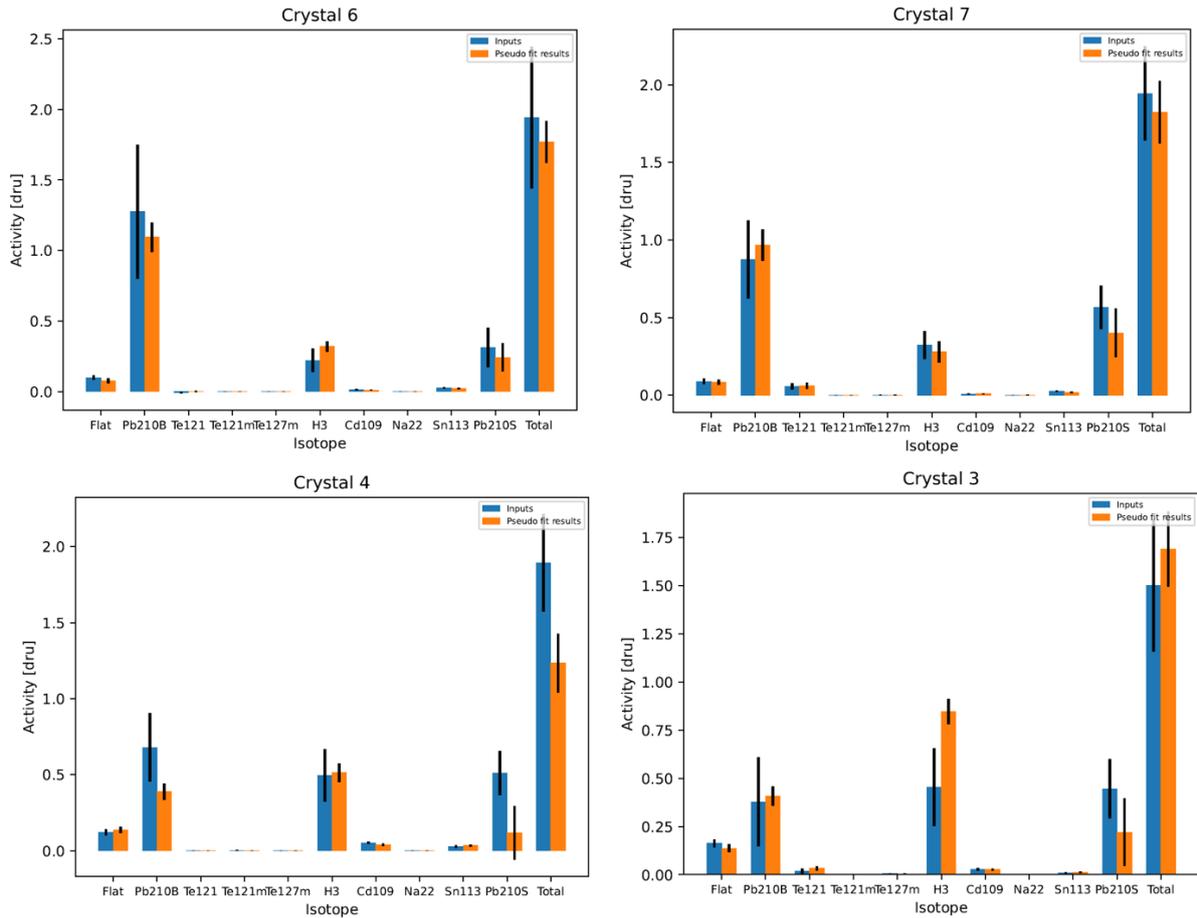


Figure 8.11: The best-fit initial activities in four of the crystals in the pseudo-experiment shown in Fig. 8.10 (orange bars) compared to the input values used in the generation of the pseudo-experiment (blue bars). These best-fit activities are used in drawing the best-fit functions in Fig. 8.10. Rather than the total activity being correlated with the observed relationship between the data and fit function shown in Fig. 8.10 for each crystal, there is seemingly no relationship between these values. The error bars show the measured uncertainty on the best-fit activities.

In the first two crystals shown in Fig. 8.11, in C6 and C7, the two crystals with the under-fitted event-rate in this pseudo-experiment, the sum of the fitted activities is shown to be low relative to the input as expected from the fit functions shown in 8.10. However, rather than the best-fit values for the background activities showing good agreement with the inputs as would be expected from the visually good fits in Fig. 8.10, the totals of the initial activities in these crystals is shown to deviate more from the input values than in the crystals where a visually poor fit is observed. In other words, at some point in the fitting and plotting procedure, either when the fit function is generated as in Fig. 8.10, or when the best-fit background activities are extracted for Fig. 8.11, an error has occurred since these two results should show the same behaviour.

The potential causes of these issues were examined as the subsequent results are discussed in more detail in Appendix. A, mainly focusing on an investigation into the performance of the fit of the surface ^{210}Pb parameters. It should be noted that during this investigation, a bug was found in the script used to carry out the fitting procedure. The effects of this bug were to exclude the new surface ^{210}Pb component from the process of modelling the inputted data, but while still producing a posterior distribution for the

parameter with realistic looking deviations from the values from the prior distributions, thus masking the issue. When this bug was removed, and the true efficacy of returning the surface ^{210}Pb parameters was revealed and found to be worse than initially thought. This bug was present in the initial testing of the split- ^{210}Pb method on SET3 data-set, but was not noticed at the time, hence the conclusions reached in Chapter 7.

Despite this issue, in all cases the best-fit annual modulation parameters returned matched the expected values, either from injected values in the case of pseudo-data analysis or the published SET3 value in the tests using the SET3 data-set. Since the primary goal of the fit model, that is correctly returning the injected annual modulation amplitude of the pseudo-data, is achieved, the new split- ^{210}Pb model will be used to analyse physics data with a single- ^{210}Pb model used to check these results.

8.5 A novel use of pseudo-data in reporting the floated-phase annual modulation search result

An issue that had to be contended with in the floated-phase annual modulation searches in data from both the COSINE-100 [1] and ANAIS [71] experiments is the tendency in such searches for a modulation amplitude greater than zero to naturally be reported in sample of data with no true modulation. This bias arises from the co-ordinates used in floated phase searches. Searches for a floated-phase annual modulation signal are most naturally carried out in polar co-ordinates in amplitude and phase, however, as discussed in more quantitative detail in Appendix A of [71], searches in this co-ordinate system are predisposed to finding non-zero values for the modulation amplitude.

In simple terms, this is because the fitter will always “try”¹ to find the most probable phase for modulation in the data, and since it is highly unlikely due to random fluctuations that all phases are found to have zero modulation, even in the case of data with no true modulation, a small modulation amplitude at the most probable phase is likely to be returned. It is possible to counter this effect by converting to a Cartesian co-ordinate system, however if this approach is chosen then the results are reported in unintuitive, unphysical units so this approach is disfavoured.

8.5.1 Measuring the modulation amplitude bias in floated-phase analyses of low modulation pseudo-data

To measure the above effect in COSINE-100 pseudo-data, the same ensemble of 1,000 pseudo-experiments with zero modulation amplitude as described in the previous sections was re-analysed using a floated-phase fit model and the distribution of the best-fit value of the modulation amplitude plotted. This is shown in comparison with the same results from the ensemble with an injected modulation amplitude and phase equal to 0.0105 dru and 152.5 days in Fig. 8.12. The values from split- ^{210}Pb model are shown, but similar results are obtained from the single- ^{210}Pb model. For reference, the accompanying distributions for the best-fit values for the phase in these two ensembles of pseudo-data are also shown in Fig. 8.13. The amplitude distribution for the 0.0105 dru pseudo-data and phase distributions for both pseudo-data ensembles have been fit to a Gaussian.

¹In reality the fitter does not “try” to do anything, but its behaviour is described in these terms for the sake of analogy.

As expected, best-fit modulation amplitudes of zero are not returned in the zero modulation pseudo-data and instead small positive values are returned. The distribution of best-fit modulation amplitudes for the ensemble with the DAMA-magnitude injected modulation is well described by a Gaussian. The bias calculated from the mean of this distribution is $-6 \pm 1 \times 10^{-4}$ dru, slightly larger in magnitude than the biases reported in the fixed phase case, but still small relative to the expected uncertainty of the value of around -4×10^{-3} dru.

The distributions of the best-fit phases are Gaussian in both cases with means and standard deviations of 184.0 ± 2.0 days and 49.0 ± 1.5 days in the no modulation case, and 150.8 ± 0.5 days and 19.6 ± 0.5 days in the 0.0105 dru modulation case. It should be noted that due to the polar nature of the co-ordinates used in generating and fitting the pseudo-data, the inputted phase of the zero-modulation pseudo-data is undefined rather than zero, hence distribution of phases being centred around approximately $365/2$ days rather than 0 days. The values reported for the best-fit phases in the 0.0105 dru modulation amplitude pseudo-data are in good agreement with the inputted value of 152.5 days.

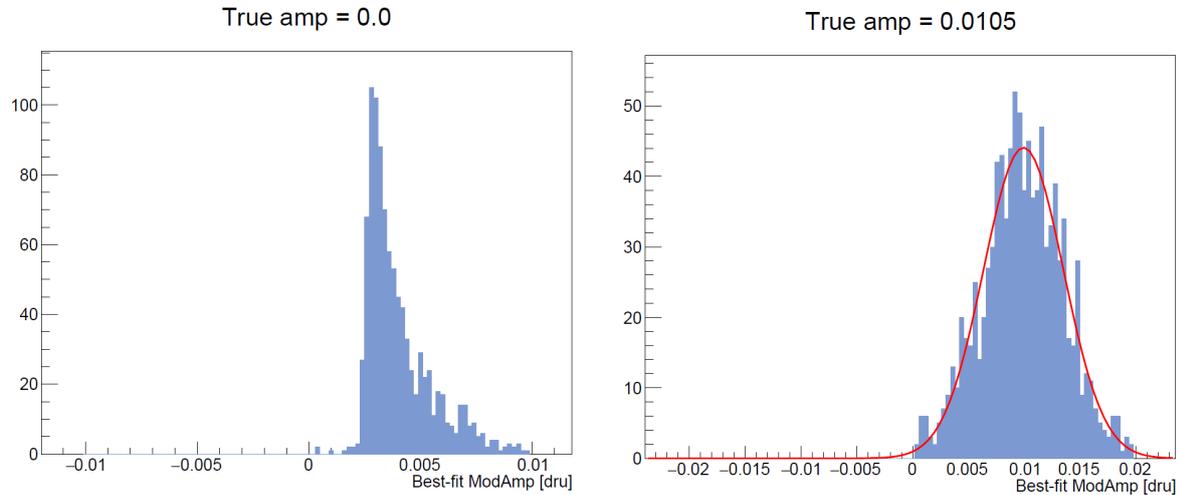


Figure 8.12: The distributions of the best-fit modulation amplitudes using a floated-phase fit model for each of the 1,000 pseudo-experiments in ensembles with injected modulation amplitudes of 0 (left) and 0.0105 dru (right), in line with the null and DAMA modulation hypotheses. The right distribution has been fit to a Gaussian and centres around the expected value of 0.0105 dru whereas in the no-modulation case, a positive bias is observed due to the co-ordinates used in the fitting.

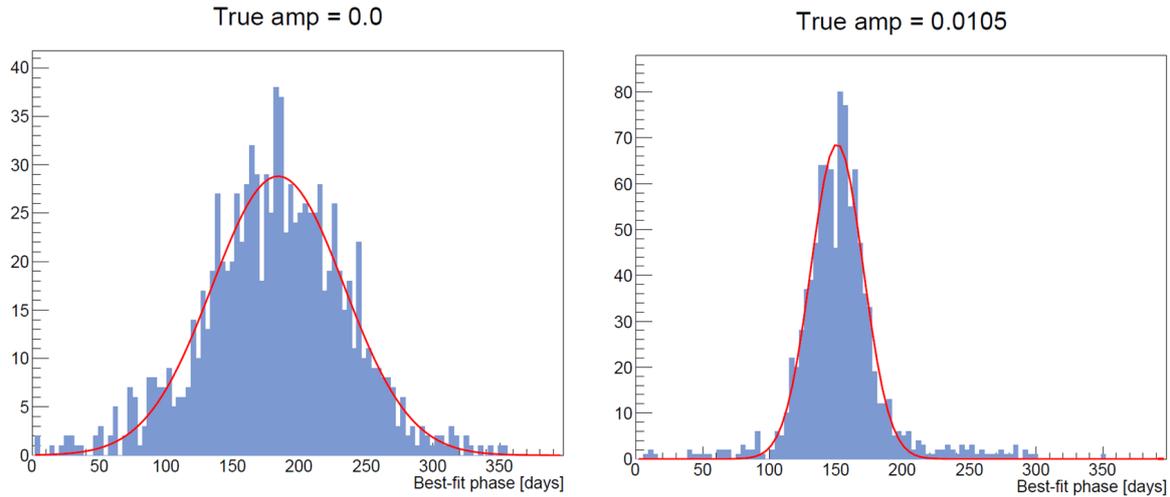


Figure 8.13: The distribution of best-fit values for the floated-phase analysis of pseudo-data described in Fig. 8.12. Both distributions have been fit to a Gaussian.

8.5.2 Accounting for the floated-phase low-amplitude bias

Attempting to deal with this apparent bias by quantifying it and applying a correction to the best-fit modulation amplitude accordingly is non-trivial and potentially unintuitive, given that the bias does not affect data at all phases equally, as well as the two-dimensional nature of the results. In the SET3 annual modulation analysis at COSINE-100, the best-fit modulation amplitude was not directly reported for the floated-phase analysis. Instead, the approach taken was to calculate the median of the best-fit modulation amplitudes in an ensemble of pseudo-data with zero injected modulation amplitude (the equivalent of calculating the median of the left plot in Fig. 8.12) and observe that the value was in agreement with the best-fit value measured in the physics data. The median rather than mean was used due to the asymmetrical nature of the distribution.

A novel approach for presenting the results of the floated-phase analysis while accounting for this bias has been developed for the SET4. Two-dimensional histograms are generated from the best-fit amplitudes and phases (calculated from their respective 1D marginalised posterior distributions) of ensembles of pseudo-data with injected modulation amplitudes of 0 and 0.0105 dru (note that DAMA also report an annual modulation amplitude of 0.0105 dru in their floated-phase analysis [49]). This is shown in Fig. 8.14 using the split- ^{210}Pb fit method where the median and mean of the amplitudes and phases respectively are also shown. Similar plots for the single- ^{210}Pb analysis method are also generated.

To visualise the degree to which the results from the SET4 floated-phase analysis, the 68.3%, 95.5%, and 99.7% highest-density credible regions (HDRs) calculated from the 2D amplitude/phase posterior distribution can be overlaid onto the 2D distributions of amplitudes and phases for different ensembles of pseudo-data. In doing so, the agreement between the physics data and the no-modulation and DAMA-modulation hypotheses can clearly be seen and visually compared. It should also be possible to integrate the 2D histogram of pseudo-data results within the bounds of the contours drawn from the results of the analysis of the physics data to quantify the degree of agreement between the pseudo and physics data. The machinery for this calculation has not yet been developed however.

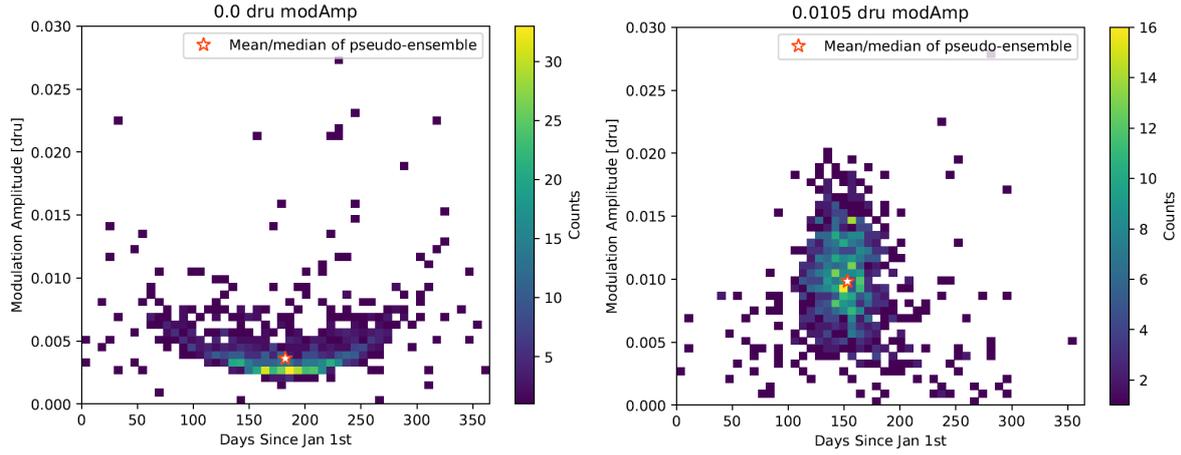


Figure 8.14: Two-dimensional histograms filled with the best-fit values (from the marginalised one-dimensional posterior distribution of the parameter) of the modulation amplitudes and phases found when the split- ^{210}Pb fit is applied to ensembles of 1,000 pseudo-experiments with injected modulation amplitudes of 0.0 dru (left), 0.0105 (right). The red and white star represents the location of the median and mean respectively of the amplitudes and phases.

8.6 Concluding remarks

Moving forward, the source of the poor performance of the fit in reporting the background components in the new fitting technique should be further investigated and rectified before the final SET4 analysis is carried out. Note that this final analysis is also contingent on some as yet incomplete updates to the event selection, designed to lower the low energy threshold, beyond the scope of this work. It seems most likely that rather than a split ^{210}Pb component being an inherently poor description of the COSINE-100 data, there is instead some procedural error introduced during the splitting of the ^{210}Pb component affecting the efficacy of the fitter in terms of correctly describing the background activities of the data, and possibly also in producing some of the plots used to test the procedure. This issue may be related to the best-fit value of the half-life being unconstrained and therefore unphysical values being returned from the fitting procedure. In future, a version of the fitting procedure with the half-life of the surface ^{210}Pb constrained to physical values should be tested.

In terms of this work, it has been shown in the analysis above in Section 8.3.1 that while issues may exist with describing the background components of the data, the injected annual modulation components injected into the pseudo-data are accurately returned. It should also be stressed that no information about the injected modulation amplitude of a ensemble of pseudo-data is inputted into the model when doing these fits. Therefore this fitting procedure is still a valid test of whether annual modulation exists in the COSINE-100 physics data. The same is also true when the single- ^{210}Pb fit model is used. Consequently both these procedures will be used in the following chapters of this work to analyse the first 4.7 years of COSINE-100 physics data and hence give a new test of the DAMA signal.

Chapter 9

Searches for Annual Modulation in the SET4 COSINE-100 Physics Data

As has been laid out in the previous sections, an analysis procedure has been developed in order to search for an annual modulation signal in the first 4.7 years (SET4) of COSINE-100 hundred single-hit data from the 20th of October 2016 to the 29th of June 2021. 61.3 kg of active NaI(Tl) scintillation material is used, giving a total exposure of 261 kg×years. Such a search represents the most up-to-date and stringent test of the positive annual modulation signal reported in the DAMA experiments and attributed to dark matter by the DAMA collaboration [49, 66, 68, 70], expanding on the already published modulation searches with 1.7 and 2.8 years of COSINE-100 data. [1, 75].

First in this chapter checks of the detector stability are described in Section 9.1, after which the event selection and data preparation is described, focusing on the rejection of low-energy noise events in Section 9.2.1 and the procedure of selecting “good” data in Section 9.2.2. Then follows discussions of the results of annual modulation searches in these data, first using a fixed-phase annual modulation model in Section 9.3, then a floated-phase model in Section 9.4.

9.1 Detector stability

Since the signal searched for at COSINE-100 is a modulation with a very small amplitude ($\mathcal{O}(0.1\%)$ above the background rate) over multi-year timescales, ensuring the detector performance is stable over these timescales is crucial, as any unaccounted for instability could manifest as an apparent change in event rate and hence mask or mimic the small annual modulation that is being searched for. As described in Section 4.3.1, both environmental and physics parameters were continuously monitored during the operation of the experiment, and any unexpected behaviour was reported for further study and potential correction. Likewise, the performance of different detector components was continuously monitored. In addition to this ongoing monitoring, further checks of detector stability were also carried out prior to the final annual modulation analysis to further ensure the requisite levels of stability. These efforts to ensure stability and the resultant findings are outlined as followed

9.1.1 Ensuring consistent detector performance

As described in Section 4.3.1, the performance of detector components was continually tested to identify and correct for any changes in performance over time. In the case

of the measured response at sub-100 keV energies of the NaI(Tl) crystals, this process entails tracking the location of the 50 keV peak from ^{210}Pb decays over time and applying corrections to the light output to keV calibration factor to account for any changes in the behaviour. This process is described in more detail in Section 4.3.3. In order to verify the efficacy of this approach in the energy region of interest (ROI) of this study, a further test of the stability of the performance of the crystal PMTs was carried out.

The decay of ^{40}K into ^{40}Ar produces a dual 1460 keV gamma and 3.2 keV X-ray photon; the former can be measured in the liquid (LS) scintillator at COSINE-100 (and is used in the long-term monitoring and calibration of the LS, as described in Chapter 5). The latter is measured in the crystal PMTs and is a major source of multi-hit background events in the crystals in the 1-6 keV ROI. Utilising the joint nature of the 1460 keV and 3.2 keV signals in conjunction with the standard event selection for signal-like events in the crystals (described in more details in Section 9.2), a sample of low-energy NaI events with an energy spectrum with a clear peak from these 3.2 keV ^{40}K decays is obtained. Specifically, the sample consists of crystal events in coincidence with a signal in the LS with an energy between 900 keV and 1,550 keV. A wide spread of LS energies is used due to the broadening of the 1460 keV peak to a Compton shoulder due to Compton scattering.

The resultant crystal energy spectra can be fit (as shown in Fig. 9.1) to further monitor the measured behaviour of the crystals in the energy ROI. Repeating this over the entire data-set gives an indication of the crystal response over time and hence the efficacy of the gain correction. The function used to parameterise the 3.2 keV peak is phenomenological and consists of a Gaussian plus a straight line where the location of the Gaussian peak is used to characterise the location of the peak. The data from each crystal is split into 400 subrun (800 hour) bins and then the resultant energy spectrum of each crystal has the above fit function applied. The location of the peaks in all five crystals across the SET4 data taking period are shown in Fig. 9.2. The dotted lines in Fig. 9.2 represent the means of the fitted locations of the 3.2 keV peak for that crystal and the shaded areas represent one standard deviation either side of this mean.

In all cases there is good agreement between the measured location of the 3.2 keV peak and the expected value. The average peak location in C3 is slightly below the expected value, but this is not expected to cause issues. Since C6 and C7 have the lowest levels of background ^{40}K their 3.2 keV peaks are made up of fewer events and consequently a larger uncertainty on the peak position is obtained. In some cases, the peak location in a given crystal appears to consistently rise or fall, most likely as a result of the continuous gain correction. To ensure this effect has not mistakenly manifested in a false annual modulation, a simple sinusoidal function with a fixed period of one year was fit to all crystals. In all cases the sinusoidal fit described the data significantly less well than a flat fit, hence it can be concluded that no annual modulation is inadvertently injected into the measurement of crystal events. Consequently, it can be stated in good confidence that the measured behaviour of the crystals is stable over the SET4 data-set.

9.1.2 Environmental monitoring

The full process of monitoring environmental variables within the detector is described in detail in Section 4.3.1. Ensuring the temperature inside the detector is kept constant is of particular importance since atmospheric temperature undergoes annual modulation with approximately the same phase as the DAMA signal. Thermometers in ten different locations were used to measure the temperature of the detector and surrounding area;

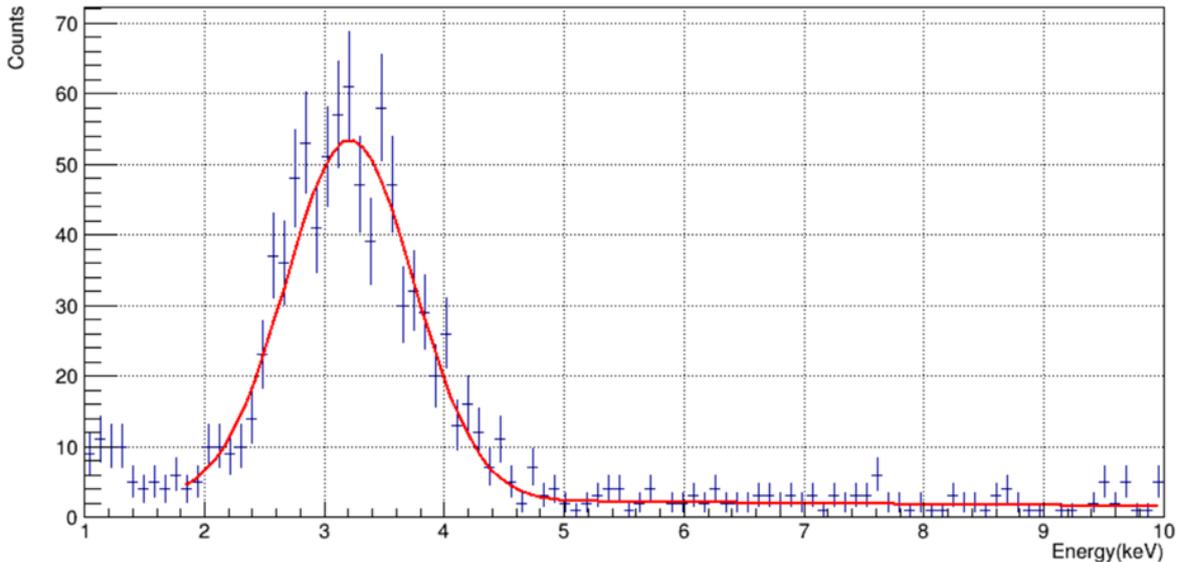


Figure 9.1: Energy spectrum of 3.2 keV events from the decay of ^{40}K measured in C2 from 1,200 hours of data taking (blue data points) and selected via coincidence with 1460 keV events in the LS. This peak is fit over time (red line) for each crystal to track the response of the crystals over time.

the most important of these is attached directly to the crystal encapsulation. The temperature measured by this thermometer was within the pre-defined acceptable range of $24.5 \pm 0.5^\circ \text{C}$ for the entirety of the SET4 data taking period with a single exception during a brief period of detector down-time (most likely a lower outage in March 2017). Generally the temperature of the inside of the detector was maintained at $24.3 \pm 0.1^\circ \text{C}$ and any deviations beyond this were reported and investigated, but no detector dead-time resulted from these reports. The values of other environmental parameters are similarly monitored to ensure optimal conditions for data taking, and as with the temperature, these parameters are stable over the SET4 data-taking period.

9.2 Event selection

The process of event selection, both in terms of data quality checks and noise rejection, is similar to that used in the SET3 annual modulation search [1], but with some alterations. This process will now be described in more detail, where first the data quality checks are described, followed by the event selection used for maximal noise rejection.

9.2.1 Event selection and noise rejection

A range of techniques are employed to obtain maximal noise rejection while retaining a large proportion of signal events. This process can be broken down into two separate stages, a “pre-selection” stage, and a “selection” stage. In the pre-selection stage, events which can obviously be rejected as noise are removed. The cuts used in this process are as follows.

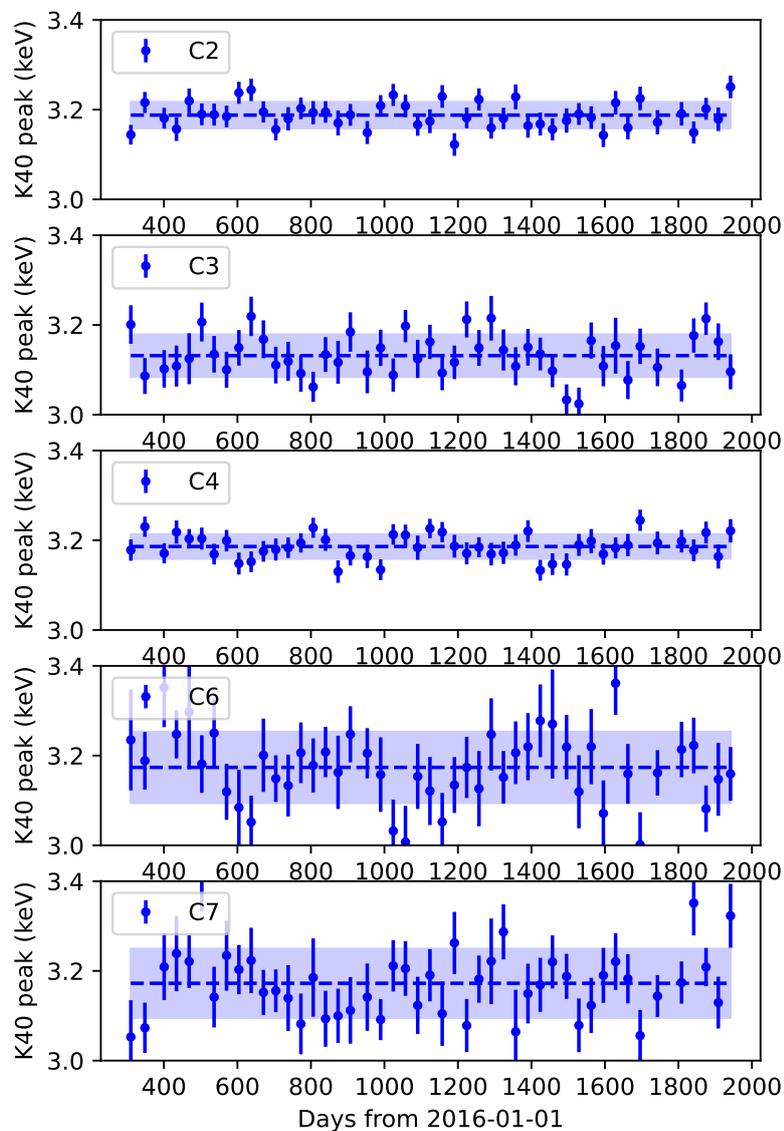


Figure 9.2: The location of the 3.2 X-ray peak (blue data points) from ^{40}K decay as measured in each of the five primary COSINE-100 NaI crystals in 400 subrun (800 hour) bins over the SET4 data-taking period. These events are selected in coincidence with a corresponding 1460 keV signal in the LS from the same decay. The means and standard deviations for each crystal are shown by the dashed lines and shaded areas respectively.

Selecting single-hit events

When searching for dark matter, any event that occurs in coincidence with another event can be rejected as noise due to the very low expected cross-section of interaction of dark matter-induced events, and therefore the very low probability of two dark matter-induced events occurring in coincidence. In practice at COSINE-100 there are three different types of detector that can induce a recordable signal, the liquid and plastic scintillator active shielding, described in Section 4.2, and the NaI crystals. When an event is recorded in one of the crystals, an 8 μs time window is opened with its centre at the time of the crystal event. If a signal above the relevant threshold from any of the other detectors is recorded during this window, then the event is marked as a “multi-hit” event and hence not used in dark matter searches.

The thresholds used for this cut depend on the type of detector, and in the case of crystal signals, when in the data-taking period the event was measured. The threshold for an event in the plastic scintillator muon veto is two plastic scintillator PMTs measuring a signal over 4,000 ADC (analogue-to-digital conversion) units in a 400 ns window. This threshold has remained constant throughout the life of the experiment. The threshold for the LS was initially set at 20 keV in early COSINE-100 analyses [2, 75], but was then increased to 80 keV in the SET3 analysis [1] due to uncertainty about the LS’ performance. This initial uncertainty was resolved, but due to the complications arising from re-calculating the ADC to keV conversion in the LS after removing the signals from one of its PMTs (described in Section 5.2.3), the higher, more conservative low energy threshold was retained.

In signals measured in the crystal PMTs, the threshold for a signal is not dependent only on the total energy deposited, but rather the number of clusters (effectively the number of peaks) in the signal. This is because signal events in PMTs such as the ones used to measure the COSINE-100 crystals are characterised by having many peaks, whereas noise events are more likely to be single-peaked. Four clusters in both PMTs of a crystal in an 8 μs time window must be recorded for a crystal event to be recorded.

This applies to all crystals across the data-taking period with the exception of C8 in the final part of SET4. In January 2021, around six months before the end of SET4 in June that year, the rate of PMT noise in one of the PMTs of C8 rose sharply such that it was no longer possible to use signals below 20 keV measured in the PMT after this point. Therefore the threshold for a signal in the crystal was changed to four clusters in the remaining functional PMT. To check what effect this change had on the final event rate, the event rate of single-hit events prior to this point was re-calculated with the updated event selection and compared to the case where a signal was required in both PMTs. This check showed that the event rate did not measurably change with the changed event selection in the earlier period, whereas an increased event rate was observed after this point if both of the C8 PMTs were used.

Rejecting PMT noise

Once single-hit events have been selected, obvious noise events with unphysical values for their start times and measured charge are rejected. Then follows a more comprehensive process of event selection utilising multivariate analysis in the form of a boosted decision tree (BDT) variable to reject PMT noise events. The process of developing this PMT-noise rejection through the various COSINE-100 analyses is described in full for context.

A significant source of noise in the detector low energy signal region (below 6 keV) originates from coincident PMT noise. That is two PMT noise events occurring sufficiently

close together in each of a crystal's PMTs to mimic a signal. This noise can be caused by many phenomena including radioactive decay in the PMT material, discharging of accumulated space charge, dark current, and afterpulses of large events [73]. These noise events become more numerous at lower energies and can be difficult to separate from signal events with simple event selection.

Using the dedicated calibration holes that allow relatively easy access inside the detector housing, multiple calibration runs have taken place over the detector's lifetime using different calibration sources. Before the detector became operational, calibration runs to establish light yields, and energy scales and resolutions using ^{57}Co , ^{60}Co , ^{241}Am , and ^{137}Cs sources took place. Then latterly, after the detector became operational, four further calibration runs using ^{60}Co and ^{22}Na were carried out to provide pure samples of signal data to aid in PMT noise rejection.

The first two of these calibration runs were carried out in the early stages of the detector operation used ^{60}Co and allowed for the development of initial parameters based on an event's waveform to distinguish between signal and PMT noise events. In the case of ^{60}Co , the Co nuclei decay via β^- decay into ^{60}Ni in an excited state which then produces 1173 keV and 1332 keV photons upon de-excitation. Measuring scintillation events associated with the β^- decay through coincidence with an 1173 keV or 1332 keV photon produces a sample of pure signal-like data. The initial parameters developed based on studies of this sample are as follows:

- The integrated charge in the first 50 ns of an event (X2 or "fast-charge") and the integrated charge between 100 and 600 ns (X1 or "slow-charge"), both normalised to the total integrated charge between 0 and 600 ns. These parameters are based on parameters used in the DAMA experiments to separate signal and PMT noise events [73]. At COSINE-100, as can be seen in Fig. 9.3, X1 and X2 cuts give good separation between signal and noise, can remove around 80% of PMT noise events above 2 keV with around 99% signal efficiency with signal events having a high X1 low X2 distribution, and the opposite for PMT noise.
- An asymmetry cut based on the fraction of the total charge of an event measured in each of the two crystal PMTs, with more asymmetric events more likely to be from PMT noise.
- A cut based on the average number of single photo-electrons (SPEs) per cluster with a signal event having roughly two or fewer SPEs per cluster. This is to remove noise events with merged pulse clusters.

Through simple linear cuts using these variables, signal and noise events can begin to be separated, but with poor noise rejection and signal efficiency up to and approaching 2 keV.

Noise rejection and signal efficiency at a 2 keV threshold were improved through the use of a boosted decision tree (BDT) multivariate analysis algorithm with the above variables used as input parameters, in addition to event energy and the charge-weighted mean time. This BDT variable was trained on a sample of noise data from a regular physics run, and a sample of signal data from an approximately three week long ^{60}Co calibration run at the end of 2016/start of 2017. The signal data from the calibration run was weighted such that its energy distribution matched that of expected signal data from a physics run.

This BDT variable allowed for the low energy threshold to be set at 2 keV for the early dark matter searches with COSINE-100 [2, 75]. The signal efficiency for each crystal between 2-20 keV as calculated during the first annual modulation search at COSINE-100 (1.7 years data taking) using the above BDT is shown in Fig. 9.4 where selection efficiency above 60% above a 2 keV threshold can be seen for all five NaI crystals used in the analysis.

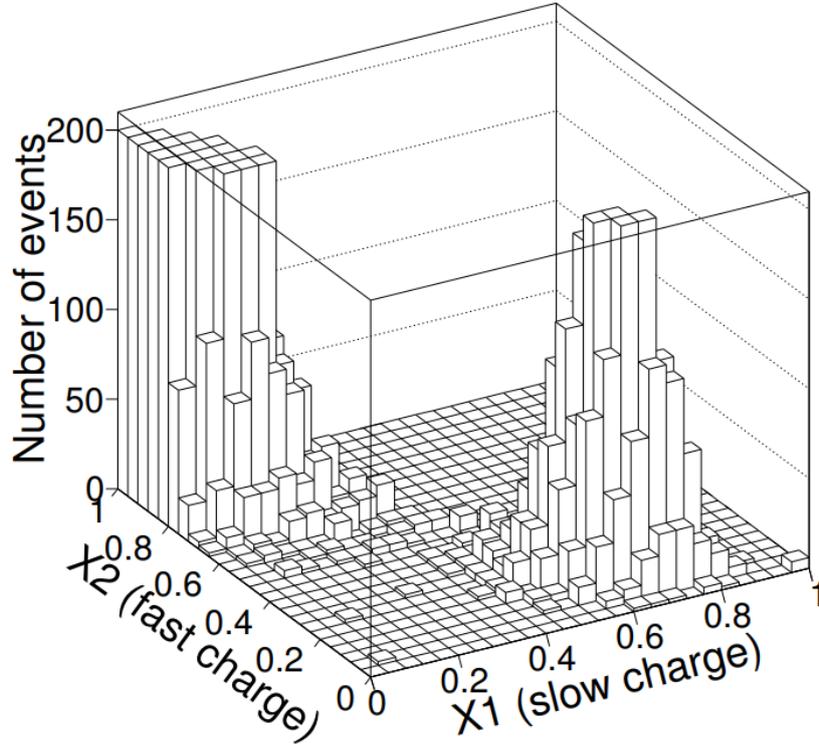


Figure 9.3: Distribution of the X1 (slow charge), and X2 (fast charge) for events above 2 keV in one of the crystals at COSINE-100. Two distinct signal (high-X1, low-X2) and noise (low-X1, high X-2) regions can be seen. From [73].

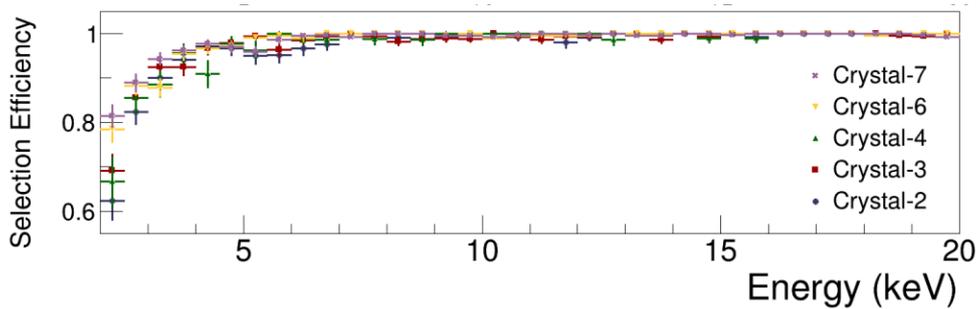


Figure 9.4: Signal efficiency between 2-20 keV for the five crystals used in modulation searches at COSINE-100 when an early BDT variable trained down to 2 keV on ^{60}Co calibration data is applied. From [75].

A requirement for a full test of the DAMA signal by COSINE-100 requires a low energy threshold that is at least as competitive as that of the DAMA/LIBRA-phase 2 experiment which, is able to search for annual modulation down to 1 keV [49]. This

necessitates developing more sophisticated techniques than those which allowed for a 2 keV low energy threshold.

The reason for the added complexity below 2 keV is that above this point, the PMT noise is dominated by signals with significantly shorter decay times (around 50 ns) than signal events which typically have decay times of around 250 ns. This is what allows the fast charge and slow charge variables to be effective at discriminating between signal and noise above 2 keV. This type of noise is known as “Type-I noise” PMT noise.

However, below 2 keV a separate population of noise events (known as “Type-II noise”) with mean-times that overlap with that of the signal region are observed, thus requiring an updated approach to noise rejection. Different types of PMT noise are described in [103]. Waveform templates characterising the shapes of the signal and the two types of noise are shown in Fig. 9.5. These templates were subsequently used in the development of an updated BDT variable, able to separate signal from Type-II noise PMT noise events.

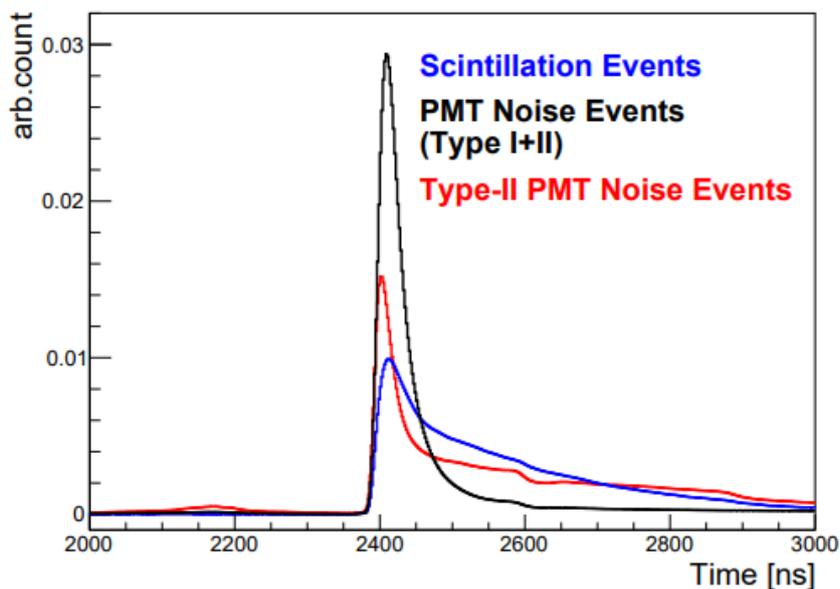


Figure 9.5: Noise templates from normalised probability density functions from accumulated events for scintillation events (blue), Type-II noise PMT noise events (red), and a combination of Type-I noise and Type-II noise (black). From [108].

The first step in this process was to better separate signal events from Type-I noise. This was achieved by taking the amplitude weighted mean-time, $\langle t \rangle$, individually for the two PMT pulses of an event using

$$\langle t \rangle = \frac{\sum_i t_i a_i}{\sum_i a_i} - T_1, \quad (9.1)$$

where t_i and a_i are the time and amplitude of the i^{th} bin of a pulse, and T_1 is the time of the first photo-electron of the pulse. The amplitude-weighted mean times of each PMT in a crystal that records a event are then combined as

$$p_m \equiv \ln(\langle t \rangle_1 + \langle t \rangle_2), \quad (9.2)$$

where in this case the 1 and 2 indices on $\langle t \rangle$ represent the two PMTs of a given crystal. p_m is known as the “mean-time parameter”. The separation power of this variable can be seen at different energies for the first 59.5 days of COSINE-100 data in Fig. 9.6. As expected

in line with the results from the initial 2 keV threshold procedure, good separation with the mean-time parameter is observed down to around 2 keV, at which point the signal and noise regions become merged.

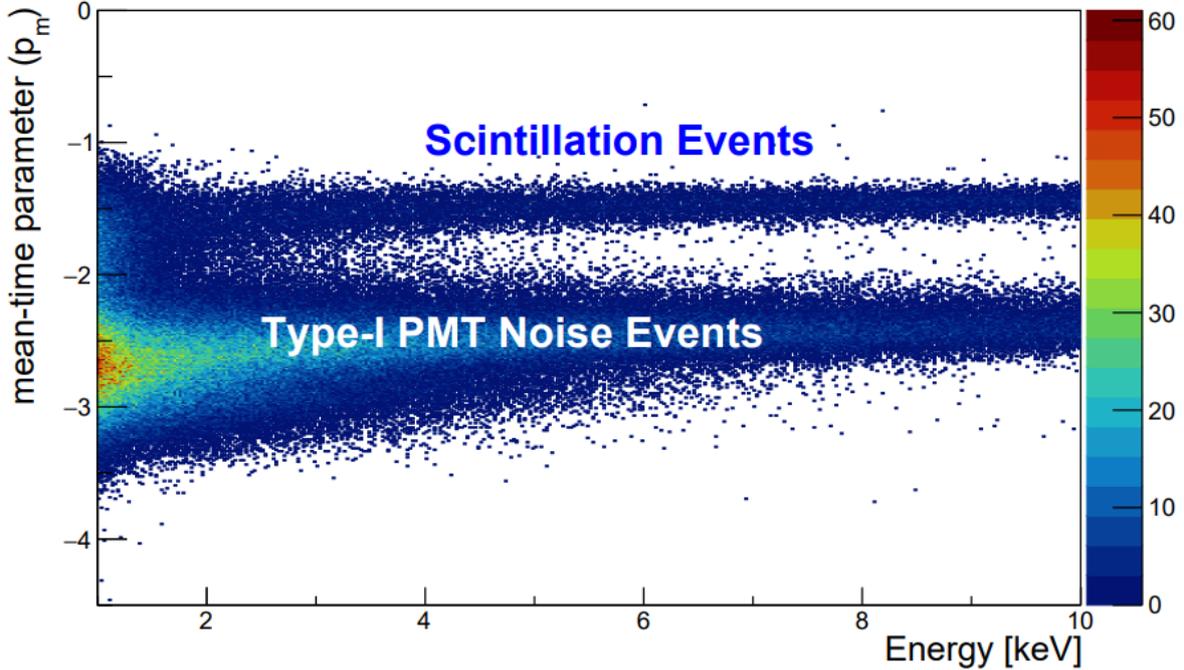


Figure 9.6: Mean-time parameter distribution of the first 59.5 days of COSINE-100 data taking shown as a function of energy. Above 2 keV there is good separation between scintillation (top band) and noise (bottom band) events. The distributions merge however below 2 keV. From [108].

To achieve separation below 2 keV, Type-II noise also must be separated from the signal events. To achieve this a new parameter characterising PMT-pulse shape is defined as

$$\lambda_{d,\text{crystal}} = \frac{\ln(Q_{\text{tail},i}/Q_{\text{head},i})}{T_{\text{tail},i} - T_{\text{head},i}}, \quad (9.3)$$

where $Q_{\text{head}} \equiv \sum_{t_i < t_c} q_i$ and $Q_{\text{tail}} \equiv \sum_{t_i > t_c} q_i$ are the amplitude sums of the first and second halves in time of the PMT's spectrum, where the halfway point, t_c , is set based on the trigger event and last SPE pulse times: t_i refers to the time of a given pulse in an event. $T_{\text{head}} \equiv (\sum_{t_i < t_c} q_i \cdot t_i)/Q_{\text{tail}}$ and $T_{\text{tail}} \equiv (\sum_{t_i > t_c} q_i \cdot t_i)/Q_{\text{tail}}$ are the amplitude-weighted mean times for the first and second halves of an event.

$$p_d \equiv \ln(\lambda_{d,1} + \lambda_{d,2}), \quad (9.4)$$

where as in the amplitude weighted mean-time calculation, the two individual PMT values are summed inside a natural logarithm.

It is observed in plots of the distribution of the mean-time and pulse-shape parameter at different energy ranges, shown in Fig. 9.7, that the distribution of Type-I noise does not change with energy, whereas Type-II noise appears almost exclusively below 2 keV. Using cuts based on the mean-time and pulse-shape parameters, and energy, a sample of Type-II noise events can be collected, thus allowing for the waveform template shown

in Fig. 9.5 to be generated. The waveform templates of the two types of noise are not sensitive to changes in mean-time and pulse-shape parameter cuts.

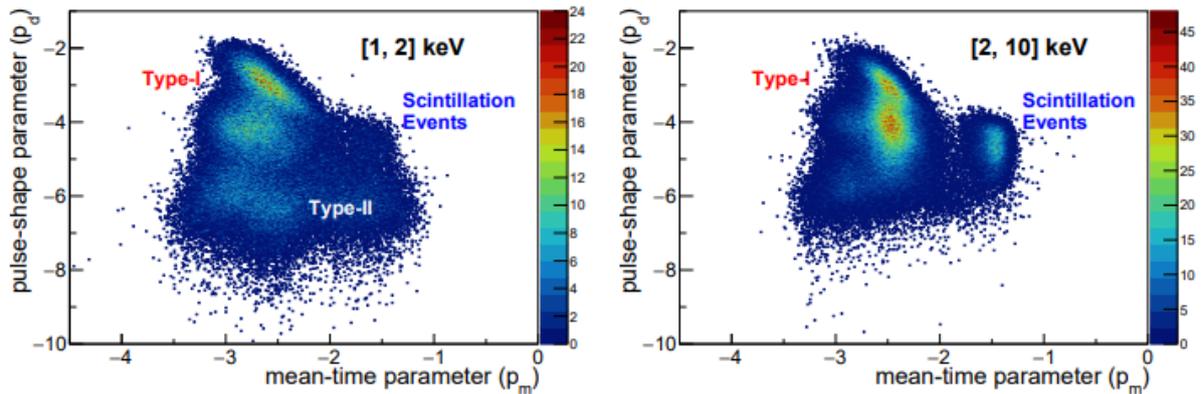


Figure 9.7: 2D-distribution of crystal pulse-shape and mean-time parameter for 1-2 keV (left) and 2-10 keV (right) events. Type-I noise can consistently be seen in the top left region of both plots, whereas the Type-II noise appears only below 2 keV. As can be seen, the Type-II noise has poor separation from signal events using the mean-time parameter. From [108].

While the pulse-shape parameter is an improvement on previous measures to identify and remove Type-II noise PMT noise, it still has limitations. Firstly, a corollary of splitting each event in half into “head” and “tail” regions is a requirement of at least two SPEs in each PMT, potentially limiting the effectiveness of a pulse-shape parameter-based cut at low energies. Additionally, the pulse-shape and mean-time parameters are inherently limited in the amount of information they can convey about a given waveform. As such, a new parameter based on the similarity of an event to a waveform template for signal and noise is developed.

To generate the data from which the signal waveform template was calculated, an additional ^{60}Co calibration run was carried out, on this occasion for 27.9 days, which provided a sample of pure data. Signal data was again selected from low energy event in coincidence with a 1173 keV or 1332 keV photon in another crystal. A sample of noise data was collected from a normal physics run. 5000 events are used to make each waveform template. The cuts for the signal and noise template are shown in Fig. 9.8 where events in the unshaded regions are selected as signal events.

Using these templates, two new parameters are defined for each event, one for each template. These are based on

$$\ln(\mathcal{L}) = \sum_i \left[T_i - W_i + W_i \ln \frac{W_i}{T_i} \right], \quad (9.5)$$

where \mathcal{L} is the likelihood parameter for a given waveform template (signal or noise), and T and W are the summed heights in the i^{th} time bin of the signal and template waveforms respectively.

The signal and noise likelihoods are combined as follows:

$$p_l = \frac{\ln \mathcal{L}_n - \ln \mathcal{L}_s}{\ln \mathcal{L}_n + \ln \mathcal{L}_s}, \quad (9.6)$$

where the n and s labels denote signal and noise respectively, and p_1 is known as the “likelihood parameter”. Larger values of p_1 indicate an event is closer to the signal template than the noise.

The distribution of p_1 is shown in Fig. 9.9 where a clear separation of signal and noise events is observed. The red line indicates the cut used to select signal events for the BDT training where events above the line are used.

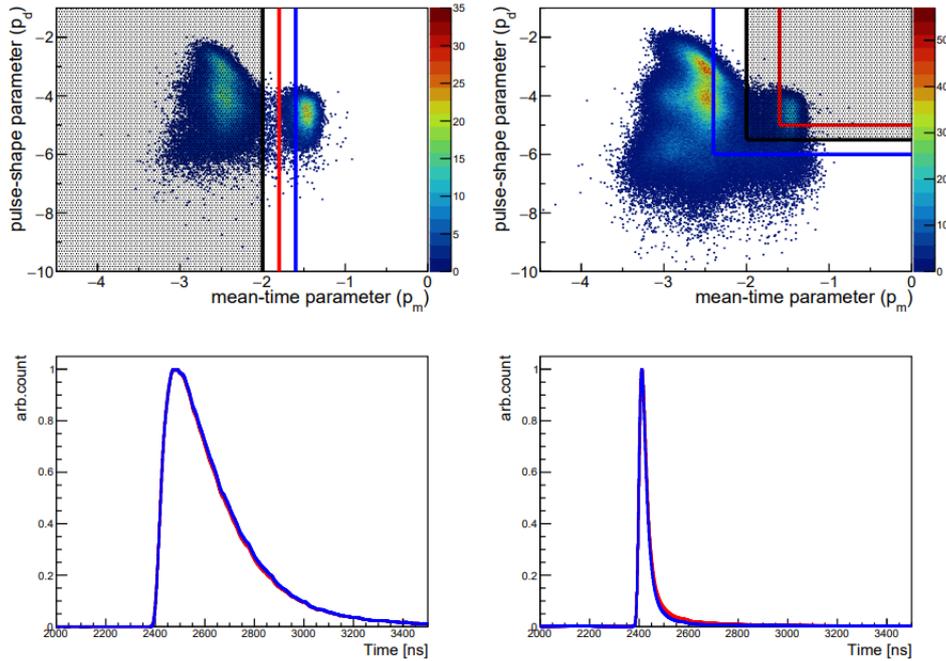


Figure 9.8: Pulse-shape and mean-time parameters distributions (top plots) for ^{60}Co calibration data (left) and physics data (right). The un-shaded regions represent regions containing events used for the selection of events for the signal (left) and noise (right) templates (bottom plots). The blue, black, and red lines show different choices of cuts, where the resulting waveform template for each cut is shown in the same colour. As can be seen, the choice of cut does not have a significant effect on the waveform template. From [108].

From Fig. 9.9 it is clear that the likelihood parameter is the most powerful parameter at distinguishing signal and noise below 2 keV of any parameter so far discussed. However, noise rejection below 2 keV can be improved further by using the parameter to select a sample of signal data down to 1 keV, then training a new BDT variable with this data, using p_1 as one of the training variables.

Signal events were chosen from multiple-hit events in the ^{60}Co calibration run that passed the energy dependent cut on p_1 as shown by the red line in Fig. 9.9. Since the ^{60}Co calibration run has different physical characteristics to physics runs and therefore a different energy spectrum, the calibration data is re-weighted to match the distribution in energy of physics data. Noise events were selected from a random 50% sample of the first 59.5 days of data, a period dominated by PMT noise.

This 1 keV-level BDT variable uses the same variables as the previous BDT trained down to 2 keV, with the additions of the mean-time and likelihood parameters discussed above. After the BDT training has been completed, a single variable for noise rejection below 2 keV is produced. This is then used to test the purity of the training sample. By

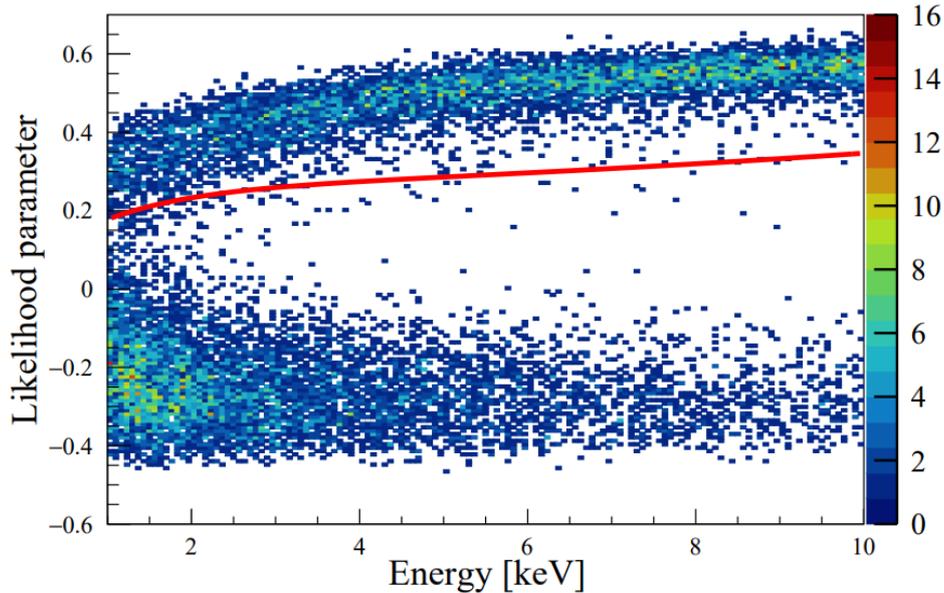


Figure 9.9: Likelihood parameter based on the similarity of a given event’s waveform to signal and noise template shown for multiple-hit events in ^{60}Co calibration data. The red line represents the cut used to select signal events for BDT training with signal events falling above the line. From [108].

modelling the signal and noise regions of the 1-D BDT distribution of the signal training sample, the signal purity in the 1-1.5 keV region was estimated to be over 99% and the signal efficiency to be around 80% as shown in Fig. 9.10. In this figure, the selection efficiencies of β/γ events generated during a ^{60}Co calibration run both in one of the analysis crystals and in a test crystal, and from events from a mono-energetic 2.42 MeV neutron beam in an early calibration run are all shown and all have good agreement.

The distribution of this new BDT variable in energy is shown for approximately 250 days of SET4 physics data in Fig. 9.11. A clear separation of signal (higher BDT values) and noise (low BDT values) events can be seen. An energy dependent cut for each crystal is developed to achieve optimal signal purity and noise rejection across in all crystals. It is this energy-dependent BDT cut that was used in the event selection in the COSINE-100 SET3 annual modulation search [1]. Since then, the value of this cut has been slightly tweaked and improved for the SET4 search described in this work.

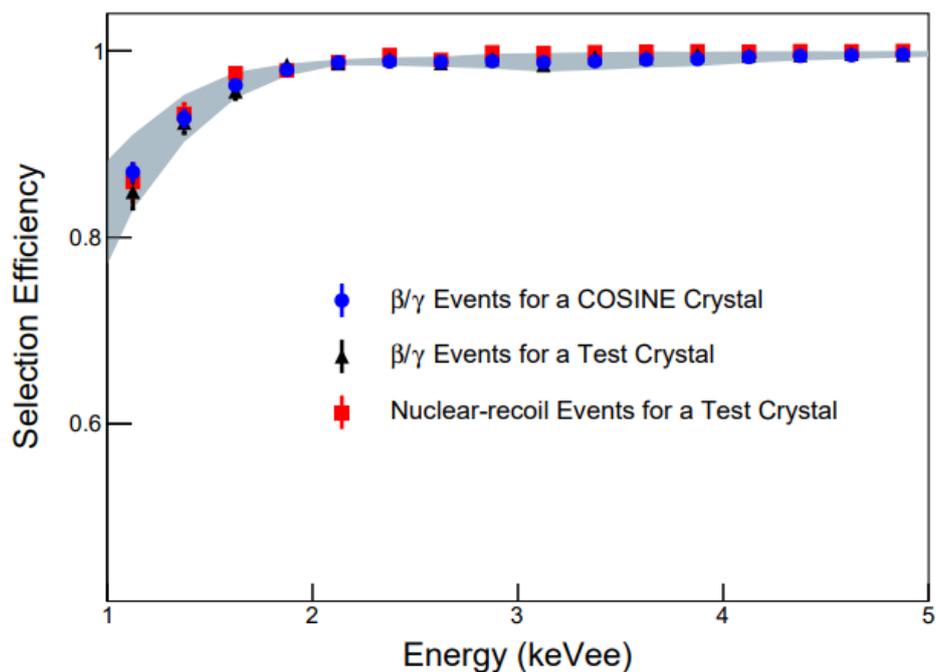


Figure 9.10: The selection efficiencies at different energies in 0.25 keV-wide bins when using the BDT variable described in the text for β/γ events from a ^{60}Co calibration source in both a test crystal (black triangles) and a crystal installed in the COSINE-100 detector (blue circles), as well as for nuclear recoils from events from a mono-energetic 2.42 MeV neutron beam (red squares). From [3].

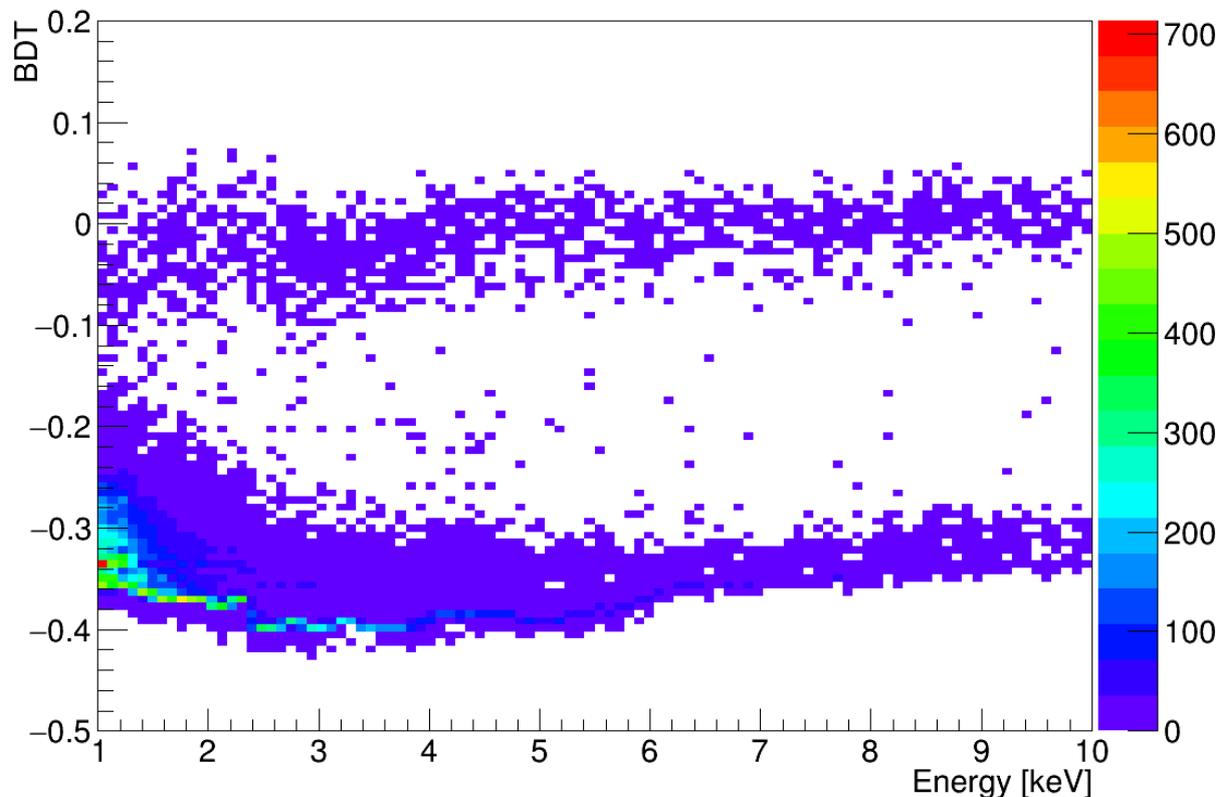


Figure 9.11: Distribution of the 1 keV BDT variable and energy for a sample of 250 days of SET4 physics data in a single crystal is shown. Simple cuts to remove obvious noise events as described in the text have been applied. There is a clear separation of signal (top) and noise (bottom) events down to 1 keV.

9.2.2 SET4 data selection and quality checks

Not all of the data recorded during the SET4 data-taking period may be used in the annual modulation searches at COSINE-100. As described in Section 4.3.2, all of the data recorded at COSINE-100 is split into two hour-long “subruns” and each subrun is manually checked by a member of the collaboration to ensure no issue has occurred during that period. In the case where there is some issue, most commonly due a muon event temporarily overloading the data acquisition system, and full confidence cannot be obtained of the quality of all the data in that subrun, the subrun is marked as “bad” and not used in analysis. In addition to this continuous monitoring, a second data quality check is performed on the SET4 data-set to remove high-noise subruns where it is likely that the increased noise rate has impacted on the signal event rate.

More specifically, there exists a third type of PMT noise at COSINE-100 known as “Gaussian noise” (described in more detail in Chapter 6) with a different pulse-shape to the other types of PMT noise already described, that appears intermittently throughout the data-taking period. An exception to this is in C1 where there is continuously a high rate of Gaussian noise. This is one of the reasons events in C1 are not used as physics events. The pulse-shape of Gaussian noise events is less well-separated from signal events than the other types of noise using the BDT variable described above so is more frequently not rejected by any of the cuts used to select signal data. During periods where large amounts of Gaussian noise are present in a crystal, the data should be discarded to avoid the additional noise contaminating the signal events. Since these extra Gaussian noise events appear only in one or a few of the crystals, it is clear that they are indeed noise events, rather than signal events which would be detected in all the crystals. The exact cause of this type of noise is unknown, but visual examinations the crystal encapsulation during the process of deconstructing the COSINE-100 apparatus to facilitate a move to a new site for the construction of an upgraded COSINE-200 experiment revealed that some of the LS liquid may have leaked into the encapsulation, causing the noise.

To identify regions of unexpectedly high event rate, the above event selection is applied to get the single-hit event rate in each subrun, then two periods with an event rate significantly above the expected rate were selected for further examination. The first of these cases is an approximately 4.5 month-long period in C7, beginning at the start of the data-taking period, and the other in a one month-long period between April and May 2018 in C2, large amounts of Gaussian noise were identified in these periods. Consequently the data from these crystals in these periods is rejected and not used in the annual modulation searches. These periods can be seen in any of the plots showing the single-hit event rate in this thesis.

Once these periods have been removed, the distributions of the number of counts per subrun in these regions are plotted for each crystal. These distributions are well described by a Poissonion function, shown for C6 and C2 in Fig. 9.12, allowing for subruns with counts above a 99.9% confidence level defined by a Poissonion fit to the distribution to be selected for further investigation. As shown in Fig. 9.12, in some crystals (C3, C4, and C6), this is only a handful of subruns, whereas in the other crystals (C2 and C7), a larger number of subruns are potential high-noise candidates. The energy spectra at low energies are then examined for these high-rate subruns and compared to the normalised spectrum from a sample of good subruns. When the spectrum of the subrun being investigated differs from the reference spectrum, the subrun is marked as bad.

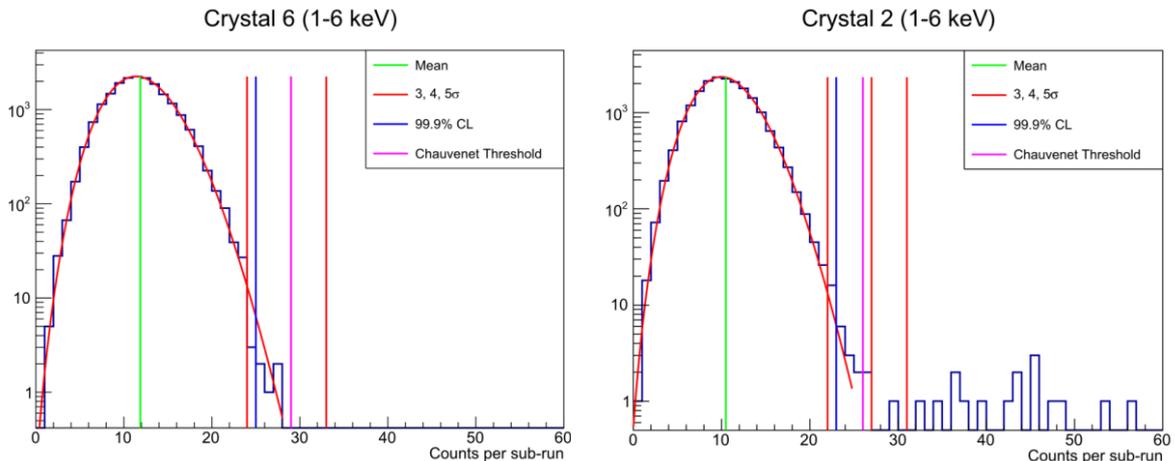


Figure 9.12: Distributions of the counts per subrun over the SET4 data-taking period for two of the COSINE-100 NaI crystals, C6 (left), and C2 (right) fit to a Poissonion (red curve). Subruns with counts above the 99.9% confidence level (blue line) defined by this Poissonion are selected for further investigation to see if they contain problematic amounts of Gaussian noise. From an internal presentation by SM Lee.

As a result of this process, 47 subruns (94 hours) in C2 were marked as bad in addition to the prolonged period removed in 2018. No other subruns were removed as a result of the data quality check, with the exception of the extended period at the beginning of data-taking in C7 as already discussed. It should be noted that while the data quality check process was repeated for the SET4 analysis across the entire data-set, the resultant list of good subruns prior to the end of the SET3 analysis (which concluded on the 21st of November 2019) is the same as in the SET3 analysis [1]. After this process was carried out, the remaining subruns in the data-set have the BDT/BDTA Gaussian noise event selection described in Section 6.2.1 applied to ensure all regions of high Gaussian noise had been identified. No further regions of Gaussian noise were identified from this.

Finally, 14 more hours of data in all crystals are marked bad immediately after the removal of one of the C8 PMTs as described in Section 9.2. The 12 hours prior to the removal of the PMT had already been marked as bad during the regular monitoring process, and while the single-hit rate appears to be restored to normal in the target crystals after the removal of the PMT, the trigger rate is still high in C8 for a further 14 hours. Therefore, since this is a very short period, the period was conservatively marked as bad and not used.

Overall, approximately 93% of the SET4 period is marked as good data, with dead-time as a result of the factors described above in addition to live-time lost due to calibration runs and power outages. The 61.3 kg of active NaI(Tl) target material in this time gives a total exposure of 261 kg \times years.

9.3 Fixed-phase annual modulation search results

The above event selection was applied to the data in the SET4 period and the resultant events were divided into a total of 115 15-day-long bins using the same binning as the pseudo-data described in Section 8.2. The beginning of the first bin, as with the previous annual modulation searches, was set at the 12th of October 2016 and the final bin ends on the 29th of June 2021. Unlike in previous annual modulation searches at COSINE-100,

the event rates are not scaled in line with 100% live-time in each bin, but rather the fit model used is instead scaled to match the live-time observed. This data-set was analysed with the method described in Section 7.2 in order to search for an annual modulation signal attributable to dark matter, and in doing so test the claim of a positive signal of the same type from the DAMA collaboration.

The annual modulation search is split into two separate parts, one where the phase of the annual modulation in the fit model used to describe the data is kept fixed at 152.5 days in line with the standard halo model. In the other the phase is allowed to float freely in the fitting process. The period used is fixed at one year in both cases. The search for dark matter and tests of the DAMA signal therein are described for each method in turn.

9.3.1 Calculation of the best-fit modulation amplitudes

As previously described in Chapters. 7 and 8, the fit model used to describe the COSINE-100 data comprises two components. A background component consisting of multiple exponential functions plus a flat function is used to characterise the event rate of background components in the detector that are not removed by the above event selection since they are genuine scintillation, single-hit events. A separate annual modulation component is also included to describe any annual modulation in the data. The background event rate is calculated independently for each crystal whereas the annual modulation component is fixed per unit mass across the detector. The primary regions of interest in energy in the context of testing the DAMA dark matter signal are 1-6 keV and 2-6 keV, in line with the published results from the DAMA experiments [49, 66, 68, 70].

The SET4 fit model includes an extra background component, designed to account for the observed effective extension of the half-life of the surface ^{210}Pb in the detector as a result of this contaminant being replenished over time by the decay products of ^{222}Rn . As described in Sec. 7.2.2, using this new fit model, the annual modulation amplitude of simulated “pseudo”-data were correctly identified, but not the activities of the background components of the pseudo-data. As such, fit models both with and without this updated surface ^{210}Pb component are used in the search for dark matter and their results compared.

The best-fit values are extracted from the means of marginalised posterior distributions of parameters of interest. These posterior distributions are calculated using Markov chain Monte Carlo techniques with 10^7 entries in the fixed-phase search and 10^8 entries in the floated-phase search. The higher number of entries is necessary in the floated-phase analysis in order to produce smooth highest-density credible regions (HDRs) of the combined two-dimensional posterior distribution of the modulation amplitude and phase components. In the fixed-phase case, 10^7 entries are sufficient to generate a one-dimensional Gaussian posterior distribution for the modulation amplitude from which the the 68.3%, 95.5%, 99.7% etc highest-density intervals (HDIs) are very well approximated by the $\pm 1\sigma$, $\pm 2\sigma$, $\pm 3\sigma$ etc. given by the standard deviation of the distribution.

Single-hit data-sets with energy ROIs of 1-6 keV and 2-6 keV are generated using the above event selection and analysed using both the split and single- ^{210}Pb modulation analysis search methods, initially both with the phase fixed at 152.5 days in line with the standard halo model. The 1-6 keV data-set with the best-fit functions from the split- ^{210}Pb fitting method is shown in Fig. 9.13 where good agreement between the data and best-fit function is observed. The best-fit function obtained from the equivalent single- ^{210}Pb analysis is visually very similar to that shown in Fig. 9.13.

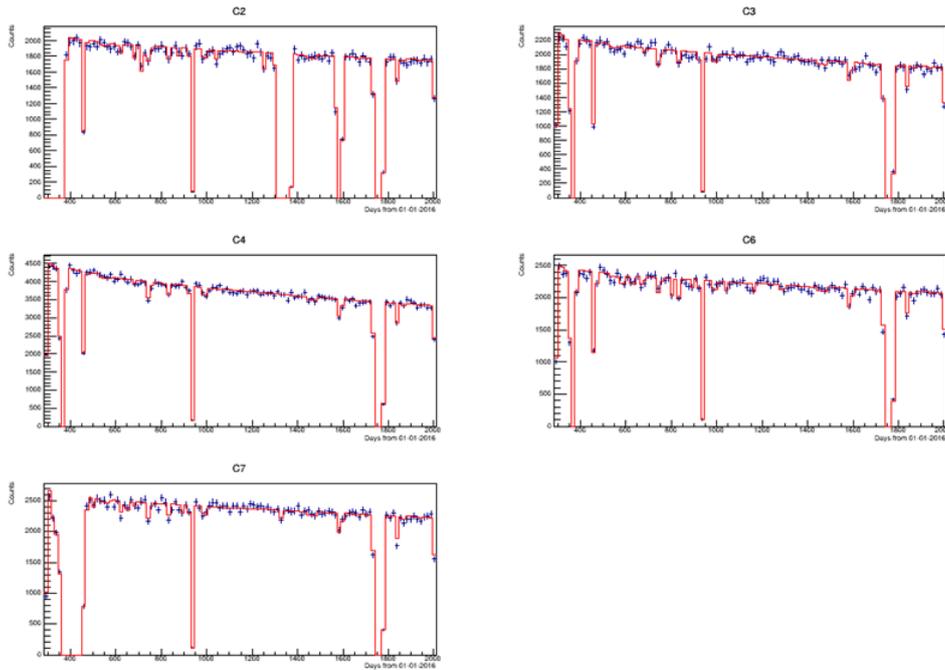


Figure 9.13: The rate against time of single-hit, signal-like events during COSINE-100 SET4 data-taking period in 15 day-long bins (blue data-points), fit (red line) to a fixed-phase annual modulation search model describing the background events in the data (including a split- ^{210}Pb component) as well as an annual modulation component. The phase of the annual modulation component is fixed at 152.5 days in line with the standard halo model.

The marginalised posterior distributions for the modulation amplitude parameter are drawn in all cases and their means calculated to give the best-fit values for the annual modulation amplitude in each case. The standard deviations of these distributions are given as the $\pm 1\sigma$ uncertainties. These best-fit modulation amplitudes and $\pm 1\sigma$ uncertainties are 0.0040 ± 0.0034 dru (0.0047 ± 0.0041 dru) and 0.0051 ± 0.0033 dru (0.0051 ± 0.0036 dru) for the 1-6 keV (2-6 keV) ROIs for the split and single- ^{210}Pb models respectively. These values are summarised along with the equivalent results from the DAMA and ANAIS experiments in Table. 9.1. The posterior distribution for the annual modulation amplitude in the 1-6 keV, split- ^{210}Pb case, overlaid with the distribution of the equivalent single- ^{210}Pb analysis is shown in Fig. 9.14. The location of the modulation amplitude in the 1-6 keV ROI reported by DAMA is also shown. The values in Table. 9.1 are presented graphically in Fig. 9.15.

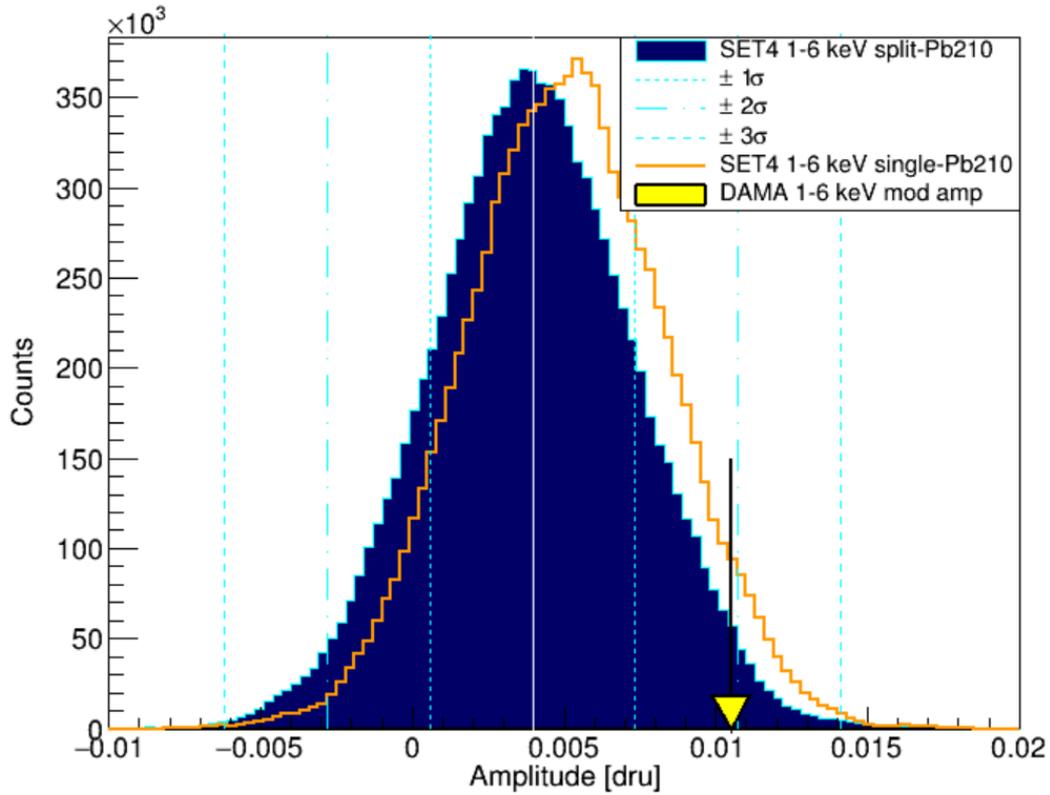


Figure 9.14: The marginalised posterior distributions of the annual modulation amplitude from the fixed-phase analysis of the 1-6 keV, single-hit SET4 data-set at COSINE-100. The filled dark blue distribution represents the posterior distribution generated from the split- ^{210}Pb analysis method and the unfilled orange distribution is from the single- ^{210}Pb analysis. The mean of the split- ^{210}Pb distribution is shown by the solid white line, whereas the various dashed light blue lines represent the $\pm 1\sigma$, $\pm 2\sigma$, and $\pm 3\sigma$ uncertainties on the best-fit value for the modulation amplitude in the split- ^{210}Pb case and are calculated from the standard deviation of the distribution. The yellow-headed arrow shows the location of the modulation amplitude reported in their 1-6 keV search.

The compatibility of the fit results with the given input data-set was evaluated using a chi-square (χ^2) goodness-of-fit test where χ^2 values were numerically calculated following the same method as used in calculating the χ^2 values in the pseudo-data, described in Section 8.3.3. The χ^2/NDF scores for the best-fit model in the 1-6 keV (2-6 keV) ROIs are equal to 555/519 (542/519) and 562/524 (549/524) for the split and single- ^{210}Pb models respectively. The resultant p-values are shown in Table. 9.1.

Configuration	Amplitude (dru)	Phase (days)	Goodness of fit
COSINE-100 SET4 1-6 keV (split-Pb)	0.0040 ± 0.0034	152.5 (fixed)	p = 0.13
COSINE-100 SET4 1-6 keV (single-Pb)	0.0051 ± 0.0033	152.5 (fixed)	p = 0.22
COSINE-100 SET4 2-6 keV (split-Pb)	0.0047 ± 0.0041	152.5 (fixed)	p = 0.24
COSINE-100 SET4 2-6 keV (single-Pb)	0.0051 ± 0.0036	152.5 (fixed)	p = 0.12
COSINE-100 SET3 [1] 1-6 keV	0.0067 ± 0.0042	152.5 (fixed)	p = 0.24
COSINE-100 SET3 2-6 keV	0.0051 ± 0.0047	152.5 (fixed)	p = 0.49
COSINE-100 SET2 [75] 2-6 keV	0.0083 ± 0.0068	152.5 (fixed)	p = 0.47
ANAIS three year result [71] 1-6 keV	-0.0034 ± 0.0042	152.5 (fixed)	p = 0.011
ANAIS three year result 2-6 keV	0.0003 ± 0.0037	152.5 (fixed)	p = 0.14
ANAIS 1.5 year result [113] 1-6 keV	-0.0015 ± 0.0063	152.5 (fixed)	p = 0.18
ANAIS 1.5 year result 2-6 keV	-0.0044 ± 0.0058	152.5 (fixed)	p = 0.67
DAMA/LIBRA-phase2 [49] 1-6 keV	0.0105 ± 0.0011	152.5 (fixed)	p = 0.51
DAMA all experiments [49] 2-6 keV	0.0102 ± 0.0008	152.5 (fixed)	p = 0.93

Table 9.1: The best-fit values, uncertainties, and p-values for the fixed-phase annual modulation amplitude found in various different NaI-based dark matter detection experiments.

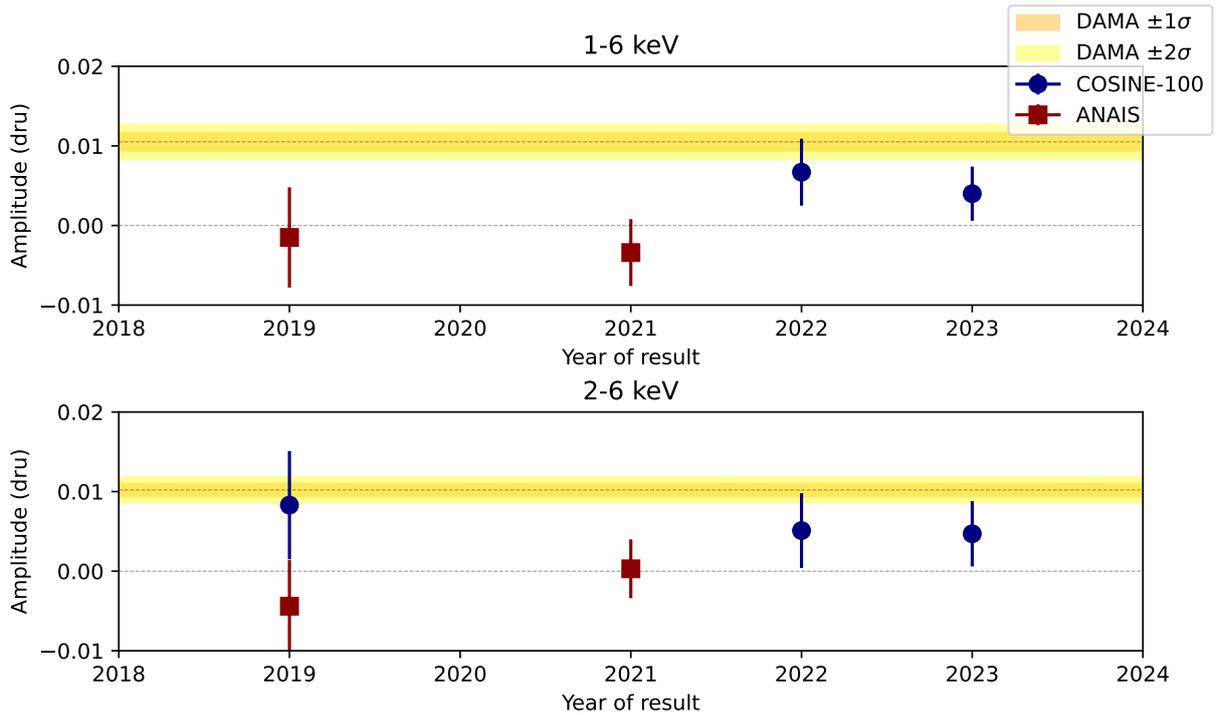


Figure 9.15: A visual representation of the fixed-phase best-fit annual modulation amplitudes presented in Table. 9.1. The blue points show results from COSINE-100 analyses ([1, 75] and this work) and the red points results from the ANAIS experiment ([71, 113]). The split- ^{210}Pb result is used in the most recent COSINE-100 points. The yellow shaded regions represent the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainties reported by DAMA [49].

Finally, an additional search for annual modulation across a wider energy ROI than the 1-6 keV ROI described was carried out. This required splitting the single-hit data-

set into individual 1 keV-wide bins from 1-2 keV to 19-20 keV. These data-sets were fit using the same single- ^{210}Pb method as described above. The single- ^{210}Pb method was chosen for this particular analysis since the activities of the split surface and bulk ^{210}Pb were not readily available in 1 keV-wide bins up to 20 keV from the time-dependent background model, as is the case for the other background components, including a single ^{210}Pb component. The best-fit modulation amplitudes and uncertainties for each energy region are shown in Fig. 9.16, along with the results from a similar analysis presented by the DAMA collaboration [49] using half a keV-wide regions.

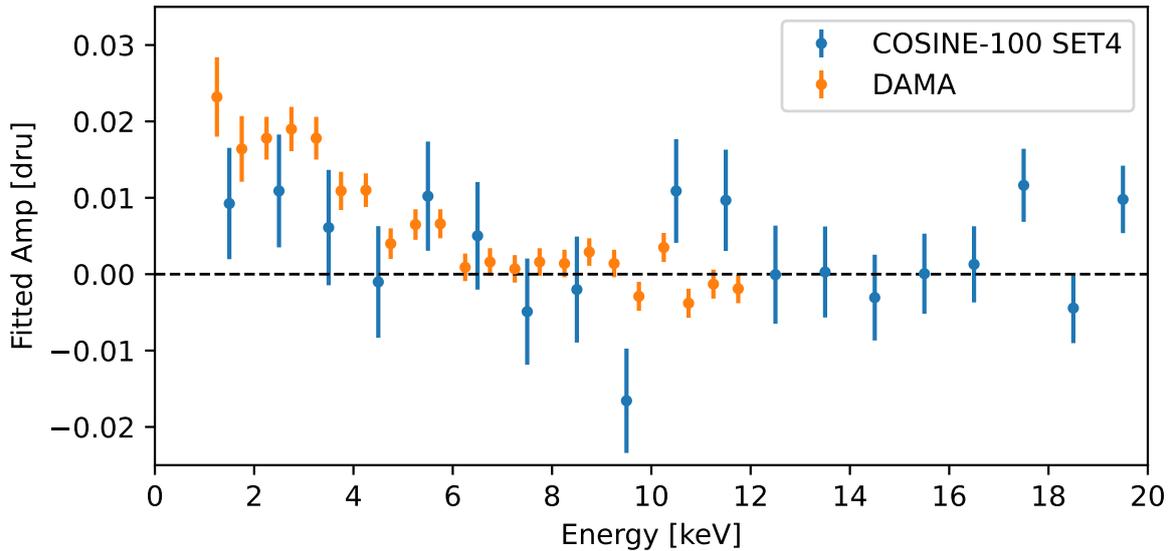


Figure 9.16: The best-fit modulation amplitudes and uncertainties in COSINE-100 SET4 single-hit data, split into 1 keV-wide energy regions (blue points). These values are compared to similar results from DAMA [49] in half-keV-wide ROIs (orange points).

9.3.2 Interpretations of the results

Once calculated, these best-fit annual modulation amplitudes can be compared both to the expected results found in the pseudo-data analysis described in Chapter 8 to judge the success and efficacy of the fitting procedure. Furthermore, the results can be compared to the DAMA results and no-modulation case as a means to potentially confirm or refute the DAMA claim.

The difference between the best-fit modulation amplitudes from the split and single- ^{210}Pb analysis methods in the 1-6 keV ROI is -0.0011 drU and therefore in agreement with the differences in best-fit amplitudes found in the pseudo-experiment-by-pseudo-experiment comparison shown in Fig. 8.7. The best-fit modulations in the 2-6 keV ROI equal for both methods. The uncertainties of the best-fit modulation amplitudes are similarly in agreement with the predicted value from the pseudo-data analysis, as are the χ^2 values. Importantly, none of the χ^2 values calculated fall into the range of “high χ^2 values” reported in Section 8.3.3.

In the 1-6 keV ROI, the best-fit modulation amplitude measured has moved away from the DAMA value and towards 0 with decreased errors relative to the SET3 result in both the single and split- ^{210}Pb cases. In the 2-6 ROI, the best-fit modulation amplitude found is equal to that reported in the SET3 analysis and as with the 1-6 keV ROI, smaller

uncertainties are reported as expected. However, as with past COSINE-100 analyses, it is not yet possible to conclusively judge in favour of either the DAMA or no-modulation hypothesis as the results found are in statistical agreement with both models. In the 1-6 keV case, the best-fit modulation amplitude found from the split- ^{210}Pb (single- ^{210}Pb) models differs from the DAMA value by 1.9σ (1.6σ) and from the no-modulation hypothesis by 1.2σ (1.5σ), where σ is the uncertainty associated with the COSINE-100 SET4 value. In the 2-6 keV ROI, these values are 1.3σ (1.4σ) for the DAMA comparison and 1.1σ (1.4σ) when comparing with no modulation. Similarly, in the analysis using the 1 keV-wide ROIs shown in Fig. 9.16, the best-fit modulation amplitudes found are generally with both the DAMA and null hypotheses at 6 keV and below, and in agreement with no modulation above 6 keV. No modulation above 6 keV is in line with the DAMA results and previous COSINE-100 results.

9.4 Floated-phase annual modulation search results

A search for an annual modulation with a phase not fixed at 152.5 days but instead a floated parameter in the fit model was also carried out on the same 1-6 keV and 2-6 keV single-hit data-sets already described. Extracting and presenting the best-fit amplitudes, phases, and related uncertainties of this floated-phase result is made more complex than in the fixed-phased case by the fact that either individual, one-dimensional, marginalised posterior distributions can be drawn for each parameter, or alternatively a single, two-dimensional, combined distribution can be drawn. The best-fit values and uncertainties extracted from these two types of posterior distribution will differ slightly.

The individual best-fit values are calculated from the means of the 1D marginalised distributions. These values and their one-dimensional uncertainties (from the standard deviation of the 1D distribution) are summarised in Table. 9.2 and are in statistical agreement with the equivalent results from the fixed-phase analysis. The floated-phase results reported by DAMA [49] are also shown, and as with the fixed-phase analysis, there is neither strong agreement nor strong tension between the COSINE-100 and DAMA results. The goodness of fit of these best-fit functions was again tested with a χ^2 goodness of fit test using the same method as in the fixed-phase analysis where 1-6 keV (2-6 keV) χ^2/NDF scores of 552/518 (542/518) and 551/523 (563/523) are found for the split and single- ^{210}Pb methods.

Configuration	Amplitude (dru)	Phase (days)
COSINE-100 SET4 1-6 keV (split-Pb)	0.0037 ± 0.0026	175 ± 73
COSINE-100 SET4 1-6 keV (single-Pb)	0.0044 ± 0.0029	157 ± 63
COSINE-100 SET4 2-6 keV (split-Pb)	0.0045 ± 0.0034	176 ± 75
COSINE-100 SET4 2-6 keV (single-Pb)	0.0056 ± 0.0035	128 ± 61
DAMA/LIBRA-phase2 [49] 1-6 keV	0.0106 ± 0.0011	148 ± 6
DAMA all experiments [49] 2-6 keV	0.0103 ± 0.0008	145 ± 5

Table 9.2: The best-fit values and uncertainties for the annual modulation amplitude and phase found in the COSINE-100 SET4 floated-phase annual modulation search in different configurations. These values are calculated from the means and standard deviations of the one-dimensional marginalised posterior distributions of each parameter. The equivalent values from the most recent DAMA results are also shown for comparison.

In addition to the extraction of best-fit values and uncertainties from the individual one-dimensional posterior distributions of the parameters of interest, combined two-dimensional distributions are also drawn and analysed, for example as shown for the 1-6 keV split- ^{210}Pb case in Fig. 9.17. Overlaid on this distribution are the means and associated 1σ uncertainties of the amplitude and phase from the 1D posterior distributions. The result shown is typical inasmuch as the best-fit value from the means of the 1D posterior distributions are in good agreement with the most probable bin in the 2D distributions, likewise the standard deviation of the 1D posterior distribution for the phase is in good agreement with the 68.3% HDR in the 2D distribution. Conversely, the equivalent 1D uncertainty for the amplitude is slightly less than the 68.3% HDR in the 2D distribution.

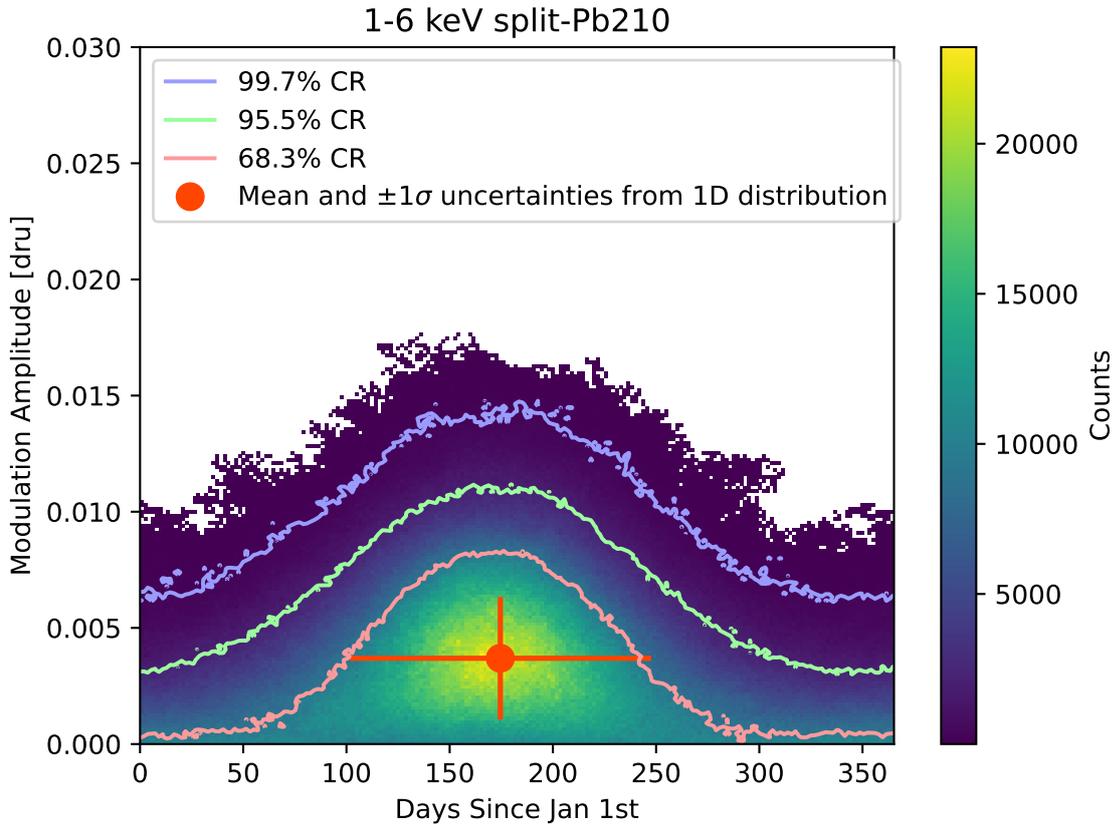


Figure 9.17: The two-dimensional marginalised posterior distribution of the annual modulation and phase in the floated-phase analysis of 1-6 keV single-hit COSINE-100 data using the split- ^{210}Pb fit model. Areas with higher counts correspond to areas of higher probability for the annual modulation amplitude and phase in the data-set as calculated in the fitting procedure. The pink, green, and blue curves represent the numerically calculated 68.3%, 95.5% and 99.7% highest-density credible regions (HDRs) of the distribution whereas the red point and error bars show the best-fit value and 1σ uncertainties calculated from the means and standard deviations of the one-dimensional histograms of the parameters.

As described in Section 8.5, comparing the results of the floated-phase analysis with the no-modulation hypothesis is non-trivial, due to the tendency of a non-zero best-fit modulation amplitude to be found, even in the case of a data-set where it is known there is no modulation. Pseudo-data with no injected modulation component was therefore used

to make a more meaningful comparison between the results generated in this analysis and the no-modulation case.

An ensemble of 1,000 such pseudo-experiments simulating data in the 1-6 keV ROI were fit using both the floated-phase single and split- ^{210}Pb analysis models and the best-fit modulation amplitudes and phases were extracted from the 1D posterior distributions for each experiment. These values were then used to populate a 2D histogram for each ensemble, representing the most likely best-fit amplitudes and phases for COSINE-100 data with no annual modulation. The same process was carried out for ensembles of pseudo-data with an injected modulation amplitude equal to that reported by DAMA [49]. The 68.3%, 95.5% and 99.7% HDRs of the equivalent analysis of the COSINE-100 single-hit physics data were overlaid on these distributions as shown in Fig. 9.18, to give a visual representation of the level of agreement between the observed data and the DAMA and no-modulation hypotheses. Both the split- ^{210}Pb (top of Fig. 9.18) and single- ^{210}Pb (bottom of Fig. 9.18) fitting models were tested. In future this level of agreement will be quantified by integrating the 2D histogram of pseudo-data results within the bounds of the HDRs from the physics data.

Visually inspecting the level of agreement between the 2D histogram of pseudo-data results and HDRs from physics results shows very good agreement between the physics data and no-modulation hypotheses in both fit methods, where the vast majority of the pseudo-data histogram is bound by the 68.3% HDR from the physics data results, and almost the entire pseudo-histogram is bound by the 99.7% HDR. In the DAMA modulation case, far less of the pseudo-histogram is bound by the 68.3% HDR and a larger fraction of the histogram lies outside of the 99.7% HDR. Nevertheless, the results observed from the physics data are in statistical agreement with both hypotheses using both methods, and in both ROIs.

9.5 Concluding remarks

In this chapter, the annual modulation search results from the first 4.7 years of COSINE-100 data representing a total exposure of 261 kg \times years are presented. This represents the largest sample of COSINE-100 data analysed to date and hence provides the most stringent test by COSINE-100 of the positive dark matter signal attributed to dark matter events by the DAMA collaboration. In both the fixed and floated-phase searches, a small positive modulation in statistical agreement with both the DAMA and no-modulation hypotheses was found. The modulation amplitudes reported for different search methods are equal or slightly smaller than those previously reported in COSINE-100 analyses and have smaller associated uncertainties. A larger exposure of COSINE-100 data is required for a conclusive test of the DAMA signal. This may come either from a future analysis of the 4.7-year data-set presented here with an updated event-selection procedure, allowing for the low energy threshold to be reduced from 1 keV to 0.75 keV, or from a larger exposure in time of COSINE-100 data.

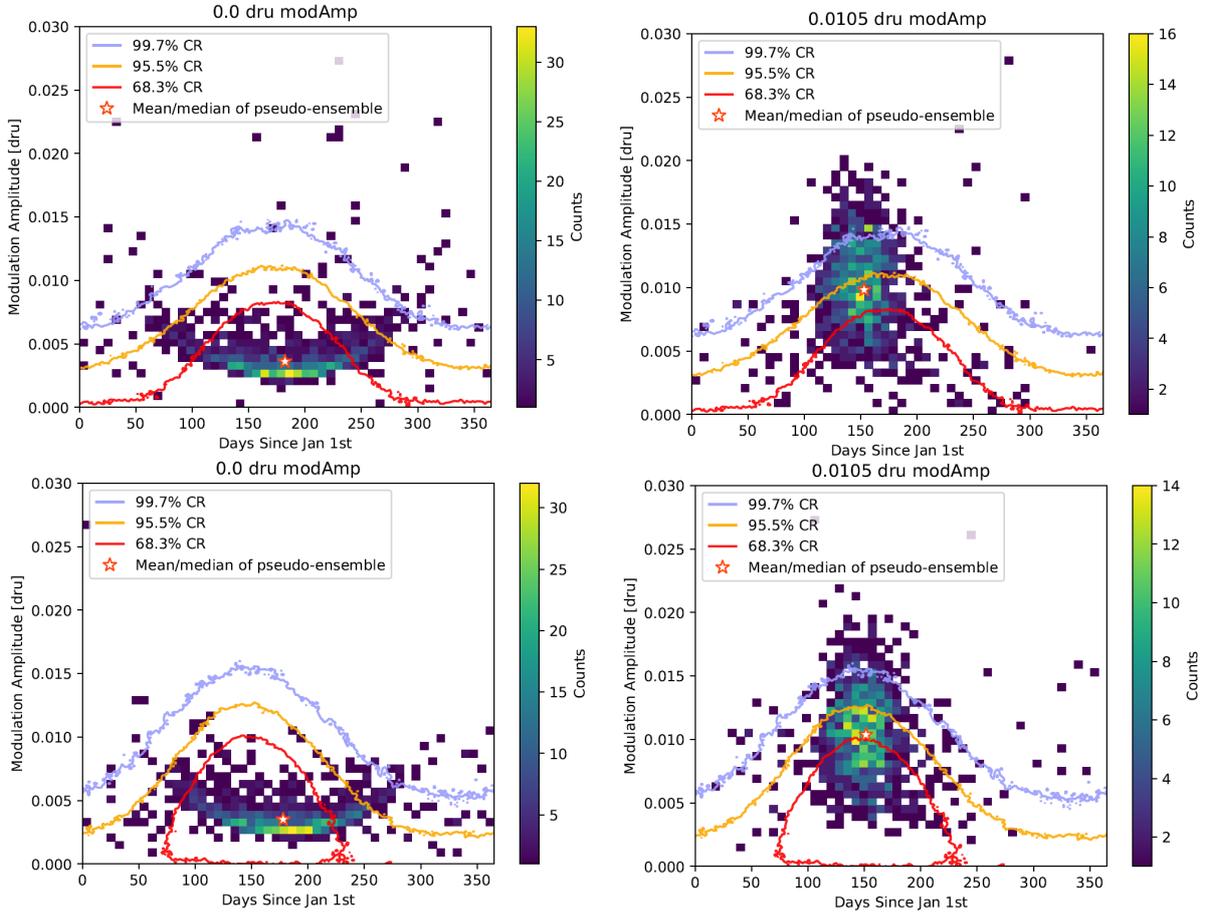


Figure 9.18: Two-dimensional histograms filled with the best-fit annual modulation amplitudes and phases found in floated-phase analyses of ensembles of 1,000 pseudo experiments simulating 1-6 keV COSINE-100 SET4 single-hit data with no modulation (left plots) an injected annual modulation with and phase and amplitude equal to the values reported by DAMA [49] (right). In the top plots the split- ^{210}Pb fit model was used and the single- ^{210}Pb model for the bottom plots. The mean of the best-fit phase and median of the best-fit amplitude of each pseudo ensemble is shown by the red and white star. Overlaid on these 2D histograms of pseudo-data results are the 68.3%, 95.5% and 99.7% HDRs (red, orange, and blue curves) found in the corresponding floated-phase search in 1-6 keV SET4 physics data. The proportion of the 2D histograms bound by the HDRs therefore represents the level of agreement between the observed physics data and the DAMA and no-modulation hypotheses.

Chapter 10

Conclusion and Outlook

In this thesis, searches for a dark matter-induced annual modulation signal in 4.7 years of COSINE-100 data are presented and examined through the lens of providing a model-independent test of the long-standing but controversial annual modulation signal, measured at the DAMA experiments and attributed to dark matter by the DAMA collaboration. This is the first time this data-set has been analysed. Consequently, the results presented in this work represent the most stringent test of the DAMA signal yet performed at the COSINE-100 experiment. Despite this, no definitive result was produced, where a small, positive annual modulation in statistical agreement with both the DAMA and no modulation hypothesis was found in all cases. In the fixed-phase modulation search using the split- ^{210}Pb method described in Chapters. 7 and 8, applied to data in the 1-6 keV ROI, the best-fit modulation amplitude reported is 0.0040 ± 0.0034 . The results of all the annual modulation searches are described in full in Chapter 9.

The search for a dark matter-induced annual modulation at COSINE-100 will continue with an increasing exposure in time due to extended detector operation, as well as with a lowered low energy threshold, thanks to the development of an improved BDT variable. Once development of this improved BDT variable and the consequent updates in event selection have been completed, the SET4 data-set will be expanded to lower energies and re-analysed to produce a final SET4 result from COSINE-100. This search will use an annual modulation search method based on the method described in this work, but with the issues in the characterisation of the background components, described in Chapter 8 and Appendix A addressed.

Data-taking at the first phase of the COSINE-100 experiment ceased on the 14th of March 2023, after nearly 6.4 years. An annual modulation search using this data-set will also be carried out, likely using an annual modulation search method based on the method described in this work. Work is also underway on the construction of an upgrade of the COSINE-100 experiment, COSINE-100U, experiment which will be re-located to the Yemilab facility in South Korea. COSINE-100U will use the same eight NaI crystals as in the original COSINE-100 experiment, but after re-machining and polishing. This will remove some of the crystal mass, but should allow for the issues in the instrumenting of crystals 1, 5, and 8 to be resolved and as such, the active detector mass will increase. Additionally, the detector will be operated at temperatures of around -35°C which in conjunction with a new method of directly attaching the crystals to their PMTs is expected to result in an improvement in light-yield of 45%. It is hoped COSINE-100U will be able to restart data-taking in early 2024. Beyond this, COSINE-200 and then COSINE-1T experiments are also planned with home-grown NaI crystals, the development of which is already underway. These three experiments are all intended to provide a stringent test

of the DAMA signal, while also setting World-leading limits for low-mass dark matter searches.

Bibliography

- [1] G. Adhikari, E. B. de Souza, N. Carlin, J. Choi, S. Choi, A. Ezeribe, L. França, C. Ha, I. Hahn, S. Hollick, *et al.*, “Three-year annual modulation search with COSINE-100,” *Physical Review D*, vol. 106, no. 5, p. 052005, 2022.
- [2] G. Adhikari, P. Adhikari, E. B. de Souza, N. Carlin, S. Choi, M. Djamel, A. C. Ezeribe, C. H. Ha, I. Hahn, A. J. Hubbard, *et al.*, “An experiment to search for dark matter interactions using sodium iodide detectors,” *Nature*, vol. 564, pp. 83–86, 2019.
- [3] G. Adhikari, E. B. de Souza, N. Carlin, J. J. Choi, S. Choi, M. Djamel, A. C. Ezeribe, L. E. França, C. H. Ha, I. S. Hahn, *et al.*, “Strong constraints from COSINE-100 on the DAMA dark matter results using the same sodium iodide target,” *Science Advances*, vol. 7, no. 46, p. eabk2699, 2021.
- [4] G. Adhikari, E. B. de Souza, N. Carlin, J. Choi, S. Choi, M. Djamel, A. Ezeribe, L. França, C. Ha, I. Hahn, *et al.*, “The COSINE-100 liquid scintillator veto system,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 1006, p. 165431, 2021.
- [5] G. Adhikari, N. Carlin, J. Choi, S. Choi, A. Ezeribe, L. E. França, C. H. Ha, I. S. Hahn, S. J. Hollick, E. Jeon, *et al.*, “An induced annual modulation signature in cosine-100 data by dama/libra’s analysis method,” *Scientific Reports*, vol. 13, no. 1, p. 4676, 2023.
- [6] N. Aghanim, Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. Banday, R. Barreiro, N. Bartolo, S. Basak, *et al.*, “Planck 2018 results-vi. cosmological parameters,” *Astronomy & Astrophysics*, vol. 641, p. A6, 2020.
- [7] Particle Data Group, R. Workman, V. Burkert, V. Crede, E. Klempt, U. Thoma, L. Tiator, K. Agashe, G. Aielli, B. Allanach, *et al.*, “Review of particle physics,” *Progress of theoretical and experimental physics*, vol. 2022, no. 8, p. 083C01, 2022.
- [8] V. Kuzmin and I. Tkachev, “Matter creation via vacuum fluctuations in the early universe and observed ultrahigh energy cosmic ray events,” *Physical Review D*, vol. 59, no. 12, p. 123006, 1999.
- [9] G. Bertone and D. Hooper, “History of dark matter,” *Reviews of Modern Physics*, vol. 90, no. 4, p. 045002, 2018.
- [10] V. C. Rubin and W. K. Ford Jr, “Rotation of the andromeda nebula from a spectroscopic survey of emission regions,” *The Astrophysical Journal*, vol. 159, p. 379, 1970.

- [11] G. B. Airy, “Account of some circumstances historically connected with the discovery of the planet exterior to uranus,” *Monthly Notices of the Royal Astronomical Society*, vol. 7, pp. 121–144, 1846.
- [12] J. C. Adams, “Explanation of the observed irregularities in the motion of uranus, on the hypothesis of disturbance by a more distant planet; with a determination of the mass, orbit, and position of the disturbing body,” *Monthly Notices of the Royal Astronomical Society*, vol. 7, pp. 149–152, 1846.
- [13] F. Zwicky, “The redshift of extragalactic nebulae,” *Helvetica Physica Acta*, vol. 6, pp. 110–127, 1933.
- [14] K. C. Freeman, “On the disks of spiral and S0 galaxies,” *Astrophysical Journal*, vol. 160, p. 811, 1970.
- [15] M. Milgrom, “A modification of the newtonian dynamics as a possible alternative to the hidden mass hypothesis,” *Astrophysical Journal, Part 1 (ISSN 0004-637X)*, vol. 270, pp. 365–370, 1983.
- [16] D. Clowe, M. Bradač, A. H. Gonzalez, M. Markevitch, S. W. Randall, C. Jones, and D. Zaritsky, “A direct empirical proof of the existence of dark matter,” *The Astrophysical Journal*, vol. 648, no. 2, p. L109, 2006.
- [17] G. W. Angus, B. Famaey, and H. Zhao, “Can MOND take a bullet? analytical comparisons of three versions of mond beyond spherical symmetry,” *Monthly Notices of the Royal Astronomical Society*, vol. 371, no. 1, pp. 138–146, 2006.
- [18] R. Massey, T. Kitching, and J. Richard, “The dark matter of gravitational lensing,” *Reports on Progress in Physics*, vol. 73, no. 8, p. 086901, 2010.
- [19] C. Bennett, D. Larson, J. Weiland, N. Jarosik, G. Hinshaw, N. Odegard, K. Smith, R. Hill, B. Gold, M. Halpern, *et al.*, “Nine-year wilkinson microwave anisotropy probe (WMAP) observations: final maps and results,” *The Astrophysical Journal Supplement Series*, vol. 208, no. 2, p. 20, 2013.
- [20] A. Coc, “Primordial nucleosynthesis,” in *Journal of Physics: Conference Series*, vol. 665, p. 012001, IOP Publishing, 2016.
- [21] T. Lin, “TASI lectures on dark matter models and direct detection,” *arXiv preprint arXiv:1904.07915*, 2019.
- [22] B. Carr, F. Kühnel, and M. Sandstad, “Primordial black holes as dark matter,” *Physical Review D*, vol. 94, no. 8, p. 083504, 2016.
- [23] S. W. Hawking, “Particle creation by black holes,” *Communications in mathematical physics*, vol. 43, no. 3, pp. 199–220, 1975.
- [24] B. Carr, K. Kohri, Y. Sendouda, and J. Yokoyama, “Constraints on primordial black holes from the galactic gamma-ray background,” *Physical Review D*, vol. 94, no. 4, p. 044029, 2016.
- [25] F. Capela, M. Pshirkov, and P. Tinyakov, “Constraints on primordial black holes as dark matter candidates from capture by neutron stars,” *Physical Review D*, vol. 87, no. 12, p. 123524, 2013.

- [26] H. Niikura, M. Takada, N. Yasuda, R. H. Lupton, T. Sumi, S. More, T. Kurita, S. Sugiyama, A. More, M. Oguri, *et al.*, “Microlensing constraints on primordial black holes with subaru/HSC andromeda observations,” *Nature Astronomy*, vol. 3, no. 6, pp. 524–534, 2019.
- [27] H. Karttunen, P. Kröger, H. Oja, M. Poutanen, and K. J. Donner, *Fundamental astronomy*. Springer, 2016.
- [28] P. Tisserand, L. Le Guillou, C. Afonso, J. Albert, J. Andersen, R. Ansari, É. Aubourg, P. Bareyre, J. Beaulieu, X. Charlot, *et al.*, “Limits on the macho content of the galactic halo from the EROS-2 survey of the magellanic clouds,” *Astronomy & Astrophysics*, vol. 469, no. 2, pp. 387–404, 2007.
- [29] “The nobel prize in physics 2015,” 2015.
- [30] L. Roszkowski, E. M. Sessolo, and S. Trojanowski, “WIMP dark matter candidates and searches—current status and future prospects,” *Reports on Progress in Physics*, vol. 81, no. 6, p. 066201, 2018.
- [31] A. Boyarsky, M. Drewes, T. Lasserre, S. Mertens, and O. Ruchayskiy, “Sterile neutrino dark matter,” *Progress in Particle and Nuclear Physics*, vol. 104, pp. 1–45, 2019.
- [32] M. Drewes, “The phenomenology of right handed neutrinos,” *International Journal of Modern Physics E*, vol. 22, no. 08, p. 1330019, 2013.
- [33] G. Steigman, B. Dasgupta, and J. F. Beacom, “Precise relic WIMP abundance and its impact on searches for dark matter annihilation,” *Physical Review D*, vol. 86, no. 2, p. 023506, 2012.
- [34] G. Gelmini and P. Gondolo, *Particle Dark Matter: Observations, Models and Searches*, ch. Chapter 7: DM production mechanisms, pp. 121–141. Cambridge University Press, 2010.
- [35] H. Baer, K.-Y. Choi, J. E. Kim, and L. Roszkowski, “Dark matter production in the early universe: beyond the thermal WIMP paradigm,” *Physics Reports*, vol. 555, pp. 1–60, 2015.
- [36] J. Wess and B. Zumino, “Supergauge transformations in four dimensions,” *Nuclear Physics B*, vol. 70, no. 1, pp. 39–50, 1974.
- [37] B. Zumino, “Supersymmetry then and now,” *Fortschritte der Physik*, vol. 54, no. 2-3, pp. 199–204, 2006.
- [38] S. Martin, “A supersymmetry primer,” *arXiv preprint hep-ph/9709356*, vol. 7, 2016.
- [39] S. Dimopoulos and H. Georgi, “Softly broken supersymmetry and SU (5),” *Nuclear Physics B*, vol. 193, no. 1, pp. 150–162, 1981.
- [40] R. D. Peccei and H. R. Quinn, “CP conservation in the presence of pseudoparticles,” *Physical Review Letters*, vol. 38, no. 25, p. 1440, 1977.
- [41] R. D. Peccei and H. R. Quinn, “CP conservation in the presence of pseudoparticles,” *Physical Review Letters*, vol. 16, no. 6, pp. 1791–1797, 1977.

- [42] R. D. Peccei, “The strong CP problem and axions,” in *Axions*, pp. 3–17, Springer, 2008.
- [43] S. Borsányi, Z. Fodor, J. Guenther, K.-H. Kampert, S. Katz, T. Kawanai, T. Kovacs, S. Mages, A. Pasztor, F. Pittler, *et al.*, “Calculation of the axion mass based on high-temperature lattice quantum chromodynamics,” *Nature*, vol. 539, no. 7627, pp. 69–71, 2016.
- [44] J. E. Kim, “Weak-interaction singlet and strong CP invariance,” *Physical Review Letters*, vol. 43, no. 2, p. 103, 1979.
- [45] M. A. Shifman, A. Vainshtein, and V. I. Zakharov, “Can confinement ensure natural CP invariance of strong interactions?,” *Nuclear Physics B*, vol. 166, no. 3, pp. 493–506, 1980.
- [46] M. Dine, W. Fischler, and M. Srednicki, “A simple solution to the strong CP problem with a harmless axion,” *Physics letters B*, vol. 104, no. 3, pp. 199–202, 1981.
- [47] A. Zhitnitsky, “On possible suppression of the axion hadron interactions.,” *Sov. J. Nucl. Phys*, vol. 31, p. 260, 1980.
- [48] C. Bartram, T. Braine, E. Burns, R. Cervantes, N. Crisosto, N. Du, H. Korandla, G. Leum, P. Mohapatra, T. Nitta, *et al.*, “Search for invisible axion dark matter in the 3.3–4.2 μ ev mass range,” *Physical review letters*, vol. 127, no. 26, p. 261803, 2021.
- [49] R. Bernabei, P. Belli, A. Bussolotti, F. Cappella, V. Caracciolo, R. Cerulli, C.-J. Dai, A. d’Angelo, A. Di Marco, H.-L. He, *et al.*, “First model independent results from DAMA/LIBRA–phase2,” *Universe*, vol. 4, no. 11, p. 116, 2018.
- [50] J. Billard, M. Boulay, S. Cebrián, L. Covi, G. Fiorillo, A. Green, J. Kopp, B. Majorovits, K. Palladino, F. Petricca, *et al.*, “Direct detection of dark matter—APPEC committee report,” *Reports on Progress in Physics*, vol. 85, no. 5, p. 056201, 2022.
- [51] R. Gaitskell, “Dark matter wimp searches,” in *IDM*, 2022.
- [52] The LZ collaboration, B. Mount, S. Hans, R. Rosero, M. Yeh, C. Chan, and R. Gaitskell, “LUX-ZEPLIN (LZ) technical design report,” tech. rep., Argonne National Lab.(ANL), Argonne, IL (United States); Pacific Northwest . . . , 2017.
- [53] J. Aalbers, D. Akerib, C. Akerlof, A. Al Musalhi, F. Alder, A. Alqahtani, S. Alsum, C. Amarasinghe, A. Ames, T. Anderson, *et al.*, “First dark matter search results from the LUX-ZEPLIN (LZ) experiment,” *arXiv preprint arXiv:2207.03764*, 2022.
- [54] E. Aprile, K. Abe, F. Agostini, S. A. Maouloud, L. Althueser, B. Andrieu, E. Angelino, J. Angevaare, V. Antochi, D. A. Martin, *et al.*, “First dark matter search with nuclear recoils from the XENONnT experiment,” *arXiv preprint arXiv:2303.14729*, 2023.
- [55] Y. Meng, Z. Wang, Y. Tao, A. Abdukerim, Z. Bo, W. Chen, X. Chen, Y. Chen, C. Cheng, Y. Cheng, *et al.*, “Dark matter search results from the PandaX-4T commissioning run,” *Physical Review Letters*, vol. 127, no. 26, p. 261802, 2021.

- [56] M. Ibe, W. Nakano, Y. Shoji, and K. Suzuki, “Migdal effect in dark matter direct detection experiments,” *Journal of High Energy Physics*, vol. 2018, no. 3, 2018.
- [57] R. Essig, J. Mardon, and T. Volansky, “Direct detection of sub-GeV dark matter,” *Physical Review D*, vol. 85, no. 7, p. 076007, 2012.
- [58] E. Aprile, J. Aalbers, F. Agostini, M. Alfonsi, L. Althueser, F. Amaro, V. C. Antochi, E. Angelino, J. Angevaare, F. Arneodo, *et al.*, “Excess electronic recoil events in XENON1T,” *Physical Review D*, vol. 102, no. 7, p. 072004, 2020.
- [59] E. Aprile, K. Abe, F. Agostini, S. A. Maouloud, L. Althueser, B. Andrieu, E. Angelino, J. Angevaare, V. C. Antochi, D. A. Martin, *et al.*, “Search for new physics in electronic recoil data from XENONnT,” *Physical Review Letters*, vol. 129, no. 16, p. 161805, 2022.
- [60] P. Agnes, I. Albuquerque, T. Alexander, A. Alton, M. Ave, H. Back, G. Batignani, K. Biery, V. Bocci, W. Bonivento, *et al.*, “Search for low-mass dark matter WIMPs with 12 ton-day exposure of DarkSide-50,” *Physical Review D*, vol. 107, no. 6, p. 063001, 2023.
- [61] D. Collaboration *et al.*, “Search for dark-matter-nucleon interactions via Migdal effect with DarkSide-50,” *Physical review letters*, vol. 130, no. 10, p. 101001, 2023.
- [62] E. Aprile, J. Aalbers, F. Agostini, M. Alfonsi, L. Althueser, F. Amaro, V. C. Antochi, E. Angelino, F. Arneodo, D. Barge, *et al.*, “Search for light dark matter interactions enhanced by the Migdal effect or Bremsstrahlung in XENON1T,” *Physical review letters*, vol. 123, no. 24, p. 241803, 2019.
- [63] G. Angloher, S. Banik, G. Benato, A. Bento, A. Bertolini, R. Breier, C. Bucci, J. Burkhardt, L. Canonica, A. D’Addabbo, *et al.*, “Results on sub-GeV dark matter from a 10 eV threshold CRESST-III silicon detector,” *Physical Review D*, vol. 107, no. 12, p. 122003, 2023.
- [64] B. Feldstein, A. L. Fitzpatrick, E. Katz, and B. Tweedie, “A simple explanation for DAMA with moderate channeling,” *Journal of Cosmology and Astroparticle Physics*, vol. 2010, no. 03, p. 029, 2010.
- [65] K. Freese, J. Frieman, and A. Gould, “Signal modulation in cold-dark-matter detection,” *Physical Review D*, vol. 37, no. 12, p. 3388, 1988.
- [66] R. Bernabei, P. Belli, F. Cappella, R. Cerulli, F. Montecchia, F. Nozzoli, A. Incicchitti, D. Prospero, C. Dai, H. Kuang, *et al.*, “Dark matter search,” *La Rivista del Nuovo Cimento*, vol. 26, pp. 1–73, 2003.
- [67] R. Bernabei, P. Belli, F. Cappella, R. Cerulli, C. Dai, A. d’Angelo, H. He, A. Incicchitti, H. Kuang, X. Ma, *et al.*, “New results from DAMA/LIBRA,” *The European Physical Journal C*, vol. 67, pp. 39–49, 2010.
- [68] R. Bernabei, P. Belli, F. Cappella, V. Caracciolo, R. Cerulli, C. Dai, A. d’Angelo, A. Di Marco, H. He, A. Incicchitti, *et al.*, “DAMA/LIBRA results and perspectives,” in *Journal of Physics: Conference Series*, vol. 384, p. 012005, IOP Publishing, 2012.

- [69] R. Bernabei, P. Belli, A. Bussolotti, F. Cappella, R. Cerulli, C. Dai, A. d'Angelo, H. He, A. Incicchitti, H. Kuang, *et al.*, “The dama/libra apparatus,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 592, no. 3, pp. 297–315, 2008.
- [70] R. Bernabei, P. Belli, F. Cappella, V. Caracciolo, R. Cerulli, C. Dai, A. d'Angelo, A. Incicchitti, A. Leoncini, X. Ma, *et al.*, “Dark matter: DAMA/LIBRA and its perspectives,” *SciPost Physics Proceedings*, no. 12, p. 025, 2023.
- [71] J. Amaré, S. Cebrián, D. Cintas, I. Coarasa, E. García, M. Martínez, M. Oliván, Y. Ortigoza, A. O. de Solórzano, J. Puimedón, *et al.*, “Annual modulation results from three-year exposure of ANAIS-112,” *Physical Review D*, vol. 103, no. 10, p. 102005, 2021.
- [72] M. Antonello, E. Barberio, T. Baroncelli, J. Benziger, L. Bignell, I. Bolognino, F. Calaprice, S. Copello, D. D'angelo, G. D'imperio, *et al.*, “The SABRE project and the SABRE proof-of-principle,” *The European Physical Journal C*, vol. 79, no. 4, pp. 1–8, 2019.
- [73] G. Adhikari, P. Adhikari, E. B. de Souza, N. Carlin, S. Choi, W. Choi, M. Djamal, A. Ezeribe, C. Ha, I. Hahn, *et al.*, “Initial performance of the COSINE-100 experiment,” *The European Physical Journal C*, vol. 78, no. 2, p. 107, 2018.
- [74] G. Adhikari, E. B. de Souza, N. Carlin, J. Choi, S. Choi, M. Djamal, A. Ezeribe, L. França, C. Ha, I. Hahn, *et al.*, “Background modeling for dark matter search with 1.7 years of COSINE-100 data,” *The European Physical Journal C*, vol. 81, no. 9, pp. 1–9, 2021.
- [75] G. Adhikari, P. Adhikari, E. B. de Souza, N. Carlin, S. Choi, M. Djamal, A. Ezeribe, C. Ha, I. Hahn, E. J. Jeon, *et al.*, “Search for a dark matter-induced annual modulation signal in NaI (Tl) with the COSINE-100 experiment,” *Physical review letters*, vol. 123, no. 3, p. 031302, 2019.
- [76] I. Coarasa, J. Apilluelo, J. Amaré, S. Cebrián, D. Cintas, E. García, M. Martínez, M. Oliván, Y. Ortigoza, A. O. de Solórzano, *et al.*, “Improving ANAIS-112 sensitivity to DAMA/LIBRA signal with machine learning techniques,” *Journal of Cosmology and Astroparticle Physics*, vol. 2022, no. 11, p. 048, 2022.
- [77] G. Angloher, P. Carniti, L. Cassina, L. Gironi, C. Gotti, A. Gütlein, D. Hauff, M. Maino, S. Nagorny, L. Pagnanini, *et al.*, “The COSINUS project: perspectives of a NaI scintillating calorimeter for dark matter search,” *The European Physical Journal C*, vol. 76, pp. 1–7, 2016.
- [78] E. B. De Souza, J. Cherwinka, A. Cole, A. Ezeribe, D. Grant, F. Halzen, K. Heeger, L. Hsu, A. Hubbard, J. Jo, *et al.*, “First search for a dark matter annual modulation signal with NaI (Tl) in the Southern Hemisphere by DM-Ice17,” *Physical Review D*, vol. 95, no. 3, p. 032006, 2017.
- [79] K. Fushimi, D. Chernyak, H. Ejiri, K. Hata, R. Hazama, T. Iida, H. Ikeda, K. Imagawa, K. Inoue, H. Ishiura, *et al.*, “PICOLON dark matter search project,” in *Journal of Physics: Conference Series*, vol. 2156, p. 012045, IOP Publishing, 2021.

- [80] J. H. Davis, “Fitting the annual modulation in DAMA with neutrons from muons and neutrinos,” *Physical review letters*, vol. 113, no. 8, p. 081302, 2014.
- [81] H. Araujo, V. Kudryavtsev, N. Spooner, and T. Sumner, “Muon-induced neutron production and detection with GEANT4 and FLUKA,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 545, no. 1-2, pp. 398–411, 2005.
- [82] J. Klinger and V. Kudryavtsev, “Muon-induced neutrons do not explain the DAMA data,” *Physical review letters*, vol. 114, no. 15, p. 151301, 2015.
- [83] E. Fernandez-Martinez and R. Mahbubani, “The Gran Sasso muon puzzle,” *Journal of Cosmology and Astroparticle Physics*, vol. 2012, no. 07, p. 029, 2012.
- [84] J. Pradler, B. Singh, and I. Yavin, “On an unverified nuclear decay and its role in the DAMA experiment,” *Physics Letters B*, vol. 720, no. 4-5, pp. 399–404, 2013.
- [85] R. Bernabei, P. Belli, F. Cappella, V. Caracciolo, R. Cerulli, *et al.*, “Comment on “on an unverified nuclear decay and its role in the DAMA experiment”,” *arXiv preprint arXiv:1210.6199*, 2012.
- [86] R. Bernabei, P. Belli, F. Cappella, V. Caracciolo, R. Cerulli, C. Dai, A. d’Angelo, A. Di Marco, H. He, A. Incicchitti, *et al.*, “A few final comments to arxiv: 1210.7548 [hep-ph],” *arXiv preprint arXiv:1211.6346*, 2012.
- [87] J. Pradler and I. Yavin, “Addendum to “on an unverified nuclear decay and its role in the DAMA experiment”,” *Physics Letters B*, vol. 723, no. 1-3, pp. 168–171, 2013.
- [88] D. Buttazzo, P. Panci, N. Rossi, and A. Strumia, “Annual modulations from secular variations: relaxing DAMA?,” *Journal of High Energy Physics*, vol. 2020, no. 4, pp. 1–21, 2020.
- [89] Y. Ko, G. Adhikari, P. Adhikari, E. B. de Souza, N. Carlin, J. Choi, S. Choi, M. Djamal, A. Ezeribe, C. Ha, *et al.*, “Comparison between DAMA/LIBRA and COSINE-100 in the light of quenching factors,” *Journal of Cosmology and Astroparticle Physics*, vol. 2019, no. 11, p. 008, 2019.
- [90] G. Tomar, S. Kang, S. Scopel, and J.-H. Yoon, “Is a wimp explanation of the DAMA modulation effect still viable?,” *arXiv preprint arXiv:1911.12601*, 2019.
- [91] R. Bernabei, P. Belli, F. Cappella, V. Caracciolo, R. Cerulli, C. Dai, A. d’Angelo, A. Di Marco, H. He, A. Incicchitti, *et al.*, “Improved model-dependent corollary analyses after the first six annual cycles of DAMA/LIBRA-phase2,” *Nuclear Physics and Atomic Energy*, vol. 20, pp. 317–348, 2019.
- [92] R. Bernabei, P. Belli, F. Cappella, V. Caracciolo, S. Castellano, R. Cerulli, C. Dai, A. d’Angelo, S. d’Angelo, A. Di Marco, *et al.*, “Final model independent result of DAMA/LIBRA–phase1,” *The European Physical Journal C*, vol. 73, pp. 1–11, 2013.
- [93] H. Lee, H. Bhang, J. Choi, S. Choi, I. Hahn, E. Jeon, H. Joo, W. G. Kang, B. Kim, G. Kim, *et al.*, “Search for low-mass dark matter with CsI (Tl) crystal detectors,” *Physical Review D*, vol. 90, no. 5, p. 052006, 2014.

- [94] M. H. Lee, “Amore: A search for neutrinoless double-beta decay of ^{100}Mo using low-temperature molybdenum-containing crystal detectors,” *Journal of Instrumentation*, vol. 15, no. 08, p. C08010, 2020.
- [95] C. Ha, Y. Jeong, W. G. Kang, J. Kim, K. Kim, S. Kim, Y. Kim, H. S. Lee, M. H. Lee, M. Lee, *et al.*, “Radon concentration variations at the Yangyang underground laboratory,” *Frontiers in Physics*, vol. 10, p. 1148, 2022.
- [96] H. Prihtiadi, G. Adhikari, P. Adhikari, E. B. De Souza, N. Carlin, S. Choi, W. Choi, M. Djamel, A. Ezeribe, C. Ha, *et al.*, “Muon detector for the COSINE-100 experiment,” *Journal of Instrumentation*, vol. 13, no. 02, p. T02007, 2018.
- [97] G. Adhikari, P. Adhikari, E. B. De Souza, N. Carlin, S. Choi, W. Choi, M. Djamel, A. Ezeribe, C. Ha, I. Hahn, *et al.*, “The COSINE-100 data acquisition system,” *Journal of Instrumentation*, vol. 13, no. 09, p. P09006, 2018.
- [98] H. Kim, G. Adhikari, E. B. de Souza, N. Carlin, J. Choi, S. Choi, M. Djamel, A. Ezeribe, L. França, C. Ha, *et al.*, “The environmental monitoring system at the cosine-100 experiment,” *Journal of Instrumentation*, vol. 17, no. 01, p. T01001, 2022.
- [99] K. Ianakiev, B. Alexandrov, P. Littlewood, and M. Browne, “Temperature behavior of NaI (Tl) scintillation detectors,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 607, no. 2, pp. 432–438, 2009.
- [100] G. Adhikari, P. Adhikari, C. Ha, E. J. Jeon, K. Kim, N. Y. Kim, Y. Kim, Y. J. Ko, H. S. Lee, J. Lee, *et al.*, “Study of fast neutron detector for COSINE-100 experiment,” *Journal of Instrumentation*, vol. 13, no. 06, p. T06005, 2018.
- [101] W. Thompson, the COSINE-100 Collaboration, *et al.*, “Current status and projected sensitivity of COSINE-100,” in *Journal of Physics: Conference Series*, vol. 1342, p. 012134, IOP Publishing, 2020.
- [102] H. Prihtiadi, G. Adhikari, E. B. De Souza, N. Carlin, J. Choi, S. Choi, M. Djamel, A. Ezeribe, L. França, C. Ha, *et al.*, “Measurement of the cosmic muon annual and diurnal flux variation with the COSINE-100 detector,” *Journal of Cosmology and Astroparticle Physics*, vol. 2021, no. 02, p. 013, 2021.
- [103] E. Bellamy, G. Bellettini, J. Budagov, F. Cervelli, I. Chirikov-Zorin, M. Incagli, D. Lucchesi, C. Pagliarone, S. Tokar, and F. Zetti, “Absolute calibration and monitoring of a spectrometric channel using a photomultiplier,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 339, pp. 468–476, 1994.
- [104] J. Chen, “Nuclear data sheets for $A=40$,” *Nuclear Data Sheets*, vol. 140, pp. 1–376, 2017.
- [105] G. Adhikari, E. B. de Souza, N. Carlin, J. Choi, S. Choi, M. Djamel, A. Ezeribe, L. Franca, C. Ha, I. Hahn, *et al.*, “Lowering the energy threshold in COSINE-100 dark matter searches,” *Astroparticle Physics*, vol. 130, p. 102581, 2021.

- [106] G. Adhikari, *Dark Matter Search with NaI(Tl) crystal and lowering the analysis threshold for COSINE-100 experiment*. PhD thesis, Sejong University, 2020.
- [107] W. Thompson, *Searching for Dark Matter with COSINE-100*. PhD thesis, Yale University, 2022.
- [108] G. Adhikari, E. B. de Souza, N. Carlin, J. Choi, S. Choi, M. Djamal, A. Ezeribe, L. França, C. Ha, I. Hahn, *et al.*, “Lowering the energy threshold in COSINE-100 dark matter searches,” *Astroparticle Physics*, vol. 130, 2021.
- [109] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, “Equation of state calculations by fast computing machines,” *The journal of chemical physics*, vol. 21, no. 6, pp. 1087–1092, 1953.
- [110] W. K. Hastings, “Monte Carlo methods using markov chains and their applications,” *Biometrika*, 1970.
- [111] G. Yu, C. Ha, E. J. Jeon, K. Kim, N. Y. Kim, Y. Kim, H. S. Lee, H. Park, and C. Rott, “Depth profile study of ^{210}Pb in the surface of an NaI (Tl) crystal,” *Astroparticle Physics*, vol. 126, p. 102518, 2021.
- [112] R. Brun and F. Rademakers, “ROOT—an object oriented data analysis framework,” *Nuclear instruments and methods in physics research section A: accelerators, spectrometers, detectors and associated equipment*, vol. 389, no. 1-2, pp. 81–86, 1997.
- [113] J. Amaré, S. Cebrián, I. Coarasa, C. Cuesta, E. García, M. Martínez, M. Oliván, Y. Ortigoza, A. O. de Solórzano, J. Puimedón, *et al.*, “First results on dark matter annual modulation from the ANAIS-112 experiment,” *Physical review letters*, vol. 123, no. 3, p. 031301, 2019.

Glossary

Λ CDM Λ Cold Dark Matter.

ADC Analogue to Digital Conversion.

ANAIS ANual modulation with NaI Scintillators, a NaI-based dark matter detection experiment.

BBN Big Bang Nucleosynthesis.

BDT Boosted Decision Tree.

BDTA A BDT variable designed with the goal of selecting .

CAT95 Charge Accumulation Time for 95% of the total charge of an event at COSINE-100, measured from the first pulse of the event.

CMB Cosmic Microwave Background.

cold dark matter Dark matter that is not relativistic at the time of freeze-out.

COSINE-100 The COSINE-100 dark matter detection experiment.

CP Charge Parity.

DAMA The DAMA dark matter detection experiments/collaboration.

DAMA/LIBRA The DAMA/LIBRA (Large sodium Iodide Bulk for RAre processes) experiment.

DAQ Data Acquisition system.

dru Daily Rate Units, counts/(days x kg x keV).

FADC Flash Analogue to Digital Converter.

freeze-out The time at which a particle falls of out thermal equilibrium in the early universe.

Gaussian noise An intermittent type of PMT noise at COSINE-100 with a waveform with a shape similar to a Gaussian.

HDI Highest-Density Interval.

HDR Highest-Density credible Region.

hot dark matter Dark matter that is relativistic at the time of freeze-out.

IBS Institute for Basic Science.

KIMS Korean Invisible Mass Search.

LAB Linear A kyl-Benzene.

LAr Liquid Argon.

LS Liquid Scintillator.

LSP Lightest Supersymmetric Particle.

LSS Large Scale Structure.

LXe Liquid Xenon.

M64ADC Charge sensitive Flash Analogue to Digital Converter.

MACHO Massive Compact Halo Object.

MC Monte Carlo.

MOND MODified Newtonian Dynamics.

MSSM Minimally Supersymmetric Standard Model.

nvt The variance of the charge-weighted mean-time of an event at COSINE-100.

PBH Primordial Black Hole.

PMT Photomultiplier Tube.

PSD Pulse-Shape Discrimination.

ROI Region Of Interest.

run Up to 1,000 COSINE-100 subruns.

SET2 The first 1.7 years of (live) data-taking at COSINE-100, from 20/10/2016 to 18/07/2018.

SET3 The first 2.8 years of (live) data-taking at COSINE-100, from 20/10/2016 to 21/11/2019.

SET4 The first 4.7 years of (live) data-taking at COSINE-100, from 20/10/2016 to 29/07/2021 (this work).

SPE Single PhotoElectron.

subrun 2 hours of COSINE-100 data, stored as a single ROOT file.

SUSY Supersymmetry.

TCB Trigger Control and clock Board.

TPC Time Projection Chamber.

type-I noise The most common type of PMT noise at COSINE-100 above 2 keV.

type-II noise The most common type of PMT noise at COSINE-100 below 2 keV.

warm dark matter Dark matter that was relativistic at the time of freeze-out, but subsequently slowed to non-relativistic speeds.

WIMP Weakly Interacting Massive Particle.

WIMP wind The changing flux of WIMPs as measured on Earth due to the Earth's relative motion around the galactic centre.

Y21 The Yangyang Underground Laboratory.

Appendix A

Further tests of describing background components using the split- ^{210}Pb fit method

A.1 Investigating the performance of the fit of the surface ^{210}Pb half-life

The new surface ^{210}Pb component is inherently different to the other background components in the SET4 fit model and also the background components in the SET3 fit model in that the half-life of this component is a floated parameter in the model, rather than being fixed. It is therefore useful to check the performance of the fit model in returning the input value of this parameter of 33.8 ± 8 years to see if this difference could be the cause of some of the observed issues. To do this, the best-fit value of the surface ^{210}Pb half-life for each crystal in each pseudo-experiment was extracted from its posterior distribution and plotted in a histogram for an ensemble of pseudo-data, as with the best-fit modulation amplitude described in Section 8.3.1. This is shown for C2 and C7 in the ensemble of pseudo-data with zero injected modulation amplitude in Fig. A.1. This result and indeed all the other results described from here on in this chapter are independent of the injected annual modulation amplitude.

As can be seen in the distributions of the best-fit values of the annual modulation amplitude in Fig. 8.3, the distribution of the best-fit values of a parameter in an ensemble of pseudo-data is expected to be a Gaussian with a mean equal to the input value used in generating that parameter in ensemble of pseudo-data. However, in the case of the best-fit value of the surface- ^{210}Pb half-life, the distribution is neither Gaussian or centred around the input value of 33.8 years.

In the case of C7, the distribution has a sharp, very quickly rising at some small value above zero to a peak at around 20 years before then dropping off roughly exponentially. Note that in all crystals, the bin with the most entries corresponds to the bin containing the standard half-life of ^{210}Pb of 22.3 years. A small fraction of the values are below zero. In C2 (shown) as well as the other three crystals not shown, the peak at low half-lives followed by a roughly exponential decrease is still observed, but rather than almost all of the best-fit surface ^{210}Pb half-lives being positive, a significant proportion (between a quarter and a third) of the values are negative, again with a peak slightly beyond zero followed by a roughly symmetrical decay to the positive values. In all cases the best-fit half-lives are very unlikely to be within ± 10 years of zero. Values close to zero being

disfavoured for the surface ^{210}Pb half-life in the fit model is an expected result as a half life of zero corresponds to an infinitely high decay constant (the value used in the fit model). It is therefore more favourable for the surface ^{210}Pb rate to be suppressed with a low initial activity, rather than a half-life close to zero.

The distributions of the best-fit value for the surface ^{210}Pb activity were also plotted. These distributions were well described by a Gaussian as expected, but with a negative bias of between 5% and 15%.

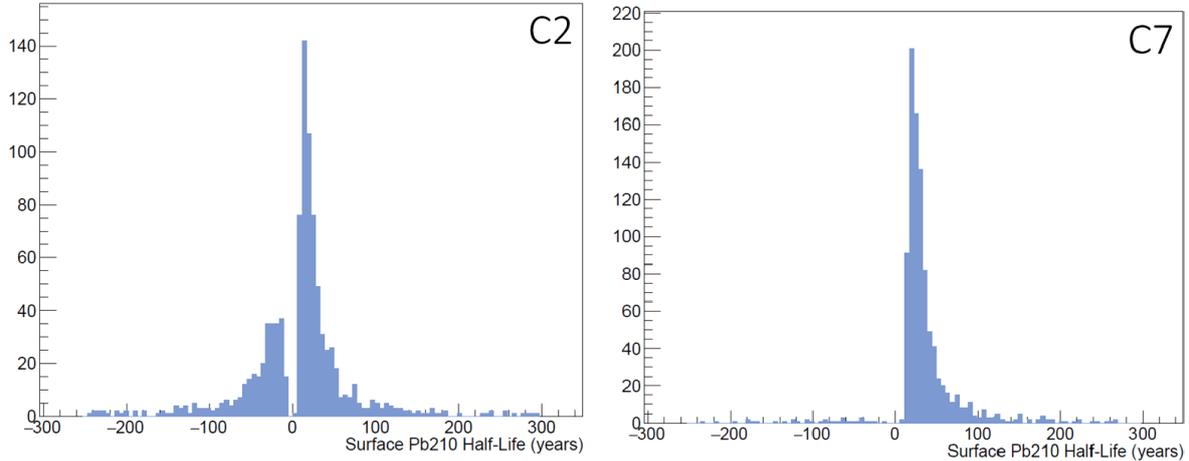


Figure A.1: Distributions of the best-fit values for the half-life of the surface ^{210}Pb components in two of the crystals in an ensemble of 1,000 pseudo-experiments with an injected modulation amplitude of zero. The distribution for C2 is typical of the distribution of the other crystals whereas the shape of C7's distribution is unique to C7.

Given the similar fraction of pseudo-experiments having higher than expected χ^2 values and negative values for the best-fit of the half-life of surface ^{210}Pb in four of the five crystals, the relationship between these two parameters should be tested to see if any causal link exists. This was done by plotting the mean of the five best-fit values of the surface ^{210}Pb half-life in each pseudo experiment against the χ^2 value of the pseudo-experiment (left plot in Fig. A.2), and the lowest of these five values against the χ^2 value (right plot in Fig. A.2). As can be seen, in both cases no meaningful relationship exists between the half-lives and χ^2 values so pseudo-experiments with a negative best-fit surface ^{210}Pb half-life are not the cause of the high χ^2 values. Furthermore, the plot showing the lowest value of the surface ^{210}Pb half-life reveals that the vast majority of pseudo-experiments have at least one negative surface ^{210}Pb half-life.

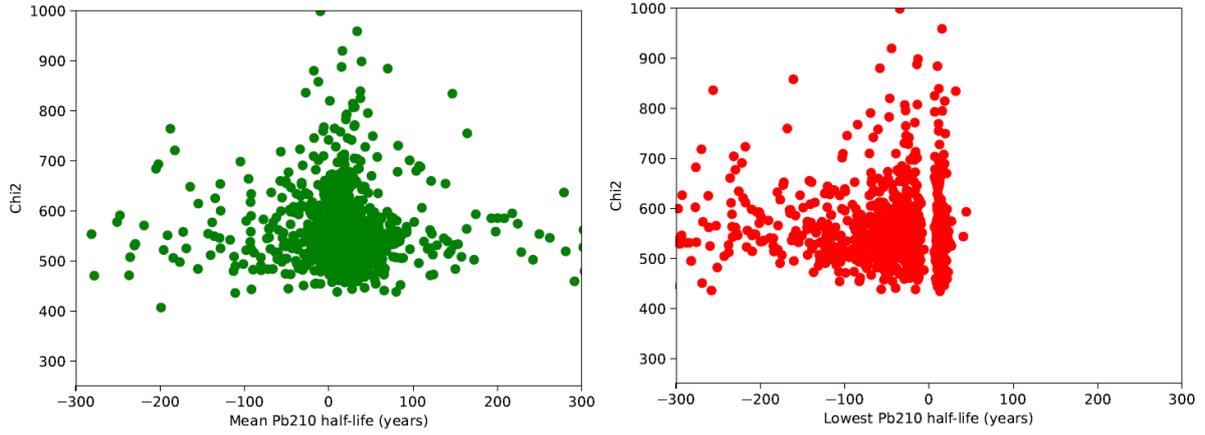


Figure A.2: Plots of the mean (left, green) and lowest (right, red) of the best-fit values of the surface ^{210}Pb half-lives across the five crystals in each pseudo-experiment in an ensemble of 1,000 pseudo-experiments against the experiment's χ^2 value for the overall fit. No relationship between these values is found

The fact there is seemingly no relationship between the negative best-fit surface ^{210}Pb half-lives and χ^2 values could potentially be explained if the best-fit activities for the surface ^{210}Pb were very low in these pseudo-experiments. This would suppress the surface ^{210}Pb component from the fit and minimise the destabilising effects of the negative surface ^{210}Pb half-lives. In this case the activity of the bulk ^{210}Pb component could equally be over-fitted to account for the lost activity from the surface component. However, when this hypothesis was tested by plotting the best fit-values for the activity and half-life of the surface ^{210}Pb (shown for C2 in Fig. A.3), no relationship was found between these values either. The input value used for the initial activity of the surface ^{210}Pb is 0.206 dru for C2.

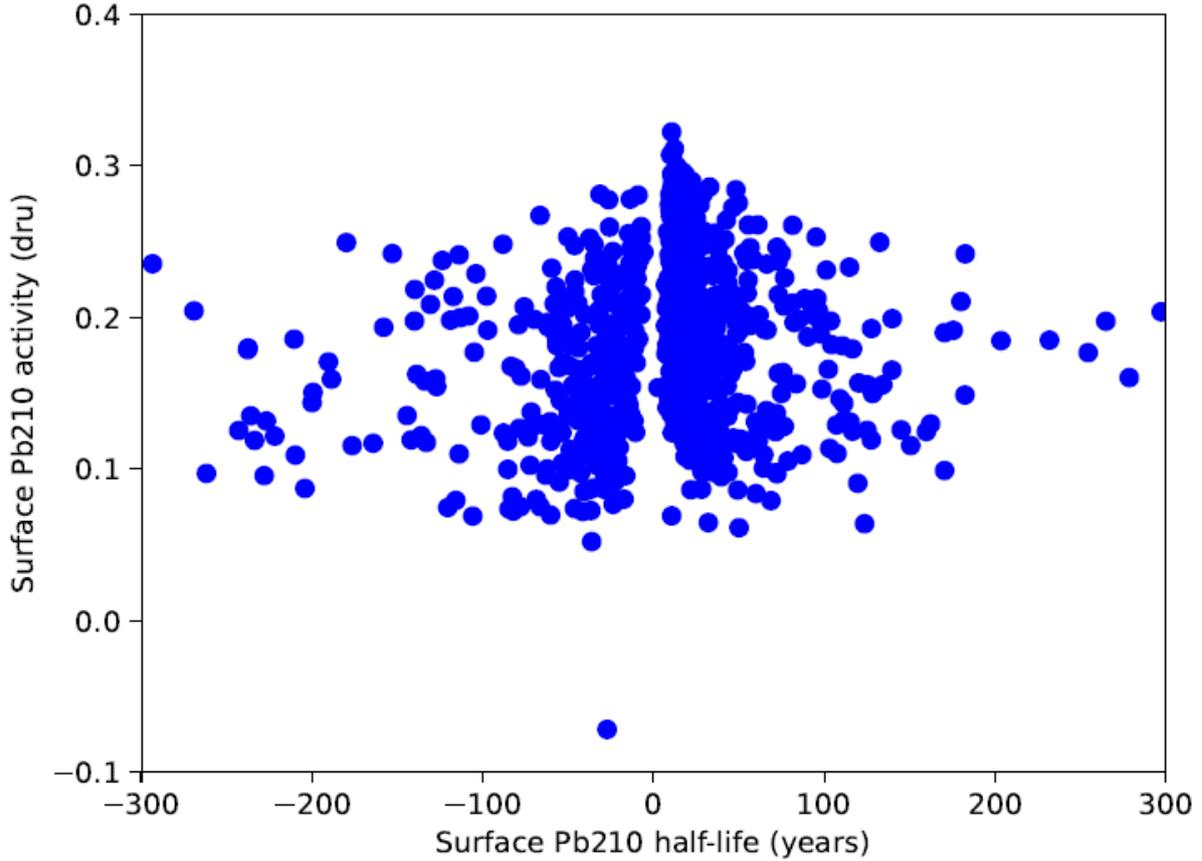


Figure A.3: The best-fit values for the half-lives and activities of the surface ^{210}Pb for C2 (shown as a representative example) in every pseudo-experiment in an ensemble of 1,000 pseudo-experiments.

As shown in Fig. A.3, it is possible for a given pseudo-experiment to have best fit values for the activity of surface ^{210}Pb that are around the expected value, while having half-lives that are either highly positive or negative, or close to zero. It is useful to understand the interpretations of these different values for the half-life of the surface ^{210}Pb . In the case where the half-life is large, be it positive or negative, the surface ^{210}Pb component is essentially modelled as flat over the roughly five year time scale described.

Smaller positive values (i.e. roughly less than 100 years) to the right of the main peak for the half-life represent the case where the surface ^{210}Pb decays as expected with an extended half-life. Positive values lower than this peak value (which corresponds to the standard half-life of ^{210}Pb) on the other hand are unexpected and likely unphysical as they correspond to the case where the surface ^{210}Pb is being removed more quickly than through decays alone. Finally small negative half-lives are even more unexpected and describe an increase in the amount of surface ^{210}Pb on the crystal. In the most extreme case this corresponds to roughly doubling the rate of the surface ^{210}Pb component. Since it is common for unphysical values for the surface ^{210}Pb half-life to be returned from the fitting procedure, in future the procedure should be updated to constrain the best-fit value of this half-life, for example disallowing values below the standard value of the half-life of ^{210}Pb . Unfortunately it was not possible to implement this alteration in this analysis.

A.1.1 Interpretations of these results

It is clear from the above findings that some issue exists with the fitting procedure in terms of correctly describing the background components, resulting from the splitting of the surface ²¹⁰Pb component in the fit model. In terms of locating and understanding the source of these issues for future analysis, it is only possible to speculate, but perhaps one source of issue arises from calculating the best-fit values after the fit has been applied to the data, rather than the fitting itself. While in theory the best-fit parameters described above are all calculated using the same method, that is from the means of the posterior distributions of each parameter, in practice there are procedural differences in how the best fit parameters are calculated in certain cases.

When drawing the fit function and hence calculating the χ^2 values (for example in Fig. 8.2), the best-fit values used to draw the fit function are produced automatically from the mean of the posterior distributions by the analysis toolkit used ¹ when the rate data is analysed and the posterior distributions are produced and saved. Conversely, in all other cases, the best-fit values are calculated in a separate step after the fitting procedure has been carried out by manually drawing the posterior distributions produced during the analysis and calculating the mean. While these approaches should produce the same results, if they do not, the divergence in the behaviour of the fit functions drawn and the best-fit values produced afterwards (i.e. the differences in Fig. 8.10 and Fig. 8.11) may be explained. A check should be carried out to ensure the correct best-fit values are returned in all cases.

It must also be noted though that given the χ^2 values for fits using a single ²¹⁰Pb component appear normal; it can be inferred that the fit function is drawn correctly in this case. Therefore it may be possible that the addition of the split ²¹⁰Pb component is somehow causing the automatic calculation of the best-fit values of the fit components to fail, and hence an incorrect fit function to be drawn and an incorrect χ^2 calculated. This requires further investigation.

Another potential source of issues when performing a fitting procedure such as this is if the model constructed using the prior distributions for the analysis do not accurately match the data being analysed, then the fit may perform poorly. In the splitting of the ²¹⁰Pb component these prior distributions are also updated and could therefore be a source of issues. However, this was checked for each crystal by drawing functions using the values for the prior distributions for both fit models and comparing the outputs. In this comparison, the values never differed by more than 6% (in C3), and by less than 3% in the other four crystals. These differences are expected to be too small to cause issues, except in the case where the fitting procedure is overly sensitive to the prior values.

Finally, it was also observed that unphysical values for the best-fit value of the surface ²¹⁰Pb are often obtained during the analysis. Possibly this is as a result of some other procedural issue, however, the pseudo-data analysis should also be repeated using a split-²¹⁰Pb fit model where the values of the surface ²¹⁰Pb half-life are constrained such that unphysical values may not be produced. This is in case some of the issues noted above are as a result of these unphysical half-lives.

¹An in-house toolkit for the analysis of COSINE-100 data